Signal Processing for Ultra-Wideband Systems

This thesis is submitted in partial fulfillment of the requirements for the degree of Master of Science in Electrical Engineering.

Submitted 28/5-2003 to the Technical University of Denmark by Lars Puggaard Bøgild Christensen

Abstract

The subject of this thesis is signal processing for ultra-wideband (UWB) systems designed for wireless communication. An introduction to UWB systems is given along with modulation and user separation. A UWB multipath propagation channel is described and advanced receivers capable of capturing the multipath energy spread out by the radio channel are introduced. A brief introduction to synchronization is given and a novel synchronization method is examined in both the AWGN and UWB multipath propagation channel. Furthermore, coexistence with narrowband systems using the same frequency band is investigated. An assessment of the feasibility of using UWB communication systems for different scenarios with respect to range, bitrate, number of users, and cost is also presented.

Keywords: Wireless Communication, WPAN, UWB, Modulation, CDMA, Time-Hopping, Direct-Sequence, PSD, RAKE, LMMSE, Synchronization, Coexistence, UWB Scenarios.

Resumé (Abstract in Danish)

Denne kandidatafhandling omhandler signalbehandling i ultra-wideband (UWB) systemer med fokus på trådløs kommunikation. Der gives en introduktion til UWB, samt hvordan modulation og brugerseparation foretages og en model for UWB udbredelseskanalen præsenteres. Herefter præsenteres flere avancerede modtager-strukturer, der er i stand til at opfange den energi, som er blevet spredt ud over tid af kanalen. Dernæst gives en kort introduktion til synkroniseringsproblemet, og en ny synkroniseringsmetode undersøges i både AWGN og UWB udbredelseskanalen. Yderligere studeres sameksistens med eksisterende smalbåndssystemer, som benytter samme frekvensbånd som UWB systemet. Til sidst gennemføres en analyse af UWB systemer benyttet i forskellige scenarier, baseret på faktorer som rækkevidde, bitrate, antal brugere samt omkostninger.

Stikord: Trådløs kommunikation, WPAN, UWB, modulation, CDMA, Time-Hopping, Direct-Sequence, PSD, RAKE, LMMSE, synkronisering, sameksistens, UWB scenarier.

Preface

This thesis serves as the final assignment in order to achieve the degree Master of Science. It has been carried out from 1. September 2002 to 28. May 2003 at Nokia Mobile Phones, Copenhagen and carries a workload of 50 ECTS points. The supervisors of the project are Søren Sennels and Carsten Juncker, Nokia and Jan Larsen, Informatics and Mathematical Modelling, DTU.

A CD is included with this thesis containing material relevant to the project. Besides this thesis, references used are included if possible. The Matlab source code used for simulations is also present along with all simulations and figures used in this thesis.

I want to thank the supervisors for guidance throughout the process of writing this thesis. A special thank to Thomas Fabricius for many interesting discussions on various topics of communication systems and signal processing. Furthermore, I would like to thank Thomas Fabricius, and Niklas Lindberg for taking the time to read this thesis and give suggestions for improvements.

Lyngby, 28. May 2003

_____

Lars P. B. Christensen

Table of contents

1 Introduction to Ultra-Wideband Systems

There has been an increasing interest in the area of Ultra-WideBand (UWB) systems over the past 5-10 years, but why is this so and what can UWB be used for? These are the questions that this chapter will try to answer.

UWB signals are generated in the time domain by transmitting very short pulses on the order of a fraction of a nanosecond to a few nanoseconds, giving an approximate bandwidth on the order of a few GHz. These signals are mainly used in three applications

- Communications
- Positioning
- Radar

If the UWB signal is used for communications, which is the main interest of this thesis, the pulses are modulated with information and the two most common ways of achieving this in UWB is Pulse Amplitude Modulation (PAM) and Pulse Position Modulation (PPM). If PAM is used then the amplitude of the pulse carries the information whereas in PPM it is done by the position of the pulse.

The positioning capability of UWB is made possible by a number of UWB nodes being able to observe the distance between each other by measuring the pulse delay from each other node. These distances then need to be communicated to a central node that can establish the positions of the individual nodes by triangulation. It is believed that the precision of such positioning will be on the order of a few centimeters, as the pulse duration is close to a nanosecond giving a spatial length of the pulses around 30cm. This is an interesting feature in conjunction with communications, as it makes more effective routing in the network possible and also enables position-based services.

Using UWB for radar applications, the transmitter and receiver are located at the same place and the distance to objects can be observed by the time delay measured by the reflected pulses. If the illuminated surface does not reflect nor absorb the pulses, it is even possible to see through the surface. The changes occurring in the pulse shape as seen by the receiver, can also give some information about the physical nature of the object, like what type of material the object is made of.

In Figure 1 an example of a common UWB pulse shape is shown modeled by the second derivative of a Gaussian pulse. This basic pulse shape is sometimes referred to as a monocycle originating from radar systems, as it has only one cycle pr. pulse.

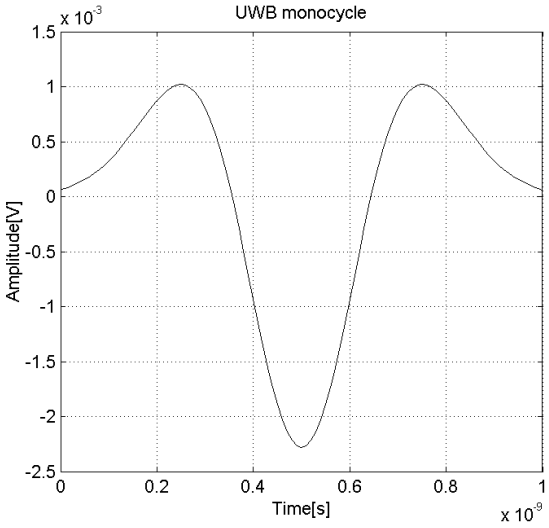

Figure 1: A typical UWB monocycle.

1.1 What is Ultra-Wideband?

Although the term UWB is commonly used to describe communications and radar systems, it is not really a term that refers to one specific usage. It is a term commonly used to describe that the relative bandwidth of the emitted signal is large, meaning that the bandwidth of the signal compared to its center frequency is large. The definitions of these terms are [1]

$$\eta = \frac{f_h - f_l}{f_c} = \frac{2\left(f_h - f_l\right)}{f_h + f_l} \quad (1.1)$$

where η is the relative bandwidth, f_h and f_l are the upper and lower frequency boundaries respectively, and f_c is the center frequency. In order for a signal to be defined as UWB, it must have at least a relative bandwidth of 0.25 and an absolute bandwidth of minimum 1.5GHz measured at the –10 dB emission point [1 §22].

An illustrative way of showing the design objective of UWB communication systems can be seen from the AWGN channel capacity as given by Shannon [2]

$$C = B\log_2\left(1 + SNR\right) \quad (1.2)$$

where C is the channel capacity in bits/s, B is the channel bandwidth in Hz, and SNR is the Signal-to-Noise Ratio.

In ordinary narrowband systems a given bandwidth is allocated to the service by a regulatory body, which guaranties that the allocated bandwidth is used only by this service. As frequency spectrum is a very scarce resource, the bandwidth will usually be selected as small as possible. The only parameter of the channel capacity that can be adjusted is therefore the SNR, which is therefore the normal design parameter that dictates the performance of the system. One obvious problem with this is the channel capacity only increasing by the logarithm of the SNR. This means that only a small gain can be achieved from improving the SNR, which is a big problem when high bitrate wireless connections are desired.

The UWB system is not band-limited as in the narrowband case, but instead power-limited. However, as the channel capacity increases proportional with the bandwidth, trading bandwidth for SNR is advantageous. Another benefit to this approach is that the SNR becomes so low that the channel capacity increases almost proportional with the SNR, making efforts into improving the SNR more beneficial than in NB systems. On the other hand doing so may require excessive coding to fulfill given Quality-of-Service (QoS) demands.

One thing in wireless UWB communications which is fundamentally different from ordinary wireless communications is the fact that UWB occupies a bandwidth of several GHz, which effectively means that it makes little sense to assign a specific frequency band to UWB services. UWB must therefore be able to co-exist in the same frequency spectrum as already allocated services without disturbing these and dealing with the interference from these services. This puts an upper limit on the emitted power of a UWB signal as well as its emitted spectrum.
The American Federal Communications Commission (FCC) has published a first report [1] on this subject where guidelines are given as to what can be expected from a regulatory point of view and it is expected that the European regulatory body will issue similar restrictions. The key limitations for wireless communications using UWB are

- Maximum average –41.3 dBm/MHz or 75nW/MHz Effective Isotropic Radiated Power (EIRP) in the frequency range 3.1GHz – 10.6GHz for indoor applications
- Even lower maximum EIRP for other frequency bands, especially for GPS bands
- Peak-to-Average Ratio (PAR) of maximum 20 dB

The total emitted power is therefore upper limited by 0.55mW and even this low transmit power is not realistic, as it would require the entire bandwidth of 7.5GHz.

The fact that UWB uses pulse widths on the order of nanoseconds give UWB the ability to deliver high bitrates on the order of several hundred Mbps, but at the expense of range. Another interesting feature is that UWB is not as prone to fading as ordinary narrowband communications and additionally the multipaths present in the radio channel, can be resolved down to a few nanoseconds in difference. Both of these properties are results of the pulsed nature of UWB. Another interesting feature is the inherent low power needed in the transmitter, as the output power is limited to a fraction of a milliwatt. This makes the usual Power Amplifier (PA) obsolete, as it is possible to drive the antenna directly from integrated transistors and thereby lowering the cost. The total power needed by a UWB system is therefore not severely limited by the transmitted power, which sometimes is the case with narrowband systems making UWB attractive for battery-powered equipment like mobile phones.
As all other technologies UWB also has drawbacks. Perhaps the most limiting factor is the power restriction that limits UWB operation to about 10m at around 100Mbps. Other UWB systems with either shorter distance and higher bitrate or longer distance and lower bitrate are of course also possible.

One of the key factors influencing the current interest in UWB systems is the fact that it has been troublesome to generate and modulate these short pulses up until now. Recent advances in semiconductor process technology [30] make it possible to integrate UWB pulse generators in a cost efficient manner and thus enable widespread use of UWB systems. However, acquisition and synchronization of UWB systems are still an open issue, as tracking the very short pulses with sufficient precision is very difficult. Another remaining question is the physical implementation, which is not without problems when dealing with such high bandwidths. It may be that the transmitter can be made easier and using less power than ordinary narrowband transmitters, but the receiver must be able to demodulate the signal in a reasonable way without using too much power and without costing too much. This could prove to be a challenge, as the signal has a bandwidth of several GHz.

1.2 Usage Scenarios for UWB Communications

In order to make UWB useful, it is needed to spot the possible applications that can benefit from this new approach to wireless communications. Having in mind the characteristics of UWB from last section different usage scenarios have been proposed for UWB [37]. One idea is to use UWB as an air-interface for new Wireless Personal Area Network (WPAN) standards that could be the next generation Bluetooth. Standardization work is currently being done in the IEEE 802.15.3a working group, which focuses on high speed WPAN solutions [30]. Other people suggest that UWB could be used as an air-interface for a Bluetooth version 3 [38], as a version 2 is already on the way and UWB is not yet a mature technology.
When looking towards the future of 3G/UMTS mobile networks, it is commonly believed that these networks will become integrated with new short-range high bitrate systems, like Wireless Local Area Network (WLAN) and WPAN to give the user seamless roaming between the fundamentally different systems. These WLAN and WPAN systems are likely to be operating in the unlicensed bands, mainly because dedicated spectrum is not available for this use and secondly because operators most likely will be reluctant to invest in new costly licenses.
As UMTS systems will continue to rely on older systems in areas with no coverage, GSM/GPRS/EDGE systems will be used for low to medium bitrates in rural areas and when a high degree of mobility is required. The UMTS system is used when medium to high bitrates are required in urban or suburban areas with a medium degree of mobility. Finally high to very high bitrate systems can be used if the terminal is within range of a high-speed WLAN or WPAN access point and having a low degree of mobility. The basic concept of this system integration can be seen in Figure 2.
This is where UWB might play a significant role giving unlicensed hotspot coverage at a high bitrate. It is also possible to achieve effective position-aware routing between UWB terminals in such a network, if the position of the terminals are known, making high bitrate UWB possible even if a UWB access point is not within range by using other terminals to route the traffic to the nearest access point.

Using UWB for medium to long-range low bitrate systems is also an interesting possibility, especially when exploiting the positioning capability of UWB. Such a system could for instance be used for remote sensors that can communicate and has knowledge of their location.

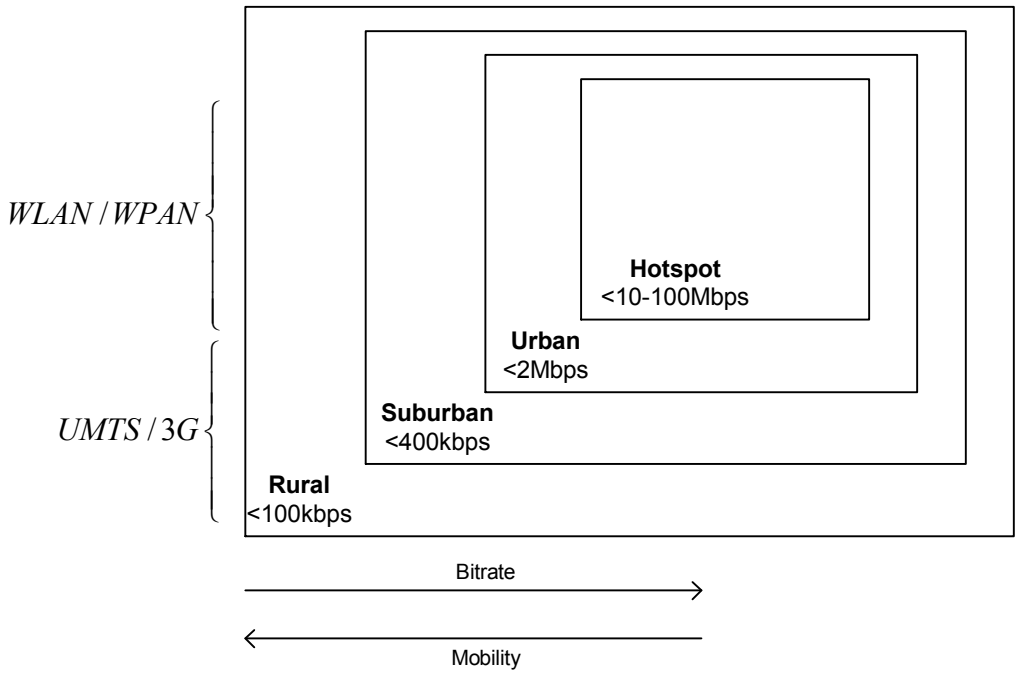

Figure 2: System concept of extended 3G.

Other uses include wireless audio and video distribution within the home. This would make it possible to replace all the cables used today between DVD/CD/MP3 players, amplifiers, TVs and other entertainment equipment. This could prove to be an interesting feature as these devices are expected to become fully digital within the next 10 years.
Another somewhat different use is called pervasive computing and is based on the vision of everything being connected to each other. This could include such diverse things as PDAs, mobile phones, TVs, refrigerators, wearable computers, cameras and all kinds of other sensors all being connected together and knowing the position of each other. This opens up a whole new range of services that can be offered, such as remote control of home appliances from mobile devices and security systems knowing who you are and opening the door for you.

1.3 The Relevance of Signal Processing

The use of signal processing is an important part of nearly all communication systems used today [11]. Future communication systems will increasingly rely on signal processing in order to push performance ever closer to fundamental bounds such as the channel capacity. This gradual increase in performance is necessary to fulfill the users expectations and is therefore a required competitive edge over competitors in the market. Efficient signal processing is thus one of the key factors determining the success of a communication system.

In the case of UWB systems this is also true. The area of signal processing for UWB systems is still being actively researched, making it an interesting and hot topic [5, 22]. One of the interesting facts of the UWB system is that it uses no carrier frequency, and the signal is therefore purely baseband in nature. This makes it possible to eliminate traditional components such as the mixer used to down-convert the signal before sampling. In turn the signal processing methods used, become ever more critical to system performance.

The aim of this thesis is therefore to present and analyze UWB communication systems from a signal processing point of view, with the purpose of using the technology for wireless short-range communications involving high bitrate links.

1.4 Comparison with other Communication Systems

As the basic idea behind UWB is now known it is time to compare UWB to its narrowband competition. The main advantage of using a narrowband system is that it is a well-known technology that has been used for many decades and that emission levels are not as strict as in the UWB case, as a dedicated spectrum is available. This raises the question: When is a UWB system preferable compared to a narrowband system?

In order to answer this question it is interesting to examine the theoretical AWGN channel capacity of realistic systems that could be used to provide a given service, such as WLAN or WPAN services. In Figure 3 the channel capacity of four different systems are compared: a UWB system, a IEEE802.11a/HIPERLAN2 WLAN, a IEEE802.11b WLAN and finally a Bluetooth WPAN system. The UWB system is operating in the 3.1GHz-6GHz band and is compliant with the FCC rules and the IEEE802.11a/HIPERLAN2 WLAN operates in the 5GHz Industrial Scientific Medical (ISM) band with a channel bandwidth of 20MHz emitting 200mW. The IEEE802.11b, which is the most common WLAN system today, operates in the 2.4GHz ISM band with a bandwidth of 22MHz and transmitting a maximum of 100mW. Finally, the Bluetooth system also operates in the 2.4GHz ISM band with a channel bandwidth of 1MHz emitting 1mW. The channel model includes free-space propagation loss, which is realistic in the indoor short range Line Of Sight(LOS) channel [16], and thermal noise at a room temperature set to 290K.

As it can be seen from Figure 3 the UWB system has a large advantage in capacity compared with the narrowband systems as long as the distance is on the order of tens of meters, but that it rapidly decreases as the distance increases. This clearly shows the potential of UWB when very high-speed short-range communications is the goal.

If the demands put on the bitrate by the service can be relaxed somewhat, then IEEE802.11a/ HIPERLAN2 or IEEE802.11b might be a strong candidate giving longer range at the expense of a higher power consumption and cost. On the other hand, if the need for a high bitrate is not required by the service, but power usage and cost is crucial to the application, then Bluetooth might be a better choice.

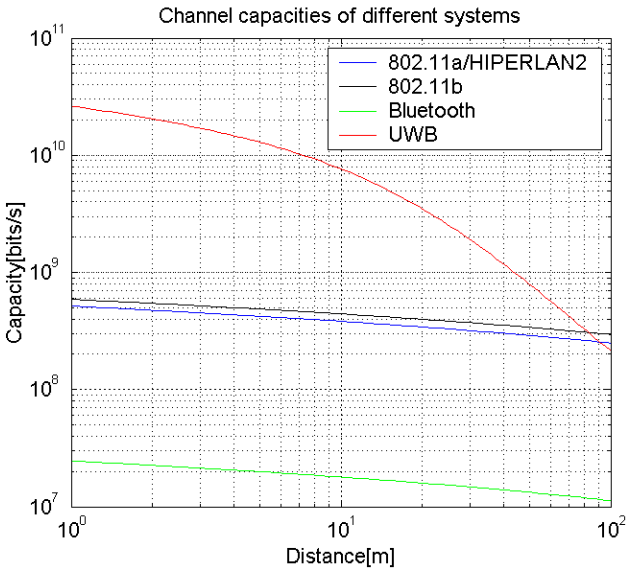

Figure 3: Channel capacity of UWB vs. IEEE802.11a/HIPERLAN2 vs. IEEE802.11b vs. Bluetooth.

1.5 Summary

In this chapter the fundamental idea of UWB systems have been introduced and the communications, positioning and radar uses have been introduced. Different applications, such as WLAN/WPAN and audio/video scenarios, have been outlined and the importance of signal processing for these systems has been stressed. It has further been shown how the UWB channel capacity is an order of magnitude better than competing narrowband systems when short-range communications is the focus. The rest of this thesis will deal mainly with signal processing for UWB systems designed for communications and will emphasize the short-range high bitrate scenario, as this is an area in which UWB has great potential.

2 Basic System Modelling

In this chapter some of the parameters affecting performance of a basic single-user UWB communication system are evaluated. This is required in order to get a better understanding of the tradeoffs needed in designing UWB communications systems.

2.1 Transceiver Model

In order to understand the elements of a UWB communications system, a block diagram is shown in Figure 4 showing the transceiver of each user and how up to K different users can share the UWB spectrum.

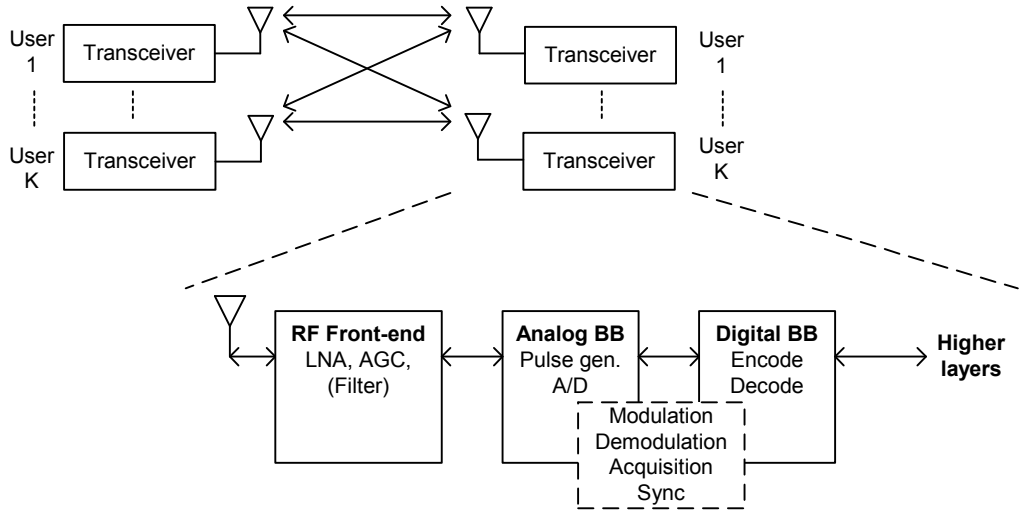

Figure 4: The UWB transceiver blocks and their function.

The typical blocks that a UWB transceiver consists of are shown in the lower part of the figure and to the left the antenna is seen. Its sole purpose is to convert the pulses to and from electromagnetic waves making wireless communication possible. Next to the antenna the RF front-end is shown. This block is responsible for conditioning the signal so that the baseband block can interface with the antenna. This includes the Low Noise Amplifier (LNA) and the Automatic Gain Control (AGC) in the receiver and, if needed, performing filtering of the transmitted pulses in the transmitter.
Next is the analog baseband block used in the transmitter to generate the pulses and do the analog-to-digital conversion needed in the receiver. The reason for having the dotted box between the analog and the digital domain is that it is not quite clear yet how they should interact and how the signal processing should be divided between the two domains. An example of this question is the correlators often used for demodulation. Should one choose to implement them using analog or digital signal processing? This question is outside the scope of this thesis. Another example is the acquisition and synchronization needed for the received signal.
Finally, encoding and decoding is done in the digital domain and is performed before the higher layers are reached.

2.2 Monocycle Model

The basic pulse shape used in UWB is sometimes referred to as a monocycle as it has only one cycle and it is often modeled as a derivative of a Gaussian pulse. The reason for using Gaussian pulses is that the techniques used to produce the narrow pulses yield pulses that are nearly derivatives of a Gaussian pulse in shape. Examples of such techniques are switching a carrier on and off very fast and feeding pulses to a differential mixer [12, 13].

Another source of differentiation are the antennas used in the receiver and transmitter. The reason is that an antenna reacts differently to an ultra-wideband signal than to a narrowband signal and it can be seen from Maxwell's equations that an antenna exposed to an ultra-wideband signal will perform a differentiation [7 p.33-34]. The radiated E-field will therefore be proportional to the derivative of the imposed pulse and the received pulse will be proportional to the second derivative of the pulse imposed on the transmitting antenna. These effects must be taken into account in order to perform effective demodulation of the information.

The monocycle $p_x(t)$ is therefore assumed modeled as the x^{th} derivative of a Gaussian pulse $p(t)$

$$p(t) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(\frac{-t^2}{2\sigma^2}\right) \quad (2.1)$$

$$p_x(t) = \frac{d^x p(t)}{dt^x} \quad (2.2)$$

where σ^2 is the variance of the Gaussian pulse. The Power Spectral Density (PSD) of the Gaussian pulse is then given by

$$\Sigma_p(f) = \left|P(f)\right|^2 = \left|\int_{-\infty}^{\infty} p(t)\exp(-j2\pi ft)\,dt\right|^2 = \exp\left(-(2\pi\sigma f)^2\right) \quad (2.3)$$

with $P(f)$ being the Fourier transform of $p(t)$. The PSD of the monocycle $p_x(t)$ can then be found as

$$\Sigma_{p_x}(f) = (2\pi f)^{2x}\,\Sigma_p(f) \quad (2.4)$$

The result given in (2.4) shows how performing derivations of the basic pulse shape can control the PSD of the monocycle, which in this case is assumed modeled as a Gaussian pulse.

Another issue that must be addressed is the pulse width, as an ideal Gaussian pulse has infinite duration. This is of course not practical and a limit to the effective pulse width must therefore be agreed upon. A reasonable way to define the pulse width is to choose the pulse width so that a given percentage of the power in the infinite duration pulse is captured. Throughout the rest of this thesis the pulse width $PW = k_w'\sigma$ will be defined as

$$\frac{\int_{-\frac{k_w'\sigma}{2}}^{\frac{k_w'\sigma}{2}} p_x^{\,2}(t)\,dt}{\int_{-\infty}^{\infty} p_x^{\,2}(t)\,dt} \geq 99.9\% \quad (2.5)$$

where k_w' is the smallest integer satisfying (2.5). The calculation of k_w' is done in Appendix A and an example of typical values of k_w' is $k_w' = 7$ in the case of $x = 2$.

In Figure 5 the monocycles $p_x(t)$ for $x = 0...4$ with $PW = 1ns$ are shown along with their respective PSDs and it can be seen, how the spectrum of the monocycle can be shifted upwards by using higher order derivatives. The ripple shown in the stopbands of the PSDs are the result of the rectangular window applied to truncate the pulse width to $PW = k_w'\sigma$. If a given application specifies another stopband criteria, the limit of 99.9% power capture set in this thesis may be altered to fit these demands.

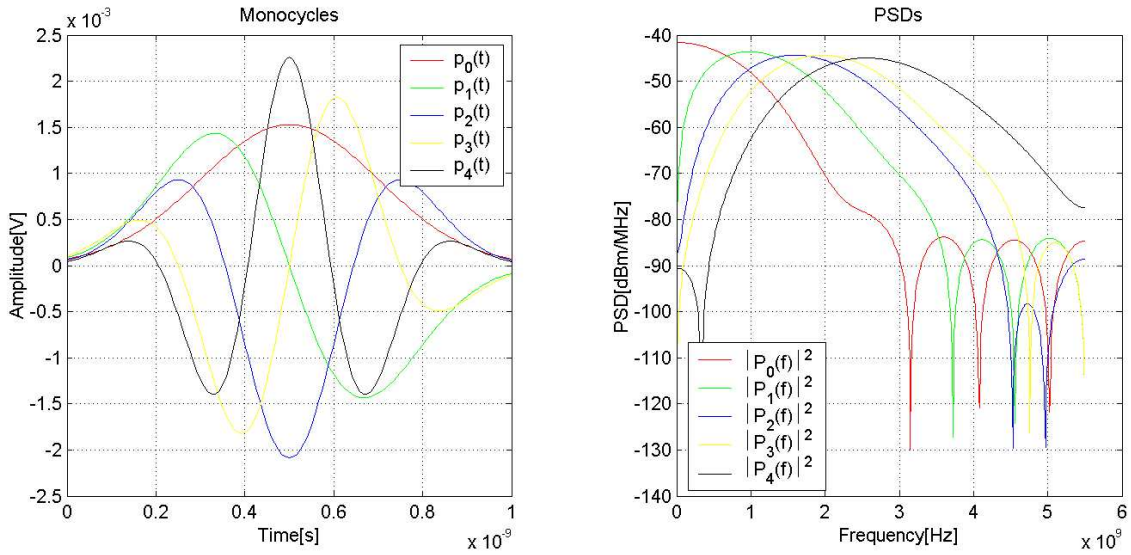

Figure 5: Monocycles $p_x(t)$ for x=0..4 with PW=1ns and their PSDs.

Now that the pulse width is known, the number of samples pr. pulse can be found. This is important in order to represent the signal digitally as used in digital signal processing and simulations. The calculations are done in Appendix B, showing that $p_2(t)$, often used in the literature for simulations [4, 5], can be represented by 10 samples pr. pulse, if 50dB attenuation between maximum of the PSD and the PSD at half the samplerate is sufficient.

2.3 Modulation Schemes

In order to make the UWB signal carry any information bits the signal must be modulated with the information. This can be done in several ways, but because UWB is radiated as a baseband signal it is not possible to use standard quadrature modulation schemes as would otherwise normally be preferred, because this would require a complex-valued baseband signal. It is also not possible to use any kind of frequency modulation because of the pulsed nature of UWB.
This leaves the possibility of using either M-ary Pulse Amplitude Modulation (M-PAM) or, because of the pulsed nature of UWB, M-ary Pulse Position Modulation (M-PPM). When using M-PAM to modulate the signal, transmitting pulses with M different amplitudes carries $\log_2(M)$ information bits, whereas in M-PPM shifting a pulse to M different positions in time carries the same amount of information. In M-PPM it is possible to place the M different pulse positions so that they will overlap. This makes it possible to trade off channel capacity for increased sensitivity to noise as placing the symbols ever close will decrease the difference between two decisions. The optimal overlap is therefore dependent on the SNR and can be shown to approach the minimum of the monocycle's autocorrelation function when the SNR goes to zero [3]. This thesis will therefore only consider non-overlapping M-PPM and 2-PPM with overlap given by the minimum of the autocorrelation function.
Another interesting possibility is to use both M-PAM and M-PPM at the same time when possible. This is made possible by the fact that the pulsed nature of UWB under the right circumstances makes the two modulation types orthogonal to each other. This special type of modulation will be named M-ary Pulse Position and Amplitude Modulation (M-PPAM) in this thesis, as a recognized name for this modulation has not been found. Further an approximation to the BER of this modulation is found in Appendix C.4.

As stated before UWB is a baseband system that does not use a carrier frequency. This is, however, a truth with modifications. UWB systems have recently been proposed to use carrier frequencies to shift a monocycle, with a bandwidth on the order of 500MHz, up in frequency. Several of these carrier frequencies are then used to form a UWB signal. Some of the advantages of this approach are discussed in chapter 3, but the effect of this on the modulation scheme is that it becomes possible to use quadrature modulation. This approach is, however, not pursued further in this thesis, as the multiband solution had

little backing when this thesis work was started. The modulation schemes studied are therefore restricted to the ones mentioned above.

In ordinary narrowband communications the selection of the parameter M in a M-ary modulation scheme is rather complicated depending on channel parameters such as the SNR. This give rise to adaptive modulation schemes in which M can be adjusted as needed in order to maximize performance. An example of such a system is the dial-up modems used today that during the connection establishment phase determines the maximum value of M that can be used for reliable communications.
Another way to adapt the system to the channel is to adjust the spreading factor of a CDMA system. This is made possible by the fact that the SNR after despreading increases proportionally with the spreading factor whereas the bitrate is inversely proportional to the spreading factor. This approach is often better because of the logarithmic relation between SNR and channel capacity at a given BER dictated by Shannon's law. The reason for assuming that the process of letting multiple users communicate simultaneously is done, at least to some extend, by Code Division Multiple-Access (CDMA) in UWB will be discussed in chapter 3.

The optimal choice of M in a UWB system can be readily determined by looking at the channel capacity pr. unit bandwidth over an AWGN channel given by Shannon [2]

$$\overline{C} = \frac{C}{B} = \log_2\left(1 + SNR\right) \qquad (2.6)$$

This determines the highest possible channel capacity that can be achieved by any modulation scheme and is therefore a natural benchmark. To visualize how the channel capacity of the different modulation schemes compares to the Shannon limit, a graph showing the relationship is presented in Figure 6. The derivations of the channel capacities are done in Appendix D.

In order to determine the gain in channel capacity experienced when increasing the SNR, the derivative of (2.6) can be observed

$$\frac{d\overline{C}}{dSNR} = \frac{1}{\ln\left(2\right)\left(1 + SNR\right)} \qquad (2.7)$$

On the other hand adjusting the spreading factor so that when the SNR before despreading changes, the SNR after despreading remains constant, results in a slope of unity. The reason is that every time the SNR before despreading is doubled, the bitrate also doubles because the SNR after despreading is to remain constant. This basically means that it is possible to scale the bitrate linearly with the SNR by adjusting the spreading factor. It is now interesting to compare this slope of unity to the maximum slope of the channel capacity

$$\arg\max_{SNR}\left(\frac{d\overline{C}}{dSNR}\right) = \frac{1}{\ln\left(2\right)\left(1 + 0\right)} \simeq 1,44 \qquad (2.8)$$

As it can be seen from (2.8) the slope is larger than unity when the SNR is very low and becomes unity when the SNR becomes

$$\frac{d\overline{C}}{dSNR} = 1 \Rightarrow SNR = \frac{1}{\ln\left(2\right)} - 1 \simeq 0,44 \qquad (2.9)$$

In a CDMA system it is therefore only in the region $0 \le SNR \le 0,44$ that using a higher order modulation scheme, that is increasing M, pays off and in the case of such low SNRs it is impractical to do so, as excessive coding will be required to give an acceptable QoS.

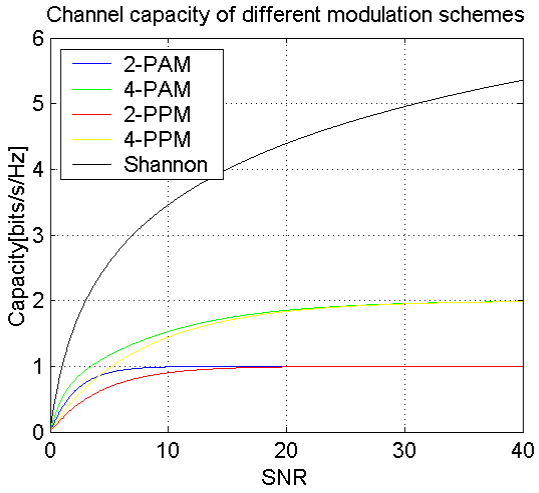

Figure 6: Comparison of channel capacity for 2- and 4-level PAM and non-overlapping PPM vs. the Shannon limit.

The result is that it is not advantageous to use higher-order modulation schemes in CDMA systems as a higher throughput can be achieved by adjusting the spreading factor appropriately instead. On the other hand adjusting the spreading factor in this manner may not be desirable, as a low spreading factor could potentially introduce more Inter Symbol Interference (ISI) in a multi-path channel, but it is shown in [6] that this is not the case if a proper receiver design is used such as a RAKE receiver.
Situations exist where maximizing the bitrate by adjusting the spreading factor is not desired or possible. An example of this is short-range communications where the SNR is very large and the optimal spreading-factor would therefore be unity, but this is not possible if multiple users must be active. Even if the spreading-factor was lowered to unity the SNR might still be large enough to support a higher-order modulation scheme.
Adjusting the spreading-factor should therefore be considered a first choice and using a higher-order modulation scheme should only be considered as the second best way to increase the bitrate.
This is the reason why the rest of this section on modulation only deals with 2-PAM and 2-PPM or BPSK and BPPM respectively, as they will be named from now on. The combination of these two binary modulation schemes named Quaternary PPAM (QPPAM) will also be considered.

2.4 Signal Model

A signal model will now be presented that will be used as a basis throughout the rest of this thesis. Although ways of letting more than one UWB system communicate at a time has not yet been discussed (see Chapter 3 for more information) some understanding of the possible multiple-access strategies are needed in order to set up a signal model. In UWB there are generally two different ways of doing this

- Time-Hopping
- Direct-Sequence

Time-Hopping (TH) is based on time separation of the different users by assigning each user with a unique TH code, which determines the time-slots in which the user is allowed to transmit.
Direct-Sequence (DS) separates the different users by assigning a unique spreading code to each user, which is used to expand the bandwidth of the signal. In the receiver the signal is despread using the same spreading code and thereby suppressing the other users.
Both TH and DS are effectively separating the users by unique codes and can therefore be considered a CDMA system. The term Spread Spectrum (SS) is sometimes also used for CDMA systems, as the signal bandwidth used by a single-user CDMA system is larger than necessary.

The signal model that will be presented now consists of two different, but closely related, models as both TH and DS CDMA systems will be examined. First the TH signal model will be presented and the DS signal model will then be given as an extension to this model.

2.4.1 Time-Hopping Model

A UWB signal using TH can be modeled as:

$$TH: \quad s^{(k)}(t) = \sum_{j=0}^{\infty} \alpha^{(k)}\left(\left\lfloor \frac{j}{N_s} \right\rfloor\right) \cdot p_{tx}\left(t - jT_f - c_{TH}^{(k)}(j)T_c - \beta^{(k)}\left(\left\lfloor \frac{j}{N_s} \right\rfloor\right) - \tau^{(k)}\right) \quad (2.10)$$

where $s^{(k)}(t)$ is the resultant signal coming from user k, $p_{tx}(t)$ is the basic transmitted pulse shape (monocycle) located from 0 to T_{mono}, which is assumed to be modeled by $p_x(t)$. The frame time T_f is the time between the uniform spaced pulses in the signal $\sum_j p_{tx}\left(t - jT_f\right)$. The time hopping code $c_{TH}^{(k)}(j)$ is an integer used to perform pseudo-random time hops by the k^{th} user at the j^{th} pulse and T_c is the minimum resolvable time delay that the hopping code can address. Finally N_s is the number of pulses pr. information bit and $\alpha^{(k)}\left(\left\lfloor \frac{j}{N_s} \right\rfloor\right)$ and $\beta^{(k)}\left(\left\lfloor \frac{j}{N_s} \right\rfloor\right)$ are the variables carrying the information from user k. If users communicate asynchronously the timing offset $\tau^{(k)}$ of the k^{th} user may be used to account for this time difference.

The reason for having two variables carrying information is that both TH-BPSK, TH-BPPM and TH-QPPAM should be supported by this model and the value of these variables are given as

$$\begin{aligned}
&TH-BPSK: && \alpha^{(k)}(n) = 2d^{(k)}(n) - 1 \quad \wedge \quad \beta^{(k)}(n) = 0 \\
&TH-BPPM: && \alpha^{(k)}(n) = 1 \quad \wedge \quad \beta^{(k)}(n) = \delta \cdot d^{(k)}(n) \\
&TH-QPPAM: && \alpha^{(k)}(n) = 2d^{(k)}(2n) - 1 \quad \wedge \quad \beta^{(k)}(n) = \delta \cdot d^{(k)}(2n-1)
\end{aligned} \quad (2.11)$$

where δ is the modulation index of the BPPM signal determining how much the pulse should be offset by the information and $d^{(k)}(n) \in \{0,1\}$ being the n^{th} information bit from the k^{th} user.

The time hopping code $c_{TH}^{(k)}(j)$ has two purposes in the TH signal. First of all it must provide pseudo-random hopping of the pulse positions from the k^{th} user in order to minimize the chance of collisions with other transmitters within range. It thus provides multiple-access capabilities when different users have different hopping codes.

Secondly, as the signal without hopping and modulation is periodic with the time T_f, spectral lines will appear in the Power Spectral Density (PSD) of the signal spaced $1/T_f$ apart in the frequency domain. When a time hopping code with period N_p is applied the spacing becomes $1/(N_p T_f)$, which is much less when N_p is large. This has the effect of spreading out the power more evenly, which is very desirable in order to meet PSD limitations set out by the regulatory body.

Another important thing to understand is the valid range in which the code $c_{TH}^{(k)}(j)$ is allowed to operate. As the maximum delay possible is determined by the frame time T_f then the maximum value of $c_{TH}^{(k)}(j)$ named N_h is bound by $0 \le N_h T_c < T_f$. The reason that equality is not an option in the upper bound, is that there will always be a finite delay from the point when the correlator has finished collecting pulse energy from one pulse until it is ready to collect energy from the next pulse.

One parameter often used in the description of a CDMA system is the spreading-factor, which determines the bandwidth expansion experienced by applying the code to the signal. In the TH signal model the spreading-factor SF is therefore given by

$$TH: \quad SF = \frac{1}{\varepsilon} = \frac{T_f}{T_{mono}} \quad (2.12)$$

where T_{mono} is the duration of the monocycle. The duty-cycle ε is a direct measure of how large a percentage of the time the transmitter is active and this also indicates that the Signal-to-Noise Ratio after despreading SNR' is given by

$$SNR' = SF \cdot SNR \quad (2.13)$$

where SNR is the average Signal-to-Noise Ratio on the physical channel termed channel SNR. The Signal-to-Noise Ratio after despreading will be named detection SNR, as this is the actual SNR at the decision device at the time of the decision.

For later use with matched filtering the basic waveform template used by the k^{th} user is termed $\varphi^{(k)}(t)$ and in the case of TH it is therefore given by

$$\varphi^{(k)}(t) = p_{tx}(t) \quad (2.14)$$

and it is assumed normalized so that

$$\int_0^{T_f} \varphi^{(k)}(t) \cdot \varphi^{(k)}(t) dt = SF \quad (2.15)$$

in order to have a channel SNR of

$$SNR = \frac{1}{\sigma^2} \quad (2.16)$$

where σ^2 is the noise variance of the Gaussian noise.

2.4.2 Direct-Sequence Model

Using the notation from the TH model a DS UWB signal can be modeled as

$$\begin{aligned} DS: \quad s^{(k)}(t) = \sum_{j=0}^{\infty} \alpha^{(k)}\left(\left\lfloor \frac{j}{N_s} \right\rfloor\right) \cdot \varphi^{(k)}\left(t - jT_f - \tau^{(k)}\right) = \\ \sum_{j=0}^{\infty} \alpha^{(k)}\left(\left\lfloor \frac{j}{N_s} \right\rfloor\right) \sum_{i=0}^{N_f-1} c_{DS}^{(k)}(i)\, p_{tx}\left(t - jT_f - iT_{mono} - \tau^{(k)}\right) \end{aligned} \quad (2.17)$$

where N_f is the number of pulses transmitted in the frame time T_f and $c_{DS}^{(k)}(i) \in \{-1, +1\}$ is the binary spreading code. All other variables are the same as in the case of the TH model. The pulse $\varphi^{(k)}(t)$ may be viewed as being a basic waveform template transmitted in every frame only multiplied by the information like in a TH-BPSK system. As transmission is continuous the possibility of positioning pulses in time as needed by PPM is not available and only BPSK using $N_s = 1$ is therefore applicable. The modulation variable is then given by

$$DS - BPSK: \quad \alpha^{(k)}(n) = 2d^{(k)}(n) - 1 \quad (2.18)$$

The reason for introducing the inner summation in (2.17) is that when using a DS system, transmission of pulses is continuous and thus N_f pulses are transmitted for every T_f with each of these commonly known as chips. This also means that the spreading-factor is given by

$$DS: \quad SF = N_f = \frac{T_f}{T_{mono}} \quad (2.19)$$

and as in the case of TH the detection SNR is given by

$$SNR' = SF \cdot SNR \quad (2.20)$$

2.5 Performance of Modulation Schemes in the AWGN Channel

It is shown in Appendix C that the correlation receiver is optimal in the memory-less AWGN channel and in order to achieve the optimal detector, M correlations with the orthonormal basis functions of the waveforms $s_m(t)$ must be performed, where M is the dimension of the signal and $m = 1...M$. A hard-decision can then be performed in order to determine the value of m that minimizes the Euclidian distance between the received waveform and $s_m(t)$.

In the case of BPSK and BPPM only a single correlator is needed and the estimate of the information bit is given by

$$\hat{d}^{(k)}(n) = \frac{\operatorname{sgn}\left(\int_{nN_sT_f}^{(n+1)N_sT_f} s^{(k)}(t) \cdot s_{corr}(t - nN_sT_f)dt \right) + 1}{2} \quad (2.21)$$

where $s_{corr}(t)$ is the template waveform. If the modulation is BPSK then $s_{corr}(t) = p_{rx}(t)$, where $p_{rx}(t)$ is the received monocycle and the BER can readily be found by observing the signal- and noise power after the correlator (see Appendix C.2). When including the fact that the SNR after despreading is increased by a factor equal to the spreading factor the BER becomes

$$BER_{BPSK} = \frac{1}{2} erfc\left(\sqrt{\frac{SF \cdot N_s \cdot SNR}{2}} \right) \quad (2.22)$$

If BPPM modulation is used then $s_{corr}(t) = p_{rx}(t - \delta) - p_{rx}(t)$ and the BER is then given by (see Appendix C.3)

$$BER_{BPPM} = \frac{1}{2} erfc\left(\sqrt{\frac{SF \cdot N_s (1 - \rho(\delta)) \cdot SNR}{4}} \right) \quad (2.23)$$

where

$$\rho(\delta) = \frac{\int_{-\infty}^{\infty} p_{rx}(t)\, p_{rx}(t - \delta)\, dt}{\int_{-\infty}^{\infty} p_{rx}(t)\, p_{rx}(t)\, dt} \quad (2.24)$$

is the normalized autocorrelation function of the monocycle at lag δ.

It can be seen from (2.22) and (2.23) that BPSK has a 3dB advantage over BPPM when $\rho(\delta) = 0$. This should come as no surprise as the effective integration period is twice as long in non-overlapping BPPM as in BPSK and therefore twice the noise power is collected with the same signal power. If overlapping BPPM is used the advantage of BPSK becomes smaller, but it will never reach zero.

When using QPPAM two correlators are needed. One of the correlators has $p_{rx}(t - \delta)$ as template and the other has $p_{rx}(t)$. Observing the numerical value of the correlators and selecting the largest performs demodulation of the bit modulated by BPPM and the sign of the selected correlation determines the bit modulated using BPSK. The calculations needed to find the BER are done in Appendix C.4, showing that

it is not possible to express the BER analytically. Instead an approximation to the BER is derived in the appendix as

$$BER_{QPPAM} = \frac{3}{2} BER_{BPPM} + \left(\frac{1}{2} - BER_{BPPM} \right) BER_{BPSK} \quad (2.25)$$

under the assumption that $\rho(\delta) = 0$ and $SNR >> 1$.

Simulations of these different modulation schemes have been performed as shown in Figure 7 in order to verify the analytical results presented above. Both the TH and DS system uses a 7th derivative of a Gaussian pulse with PW=2/3ns as the received monocycle in order to place the PSD of the free-space signal in the frequency band allocated by the FCC for UWB operation (see Appendix 3.3). Secondly, the selection of the pulse width makes it possible to deliver approximately 100Mbps raw bitrate using a spreading factor of around 15, making it possible to add more users to the system later. The spreading-factor of the TH and DS systems are not the same and is fixed at $SF = 16$ in the case of TH and $SF = 15$ in the case of DS. The reasoning behind selecting these parameters are first outlined in Chapter 3 and are therefore not discussed further at present time. The TH system uses one pulse pr. bit i.e. $N_s = 1$.

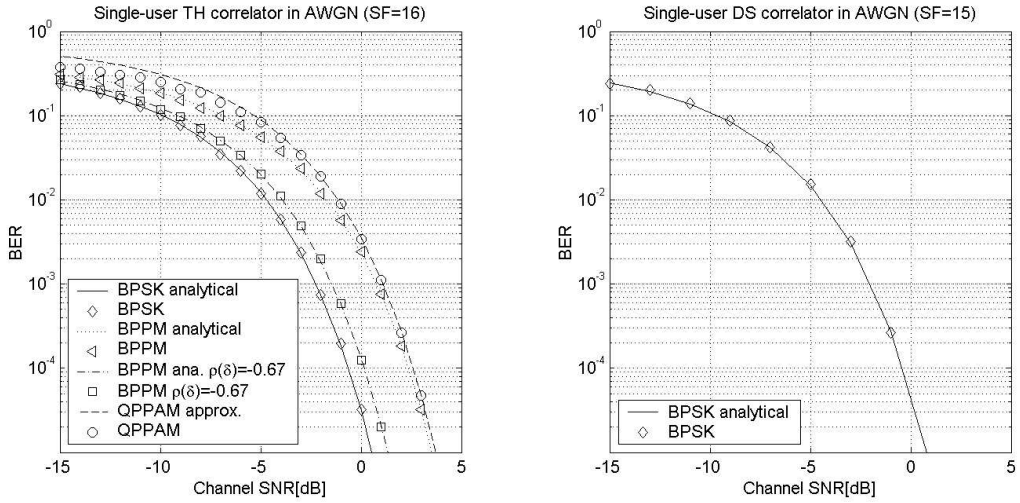

Figure 7: Simulation of single-user UWB systems in AWGN channel.

The reason for using $\rho(\delta) = -0.67$ in the case of overlapping BPPM is that the corresponding value of the overlap δ is the minimum of the normalized autocorrelation function $\rho(\delta)$. This value of $\rho(\delta)$ therefore indicates the best possible performance achievable by using a BPPM system with overlap. It should furthermore be noted how the analytical approximation to the QPPAM modulation scheme given in (2.25) fits well with the simulation when $SNR > 0\ dB$. Although it looks as if QPPAM is the worst performing modulation scheme of them all, it must be remembered that this scheme actually delivers twice the bitrate using the same number of pulses compared to the rest of the modulation schemes. Including this effect in the evaluation will shift the BER of QPPAM by 3dB, making QPPAM or overlapping BPPM the second best choice after BPSK.

The conclusion is therefore that in order to maximize performance of a UWB system, BPSK modulation should always be used. The rest of this thesis will therefore exclusively focus on BPSK modulation.

2.6 Summary

The basic TH and DS UWB communication systems have been introduced including the modelling of the monocycle. The single-user performance of these two systems in the AWGN channel has also been given

analytically and by simulations. It has further been shown how the BPSK modulation is superior to other schemes and it has therefore been selected as the only modulation to be used throughout the rest of this thesis.

3 Multiple-Access

The possibility of having multiple users in the system is important in UWB communications, as the typical applications will require more than one user operating in the environment at a time. This section will therefore investigate the different strategies to allow this.

Basically there are three ways of separating different users who use the same media

- **Frequency Division Multiple-Access (FDMA):** Users are separated by having a central node assign a unique frequency band to each user.
- **Time Division Multiple-Access (TDMA):** The channel is split into a number of disjoint periods of time named timeslots, which occur periodically. Each user is then assigned a given timeslot by a central node.
- **Code Division Multiple-Access (CDMA):** Each user has a unique code, which then codes the transmission in such a way that the user of interest can be demodulated at the receiver. The users are therefore separated by their codes. There are basically three different ways of performing CDMA
  - *Frequency-Hopping (FH):* Works like FDMA except for the frequency band used is determined by the code for each transmission.
  - *Time-Hopping (TH):* The channel is split into timeslots like in TDMA, but the code determines the timeslot to be used at each transmission.
  - *Direct- Sequence (DS):* The transmission is multiplied by the code in both the transmitter and the receiver and the code properties will then allow for the desired user to be demodulated.

Centralized control of the system, as needed to some degree in all of the multiple-access schemes, is usually not desirable for the applications that UWB is considered for. The system should therefore be designed to eliminate as much of this need, as it is often advantageous to have self-controlling units that operate without the need to coordinate the traffic. This way of communicating is often called ad-hoc networking, as connections are established and removed as the situation requires without involving other nodes than those directly participating in the communication. Nodes that are not directly communicating with each other will therefore not necessarily be operating synchronously and the communication must be considered being asynchronous. Still, units communicating with each other or a common node must be synchronized and be able to share the bitrate provided by the common node. This type of topology is often referred to as a piconet. The common node is in charge of controlling the traffic within the piconet by means of for instance TDMA, as the communication is already synchronized. Several piconets can then operate asynchronously within the same area. Next, in order to provide data both to and from a given node to the central node, Time Division Duplexing (TDD) is often used as it is simple and because synchronism has already been achieved.

FH-CDMA systems are not looked into further in this thesis, as the performance of FH will equal that of TH, but will require a more complex system in order to jump between the frequencies. The control of the traffic within the piconet including duplexing is also not considered here and the focus is therefore on multiple-access systems based on TH- and DS-CDMA.

Hybrids of the multiple-access methods outlined above are also possible dependent on the system. An example of such a system is a TH system transmitting several pulses pr. bit and coding each pulse of the bit as in a DS system. The user separation can then be done using the TH or DS code or both.
The hybrid approach seems to gain more and more acceptance in the IEEE 802.15.3a working group as time goes by. The advantage of using hybrids lies in the fact that more parameters are available to adjust the capabilities of the system, which makes it more flexible and scalable. By using this approach, a standard can be made which includes many different configurations of UWB systems that are all compatible. The performance of the UWB devices produced can then be adapted to the production cost and market demand. An example of such a system is the multiband physical layer proposal presented by the company General Atomics to the IEEE 802.15.3a working group [28]. In this proposal the frequency

band is split into 20 channels as in FH and the UWB system can then be scaled dependent on how many bands are available and what type of modulation is supported.

Although at the present time of writing it seems likely that the standardized system will be a hybrid solution, this is not the approach followed in this thesis. Instead solutions based on either TH- or DS-CDMA multiple-access among the piconets are investigated. The hybrid approach is rather new and has only recently during the spring of 2003 been endorsed by some of the major players in the field. The task of studying such multiband systems is thus a topic for future studies.

3.1 Time-Hopping

As mentioned before in Section 2.4, Time-Hopping (TH) makes multiple-access possible by dividing the channel into $SF = \frac{T_f}{T_{mono}}$ non-overlapping timeslots as shown in Figure 8. Users are then allowed to use timeslots assigned to them by their unique hopping code. This means that no centralized control of the timeslots is needed, but this also makes it possible for users to try and use the same timeslot at the same time causing a collision to occur. On the other hand, if users are asynchronous, collisions may occur that overlap two timeslots as the time reference of the users are different. Transmissions occur only $\frac{1}{SF}$ part of the time and the transmitted power in a given timeslot must then be SF times larger than the average transmitted power. It is therefore evident that the improvement over a system without spreading using only one pulse pr. bit, termed processing gain is given by $P_G = N_s SF$. This is correct as the received signal power at the decision device is increased by a factor SF over the average signal power at the channel, while the collected noise power is the same.

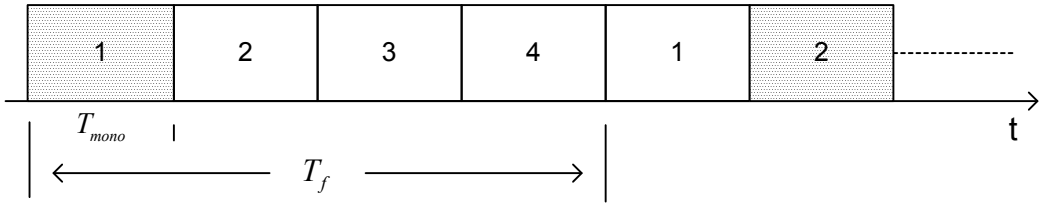

Figure 8: Splitting the channel into disjoint timeslots, in this case SF=4. Grayed-out timeslots indicate where transmissions occur.

Looking at the collision probability and assuming a uniformly distributed TH code over the available timeslots the probability of having at least one user colliding with the user of interest in a synchronous TH system is given by

$$P_{col} = 1 - \left(1 - \frac{1}{SF}\right)^{K-1} \quad (3.1)$$

where K is the total number of users in the system. When collisions occur the transmission is disrupted and given that all users are received with equal power, the approximated BER of a correlator receiver becomes

$$BER \simeq \left(1 - P_{col}\right) BER_{norm} + \frac{P_{col}}{2 \cdot 2} \quad (3.2)$$

as the BER is ¼ when a collision happen. The value of BER_{norm} indicates the normal BER of the used binary modulation scheme. It can be seen from (3.2) that collisions have a strong influence on the performance of a TH system and it is therefore important for the number of collisions in the system to be minimized.

In the asynchronous case the probability of collisions is given by

$$P_{col} = 1 - \left(1 - \frac{2}{SF}\right)^{K-1} \quad (3.3)$$

as there are now potentially two timeslots that will collide with the user in question. It is therefore natural to believe that the BER of the asynchronous system will be higher than that of the synchronous system, but it must be remembered that the BER is now dependent on the normalized autocorrelation function $\rho(\delta)$ of the monocycle for a given offset δ. The average BER of an asynchronous TH system will therefore often be a little bit lower than that of the synchronous TH system. However, it is not possible to evaluate the average BER of the asynchronous system analytically. In Appendix E it is shown how the BER can be approximated by

$$BER \simeq (1 - P_{col})\, BER_{BPSK} + \frac{SF}{8} P_{col} \int_{-T_{mono}}^{T_{mono}} erfc\left(\frac{1+\rho(\lambda)}{\sqrt{2\sigma^2}}\right) + erfc\left(\frac{1-\rho(\lambda)}{\sqrt{2\sigma^2}}\right) d\lambda \quad (3.4)$$

assuming the use of BPSK modulation and random hopping codes. The result given by (3.4) is exact for $K = 2$, but will be too optimistic when $K > 2$ as it assumes only one collider interfering with the decision.

Although using uniformly distributed TH codes as done above give a good indication of the performance achievable with TH systems, such codes are never used in real life. First of all realistic codes must be deterministic and periodic in order for the receiver to lock on to the signal. They must also be "full" meaning that during one period of the TH code all of the timeslots available must be covered in order to make full use of the system capacity. Finally the codes must minimize the number of collisions between users.
In order to quantify the performance of TH codes the ambiguity of the codes must be examined. This can be done by observing the equation [8]

$$c_{TH}^{(x)}(j+l) + m = c_{TH}^{(y)}(j) \pmod{N_p} \quad (3.5)$$

with $l \in \{0..N_p - 1\}$ and $m \in \{0..SF - 1\}$ being all possible offsets at the code- and timeslot level. Whenever $x \neq y$ and (3.5) is satisfied, a collision between different users occurs. This is also known as a hit and a hit-function giving the number of hits as a function of offset between hopping codes x and y can then be defined as

$$R_{xy}(l,m) = \sum_{j=0}^{N_p-1} \delta\left(c_{TH}^{(x)}(j+l) + m - c_{TH}^{(y)}(j)\right) \qquad , \delta(\kappa) = \begin{cases} 1, & \kappa = 0 \pmod{N_p} \\ 0, & \kappa \neq 0 \pmod{N_p} \end{cases} \quad (3.6)$$

This function is much like the ordinary correlation function because it measures the orthogonality between the TH codes for different time offsets given by l and m and the ideal code set will have the autohit-function

$$x = y: \quad R_{xy}(l,m) = \begin{cases} 1 & ,l = m = 0 \\ 0 & ,l \neq 0 \vee m \neq 0 \end{cases} \quad (3.7)$$

giving good synchronization properties and the crosshit-function

$$x \neq y: \quad R_{xy}(l,m) = 0 \quad (3.8)$$

which will eliminate any collisions from other users. Making such TH codes are simply not possible as periodic sequences will never be able to fulfill neither (3.7) nor (3.8).

The best possible autohit- and crosshit-function that a periodic TH code can achieve is realized when any combination of l and m yields no more than one hit, except for $l = m = 0$ in the autohit-function. The

reason is that when any shift is allowed, the code must collide at least once for appropriate shifts and minimizing the number of hits yields the code property [8]:

$$\max\left(R_{xy}\left(l,m\right)\right)=\begin{cases}N_p & ,x=y\wedge l=m=0\\ 1 & ,otherwise\end{cases} \quad (3.9)$$

Achieving this property is not possible as getting good autohit- and crosshit-function properties at the same time is not achievable and a tradeoff between the two is needed dependent on the application. One may then choose a code that has a property that fits this tradeoff in the best possible way.
As UWB is a relatively new technology, design of codes specifically tailored for this purpose has not yet been done, but codes designed for Frequency-Hopping (FH) CDMA systems can be used instead as the criteria for code selection is the same as outlined above for TH. Many different codes exist for FH, but some of the most used code families are the Welch-Costas [10], congruence [9], and the Reed-Solomon [14] codes. The rest of this thesis will focus on a subset of the congruence codes called hyperbolic congruence codes, as these represent the best compromise between the maximum of the autohit- and crosshit-functions having only two hits in both. These codes will be described below.

Although using the maximum number of hits in the autohit- and crosshit-function as a comparison between different types of codes is common [9], this may not be the best method as the probability of colliding with other users is not directly linked to the maximum number of hits in the hit-function. Instead the average number of hits should prove a better measure, but calculating this analytically is yet to be seen.

3.1.1 Hyperbolic Congruence Codes

The family of congruence codes is based on calculations in a finite field over a prime p denoted by $GF\left(p\right)$. A subset of these called the hyperbolic congruence code, is then defined as [8]

$$c_{TH}^{(k)}\left(j\right)=\frac{k}{j} \pmod p \quad (3.10)$$

with $k\in\left\{1..p-1\right\}$ being the user and $j\in\left\{1..p-1\right\}$ the TH code index. As any element in $GF\left(p\right)$, except the zero element, has a unique inverse element the hyperbolic congruence codes are full codes with $c_{TH}^{(k)}\left(j\right)\in\left\{1..p-1\right\}$ and having period $N_p=p-1$. It is shown in [8] that the hyperbolic congruence codes will have a maximum of 2 hits in both its autohit- and crosshit-function giving the best trade-off between the two properties.
Requiring that $p=SF+1$ makes it possible to exploit the hyperbolic congruence codes as a TH code, but at the same time limiting the possible values of the spreading factor. An interesting property of this code that has been discovered while working with it, is that although it can be shown to have maximum two hits in the autohit- and crosshit-function, it exhibits only one hit in the crosshit-function on average. It has not been possible to find material describing this phenomenon or derive it mathematically.

3.1.2 Simulation

In order to understand the impact that collisions between users in an asynchronous TH system have on the BER, a simulation of two TH systems have been performed and the result is shown in Figure 9. Both systems use hopping codes with $SF=16$ and $N_s=1$ using a correlator receiver in an AWGN channel. The used monocycle is modeled by $p_7\left(t\right)$ with a 2/3ns pulse width and it is sampled with $N_{sp}=13$ samples pr. pulse, as will be discussed in the next section.
The averaging over the asynchronous positions of the users is done over all possible sample offsets of the second user and with the rest of the users being uniformly distributed within the possible sample offsets. This is done because averaging over all possible offsets is computationally impossible when the number of users increase.

In the graph on the left the simulation using random codes is shown along with numerical results of (3.4). It should be noted how the result of (3.4) fit nicely with the simulation when $K = 2$ and how it is too optimistic when $K > 2$. This is a direct result of the fact that (3.4) includes only one collider at a time. The right graph shows simulation results of a similar system using hyperbolic hopping codes. It is seen how performance is the same, as both codes have one hit on average. The hyperbolic codes have the advantage of being deterministic and periodic as needed for realistic systems.

The simulation clearly shows how the collisions between users are a central limitation to the TH scheme, as the BER reaches a level after which increasing the SNR no longer reduces the BER. This fact can also be seen directly from the expressions in (3.2) and (3.4). Other receivers that can help mitigate the impact of the collisions are therefore important when considering multiple-access based on TH.

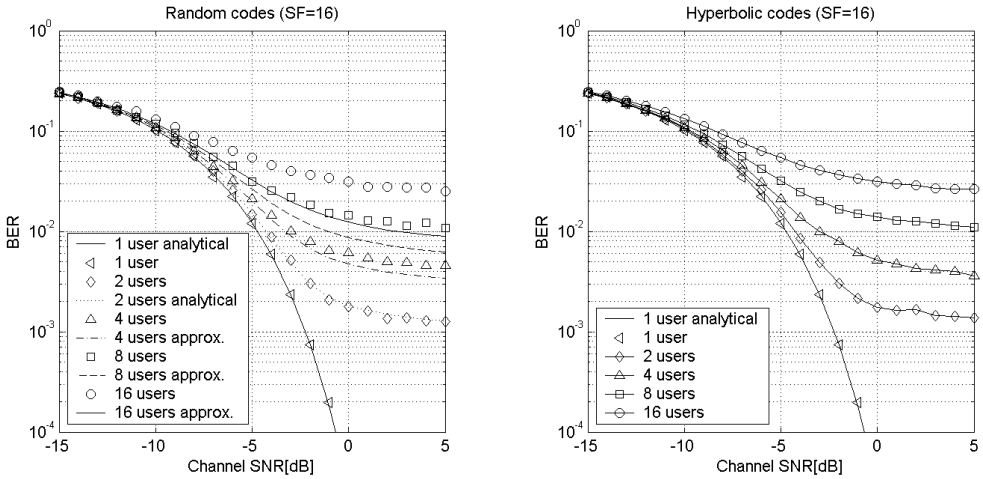

Figure 9: Simulation of a TH system as a function of the number of users.

3.2 Direct-Sequence

As opposed to TH where transmission only takes place $\frac{1}{SF}$ of the time, pulses are transmitted continuously in a Direct-Sequence (DS) system. The transmitted monocycles are then coded by multiplying them with a bipolar code of length $SF = \frac{T_f}{T_{mono}}$, which makes it possible for the receiver to separate the signals coming from different users. An example of such a code is shown in Figure 10.

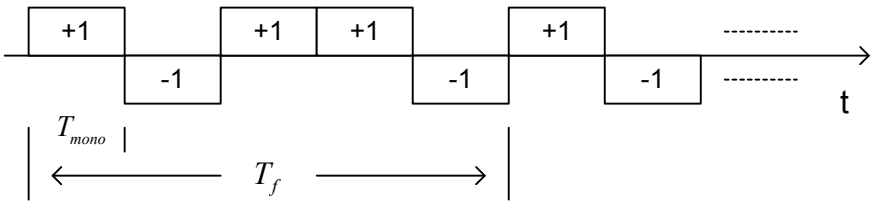

Figure 10: Example of a bipolar spreading code used in a DS system. The code shown has SF=5.

In the receiver the incoming signal $r(t)$ is multiplied by the coded monocycles $\varphi^{(l)}(t)$ and integrated in order to form sufficient statistics $y^{(l)}(n)$ for the l^{th} user

$$y^{(l)}(n) = \int_{nT_f}^{(n+1)T_f} \varphi^{(l)}(t - nT_f) \cdot r(t) dt = \int_{nT_f}^{(n+1)T_f} \varphi^{(l)}(t - nT_f) \cdot \left(\sum_{k=1}^{K} A_k s^{(k)}(t) + z(t) \right) dt \quad (3.11)$$

where $z(t)$ is white Gaussian noise with variance σ^2, A_k the amplitude of the k^{th} user and the rest is defined as in Section 2.4.2. As stated earlier it is assumed that the received monocycle $p_{rx}(t)$ is normalized so that

$$\int_0^{T_{mono}} p_{rx}(t) \cdot p_{rx}(t) dt = 1 \Leftrightarrow \int_0^{T_f} \varphi^{(l)}(t) \cdot \varphi^{(l)}(t) dt = SF \quad (3.12)$$

In a completely synchronous system the correlation coefficient between the codes of user l and k

$$\rho_{lk} = \int_0^{T_f} \varphi^{(l)}(t) \cdot \varphi^{(k)}(t) dt \quad (3.13)$$

can be used in (3.11) to yield

$$y^{(l)}(n) = \sum_{k=1}^{K} \rho_{lk} A_k \alpha^{(k)}(n) + z^{(l)}(n) = \overbrace{SF \cdot A_l \alpha^{(l)}(n)}^{Desired} + \overbrace{\sum_{k \neq l} \rho_{lk} A_k \alpha^{(k)}(n)}^{Interference} + \overbrace{z^{(l)}(n)}^{Noise} \quad (3.14)$$

with $z^{(l)}(n)$ being noise samples. It is easily seen that the correlation coefficient between any code and itself is $\rho_{ll} = SF$ and that $\rho_{lk} = \rho_{kl}$.
In the rightmost part of (3.14) the different contributions have been separated. The first part is the desired signal of user l and next to it is the interference component coming from all other users in the system. It should be noted that the interference is weighed by the correlation coefficient between the interferer and the desired user. This makes it desirable to use codes that have very low cross-correlation coefficients.

It is sometimes desirable to use the matrix notation of (3.14) given by

$$\mathbf{y}(n) = \mathbf{RA}\boldsymbol{\alpha}(n) + \mathbf{z}(n) \quad (3.15)$$

in which the correlation matrix $\mathbf{R}$ is the $K \times K$ matrix with $R_{lk} = \rho_{lk}$, $\mathbf{A}$ is the $K \times K$ diagonal matrix containing all users amplitudes and $\boldsymbol{\alpha}(n)$ is the length K column vector holding the transmitted bits from all users. Finally $\mathbf{z}(n)$ is the length K column vector consisting of the noise samples $z^{(k)}(n)$. Realizing that $\mathbf{z}(n)$ is created by a linear transform of the samples of $z(t)$ with all K different codes, the covariance matrix is given by

$$E\left[\mathbf{z}(n) \mathbf{z}(n)^T \right] = \sigma^2 \mathbf{R} \quad (3.16)$$

and the variance of $z^{(k)}(n)$ is therefore $SF \cdot \sigma^2$ as all diagonal elements of $\mathbf{R}$ are equal to SF.
In order to evaluate the performance of a DS system, it is common to model the interference as being a Gaussian random variable and thus view it as an increase in the AWGN noise introduced by the channel [15 p. 114]. This can be justified by assuming that many users are active in the system and then applying the central limit theorem on the interference component of (3.14). This results in

$$y^{(l)}(n) = SF \cdot A_l \alpha^{(l)}(n) + q(n) + z^{(l)}(n) \quad (3.17)$$

with $q(n) \sim N(0, \sigma_i^2)$ where

$$\sigma_i^2 = \sum_{k \neq l} A_k^2 \rho_{lk}^2 \quad (3.18)$$

Using an approach similar to the one used in Appendix C.2 the BER of the l^{th} user in a DS system can now be found by observing $y^{(l)}(n)$. As shown in Appendix C.2 the optimal receiver in the AWGN channel is the single-user correlator, which can be implemented by basing the decision on the sign of $y^{(l)}(n)$

$$\hat{\alpha}^{(l)}(n) = \operatorname{sgn}\left(y^{(l)}(n)\right) \quad (3.19)$$

and the decision variable will therefore be

$$\hat{\alpha} = \operatorname{sgn}\left(SF \cdot A_l + v\right) \quad (3.20)$$

with v being the sum of the noise and interference components and therefore $v \sim N\left(0, SF \cdot \sigma^2 + \sigma_i^2\right)$. The BER can then be calculated as

$$BER = \frac{1}{2} p\left(SF \cdot A_l + v < 0 \mid \alpha = +1\right) + \frac{1}{2} p\left(SF \cdot A_l + v > 0 \mid \alpha = -1\right) \quad (3.21)$$

which by symmetry is equal to

$$BER = p\left(v > SF \cdot A_l\right) = \frac{1}{\sqrt{2\pi\left(SF \cdot \sigma^2 + \sigma_i^2\right)}} \int_{SF \cdot A_l}^{\infty} e^{\frac{-\lambda^2}{2\left(SF \cdot \sigma^2 + \sigma_i^2\right)}} d\lambda = \frac{1}{2} erfc\left(\sqrt{\frac{A_l^2}{2\left(\frac{\sigma^2}{SF} + \frac{\sigma_i^2}{SF^2}\right)}}\right) \quad (3.22)$$

with the channel SNR being $SNR = \frac{A_l^2}{\sigma^2}$. As can be seen from (3.22) the BER is dependent on the spreading-factor as well as the two noise variances and of course the amplitude of the desired user. It should also be noted that a single-user system without spreading can achieve the same performance as a DS system by raising the transmitted power by a factor SF and the processing gain of the system is therefore $P_G = SF$. An issue with (3.22) that one should bear in mind is the fact that it is derived using the central limit theorem and as a consequence is only applicable when the number of users in the system is "large enough". The approximation therefore tends to be somewhat unreliable at high SNRs, as the interference is the dominant term in this case.

A major problem of using the single-user correlator receiver, as described by (3.22), occurs when $2\sigma_i^2 >> SF^2 A_l^2$, which is the case when the Signal-to-Interference Ratio (SIR) is low. In this case the receiver is overcome by the interference and performance will be bad. A scenario like this will happen when a terminal far away is received at a much lower power than a terminal close to the receiver and it is therefore often referred to as the near-far effect. To improve performance power control is often used such that the received power from all users will be the same, but this requires some form of centralized control of the power levels. In order to have a specific measure of the level of interference compared to the desired signal, the SIR is defined as

$$SIR_{UWB} = \frac{A_l^2}{\sum_{k \neq l} A_k^2} \quad (3.23)$$

This definition is not very useful for simulations, as it is very hard to say anything about the general distribution of A_k. One possible solution to this problem is to set the power of all interferers to the same

level being N_{SIR} times higher than the desired user, i.e. $A_{k\neq l}^2 = N_{SIR} A_l^2$. The SIR from (3.23) can then be written as

$$SIR_{UWB} = \frac{1}{N_{SIR}(K-1)} \quad (3.24)$$

In Chapter 5 more advanced receivers will be introduced capable of improving the performance over that of the correlator in situations where SIR_{UWB} is low.

Considering the asynchronous case, the overlapping of more than one bit pr. user must be accounted for. Arranging the users by delay as

$$\tau^{(1)} \leq \tau^{(2)} \cdots \leq \tau^{(K)} \quad (3.25)$$

helps simplify the expressions. The partial cross-correlation between the different users are

$$\rho_{lk} = \int_0^{\tau^{(k)}} \varphi^{(l)}(t) \cdot \varphi^{(k)}\left(t - \tau^{(k)}\right) dt \quad (3.26)$$

$$\rho_{kl} = \int_{\tau^{(k)}}^{T_f} \varphi^{(l)}(t) \cdot \varphi^{(k)}\left(t - \tau^{(k)} + T_f\right) dt \quad (3.27)$$

with ρ_{lk} and ρ_{kl} being the contribution from the previous and current bit respectively from user k onto the output of the matched of user l. Using these partial cross-correlations an analogous expression to (3.14) can be found as

$$y^{(l)}(n) = \overbrace{SF \cdot A_l \alpha^{(l)}(n)}^{Desired} + \overbrace{\sum_{k=1}^{l-1} \rho_{kl} A_k \alpha^{(k)}(n) + \sum_{k=1}^{l-1} \rho_{lk} A_k \alpha^{(k)}(n+1)}^{Interference\ from\ users\ k<l} + \overbrace{\sum_{k=l+1}^{K} \rho_{lk} A_k \alpha^{(k)}(n-1) + \sum_{k=l+1}^{K} \rho_{kl} A_k \alpha^{(k)}(n)}^{Interference\ from\ users\ k>l} + \overbrace{z^{(l)}(n)}^{Noise} \quad (3.28)$$

or in a matrix notation

$$\mathbf{y}(n) = \mathbf{R}[1]\mathbf{A}\boldsymbol{\alpha}(n-1) + \mathbf{R}[0]\mathbf{A}\boldsymbol{\alpha}(n) + \mathbf{R}^T[1]\mathbf{A}\boldsymbol{\alpha}(n+1) + \mathbf{z}(n) \quad (3.29)$$

where

$$R_{kl}[0] = \begin{cases} SF & , k = l \\ \rho_{kl} & , k < l \\ \rho_{lk} & , k > l \end{cases} \quad (3.30)$$

$$R_{kl}[1] = \begin{cases} 0 & , k \leq l \\ \rho_{lk} & , k > l \end{cases} \quad (3.31)$$

The noise contribution can be seen to have covariance matrix

$$E\left[\mathbf{z}(n)\mathbf{z}(i)^T\right]=\begin{cases}\sigma^2\mathbf{R}^T[1] & ,i=n-1\\ \sigma^2\mathbf{R}[0] & ,i=n\\ \sigma^2\mathbf{R}[1] & ,i=n+1\\ \mathbf{0} & ,otherwise\end{cases} \quad (3.32)$$

as the noise samples at the output of each users matched filter is created by a linear transform of $z(t)$ with each users code at the given offset. An approximation using the same assumptions as in (3.17) can be constructed using the central limit theorem, resulting in

$$\sigma_i^2=\sum_{k\neq l}\left(|\rho_{kl}|+|\rho_{lk}|\right)^2 A_k^2 \quad (3.33)$$

which can be used in (3.22) to estimate the BER of the system when hard-decision on the output of the matched filter is used. Still this expression will only be an approximation to the BER at a given offset, but it is usually of more interest to know the BER averaged over all possible offsets. In order to do so (3.22) must be integrated and the same problem as outlined for asynchronous TH described in Appendix E will apply here. It is therefore not possible to give a closed-form expression of the average BER, but performance will generally be better than the synchronous case, as the average cross-correlations are generally lower. This is of course not the case if codes that rely on synchronous operation are used, such as orthogonal codes [19].

3.2.1 Selected Code Families

Some of the possible spreading codes that are candidates for an asynchronous UWB DS system will be considered here. The types of codes presented here are the maximal-length, Kasami, and random sequences. Other types of codes with other properties also exist, but the ones outlined here are the classical ones, characterized by being optimal or near optimal by specific measures.

Maximal-length Sequences

The maximal-length sequences, or simply m-sequences, are characterized by being the periodic sequences that have the longest period possible produced by a Linear Feedback Shift Register (LFSR) using n delay elements [11]. The structure of the LFSR is shown in Figure 11. The feedback connections $c_j \in \{0,1\}$ and similar for the shift register in- and outputs $a_i \in \{0,1\}$. All calculations are therefore performed modulo 2. The longest periodic sequence that can be generated using an LFSR has period $N_p = 2^n - 1$, as the all-zero vector is of no use. The output may then be represented by the length N_p vector $\mathbf{a}$.

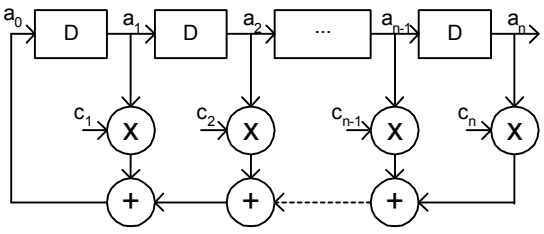

Figure 11: Structure of a LFSR.

The feedback connections are normally expressed by a polynomial as

$$f(D)=\sum_{j=0}^{n} c_j D^j \quad (3.34)$$

with D being the delay operator, similar to z^{-1} used in digital signal processing terms. This polynomial determines the behavior of the LFSR and is known as the characteristic polynomial for this reason. If the LFSR is to produce m-sequences, it can be shown that $f(D)$ must be irreducible resulting in the zero- and highest-order coefficients being one, i.e. $c_0 = c_n = 1$. An example of an irreducible characteristic polynomial is $f(D) = D^5 + D^2 + 1$ leading to $N_p = 2^5 - 1 = 31$.

In order to use m-sequences as spreading codes in a DS system the code must be bipolar, which is achievable by defining a new output as $\overline{\mathbf{a}} = 2\mathbf{a} - \mathbf{1}$. One of the interesting properties of m-sequences is the fact that its auto-correlation is two-valued

$$R_{\overline{aa}}(l) = \sum_{i=1}^{N_p} \overline{a}_i \overline{a}_{i+l} = \begin{cases} N_p & , l = 0 \left(\mathrm{mod}\, N_p\right) \\ -1 & , l \neq 0 \left(\mathrm{mod}\, N_p\right) \end{cases} \quad (3.35)$$

making it a close approximation to a delta-function for large N_p. It is therefore an interesting sequence for synchronization purposes. The cross-correlation between two different m-sequences $\overline{\mathbf{a}}$ and $\overline{\mathbf{b}}$ is given as

$$R_{\overline{ab}}(l) = \sum_{i=1}^{N_p} \overline{a}_i \overline{b}_{i+l} \quad (3.36)$$

Unfortunately the cross-correlation properties of m-sequences are not very good and thus they form bad spreading codes. They can however be used as a basis for producing codes with better cross-correlation properties. To compare different types of spreading codes, it is interesting to be able to lower-bound the cross-correlation in order to know the best possible performance. This is possible using the Welch lower-bound [19] stating

$$\left|R_{\overline{ab}}(l)\right| \geq N_p \sqrt{\frac{M-1}{N_p M - 1}} \simeq \sqrt{N_p} \quad (3.37)$$

with M being the number of codes in the set.

Kasami Codes

The set of Kasami codes is interesting as it provides very low cross-correlation between the codes. The codes are usually split into two sets, one named *small-set* and one named *large-set*. The small-set Kasami codes provide the best cross-correlation properties, but only a small number of such codes are available. The large-set has more codes and includes the small-set as a subset, but has higher cross-correlation values [19].

To generate the small-set the m-sequence $\mathbf{a}$ must be decimated by a factor of $2^{n/2} + 1$ to produce the sequence $\mathbf{a}'$ with n being even. It can be verified that $\mathbf{a}'$ is itself an m-sequence of period $N_{p,a'} = 2^{n/2} - 1$. By taking $2^{n/2} + 1$ replicas of $\mathbf{a}'$ and adding it to $\mathbf{a}$ using all $N_{p,a'} - 1$ different cyclic shifts, a new set of sequences is formed. Finally by including $\mathbf{a}$ in the set, the small-set Kasami codes are produced having $M = 2^{n/2}$ different codes in the set. The auto- and cross-correlation functions for this code set can be shown to be three-valued taking on values

$$R_{small-set}(l) \in \left\{-1, -1 \pm 2^{n/2}\right\} \quad (3.38)$$

The Welch lower bound on the cross-correlation is therefore approached for large N_p and the small-set Kasami codes are therefore optimal.

Turning to the large-set Kasami codes these are produced in a similar way. Besides the m-sequences $\mathbf{a}$ and $\mathbf{a}'$ used for the small-set, the m-sequence $\mathbf{a}''$ must be formed by decimation of $\mathbf{a}$ by a factor of $2^{n/2+1}$. It is now possible to generate the large-set Kasami codes by adding $\mathbf{a}$, $\mathbf{a}'$ and $\mathbf{a}''$ for all possible shifts of $\mathbf{a}'$ and $\mathbf{a}''$. By carefully counting, the total number of different sequences can be seen to be

$$M = \begin{cases} 2^{3n/2} & , n = 0 \pmod 4 \\ 2^{3n/2} + 2^{n/2} & , n = 2 \pmod 4 \end{cases} \quad (3.39)$$

The auto- and cross-correlation for this code set can attain five different values

$$R_{\text{large-set}}(l) \in \left\{-1, -1 \pm 2^{n/2}, -1 \pm 2^{n/2+1}\right\} \quad (3.40)$$

and is therefore not capable of touching the Welch lower bound. On the other hand more codes are available and the choice of whether to use small- or large-set Kasami codes is therefore dependent on the number of codes needed.

Random Codes

It is sometimes of interest to use random spreading codes to evaluate system performance in order to separate the choice of spreading codes from the system. This is done by letting $c_{DS}^{(x)}(j) \in \{-1, +1\}$ be a Bernoulli random variable with probability ½. The cross-correlation of this code will therefore also be stochastic with mean

$$E\left[\sum_{j=0}^{SF-1} c_{DS}^{(x)}(j)\, c_{DS}^{(y\neq x)}(j-l)\right] = 0 \quad (3.41)$$

where the expectation is taken with respect to the lag l. The variance can be found from the fact that the correlation is the sum of SF independent contributions each with variance one and is therefore

$$E\left[\left(\sum_{j=0}^{SF-1} c_{DS}^{(x)}(j)\, c_{DS}^{(y\neq x)}(j-l)\right)^2\right] = SF \quad (3.42)$$

The results given by (3.41) and (3.42) also apply for the auto-correlation when $l \neq 0$ as all values of the code are independent.

3.2.2 Simulation

A simulation of an asynchronous DS UWB system with multiple users has been performed in order to quantify the performance of such a system in the AWGN channel. The system uses $SF = 15$ and is based on large-set Kasami codes with all small-set Kasami codes being allocated first. The averaging over the asynchronous positions of the users are done over all possible sample offsets of the second user and with the rest of the users being uniformly distributed within the possible sample offsets.

The result shown in Figure 12 clearly shows how the performance of the correlator is degraded when more users are added to the system raising the level of multiple-access interference at the receiver. In the case where the interferers are received at 10 times the power of the desired user i.e. $N_{SIR} = 10$, the problem becomes very severe as discussed earlier. It is therefore of great interest to design receivers that can reduce the impact of multiple-access interference making it possible to serve more users with the system.

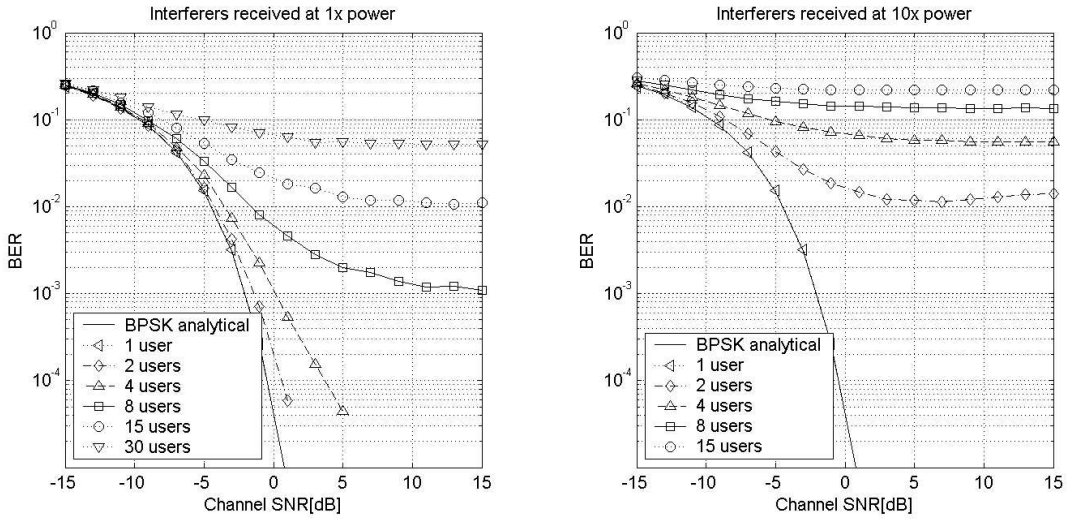

Figure 12: Simulation of an asynchronous DS-BPSK UWB system in the AWGN channel.

3.3 PSD of the UWB Signal

As discussed previously the PSD of the monocycle $\Sigma_{p_x}(f)$ is given by (2.4), but this expression does not include the effects of modulation and spreading on the signal. As UWB systems are power limited it is important to know how this affects the PSD in order to maximize transmission power without exceeding the limits set out by the regulatory body.

It is shown in Appendix F how no spectral lines are present in the PSD of the signal if BPSK modulation is used. This is a great advantage, as the energy will be spread out over a continuous spectrum making a higher transmission power possible. In Appendix F the PSD of a TH-BPSK signal with $N_s = 1$ is derived as

$$\Sigma_{TH}^{(k)}(f) = \frac{\left|P_{tx}(f)\right|^2}{T_f} = \frac{\Sigma_{p_x}(f)}{T_f} \quad (3.43)$$

where $P_{tx}(f)$ is the Fourier transform of $p_{tx}(t)$. It is further shown how the PSD of a DS-BPSK signal is given by

$$\Sigma_{DS}^{(k)}(f) = \frac{\left|P_{tx}(f)\right|^2}{T_f}\left|C_{DS}^{(k)}(f)\right|^2 = \frac{\Sigma_{p_x}(f)}{T_f}\left|C_{DS}^{(k)}(f)\right|^2 \quad (3.44)$$

with $C_{DS}^{(k)}(f)$ being the spectrum of the spreading code given by

$$C_{DS}^{(k)}(f) = \sum_{i=0}^{SF-1} c_{DS}^{(k)}(i) e^{-j2\pi f i T_{mono}} \quad (3.45)$$

The periodic hopping code of a TH system consequently has no influence on the PSD of the signal whereas the DS spreading code has. However, if $N_s > 1$ the TH hopping code can potentially affect the PSD of the signal as shown in [29]. In connection with UWB systems the best spreading code, as seen from a transmission power point of view, is the one resulting in a constant PSD level across the spectrum of interest, that is

$$\left|P_{tx}(f)\right|^2 \left|C_{DS}^{(k)}(f)\right|^2 = k_c \quad , f_l < f < f_u \quad (3.46)$$

where k_c is an arbitrary constant and f_l and f_u are the lower and upper frequency limits respectively. However designing codes that satisfy (3.46) and have good correlation properties is impossible and a tradeoff between the two is needed dependent on the use.

As the spectral properties of the resulting UWB signal are now known, it is possible to decide on a reasonable choice of monocycle bandwidth, center frequency and sample rate, as these will be used throughout the rest of this thesis. The first parameters to select are the bandwidth and the center frequency. It is possible to use the expressions of Appendix B to find the frequency where the PSD is largest and this method may be used as an indication of the center frequency, but this is dependent on the chosen monocycle. Selection of the monocycle and its pulse width is therefore performed empirically. It has therefore been decided that the monocycle $p_6(t)$ with a pulse width around 2/3ns gives a reasonable use of the spectrum, as it occupies the band from 4-6.5GHz defined by its –3dB points. It also fulfils the regulatory emission maximum by being attenuated by at least 10dB outside the 3.1-10.6GHz band in which UWB transmission with up to –41.3dBm/MHz is allowed.
Selecting this monocycle as the propagating signal, the received signal will be the derivative of the monocycle i.e. $p_7(t)$ and the signal imposed on the transmitter antenna is $p_5(t)$. The choice of sample rate must therefore be based on sufficient rejection of aliasing when sampling $p_7(t)$ in the receiver as described in Appendix B.

It was decided to use $k_w' = 9$ for the pulse $p_7(t)$ before the table in Appendix A was finalized, as it was believed that this value was the correct one to use. However, when the table was finalized the conclusion was that the correct value should have been $k_w' = 7$ in order to capture more than 99.9% of the pulse energy when truncating the pulse duration. As simulations had already been conducted using $k_w' = 9$, it was decided to continue doing so to make simulations comparable. It has further been decided to use $N_{sp} = 13$ samples pr. pulse as this seems to give a reasonable rejection of aliasing at half the sampling frequency which shown in the figure below.

Figure 13 shows the PSD of the selected monocycle for $PW = \frac{2}{3} ns$ and $N_{sp} = 13$ with a transmission power of 10^{-4} Watts. The sample frequency is therefore $f_s = \frac{N_{sp}}{PW} = 19.5 GHz$, which is rather high even for a UWB system. It is possible to reduce this figure significantly by allowing more aliasing in the receiver, but this is not considered further as this is not within the scope of this thesis.

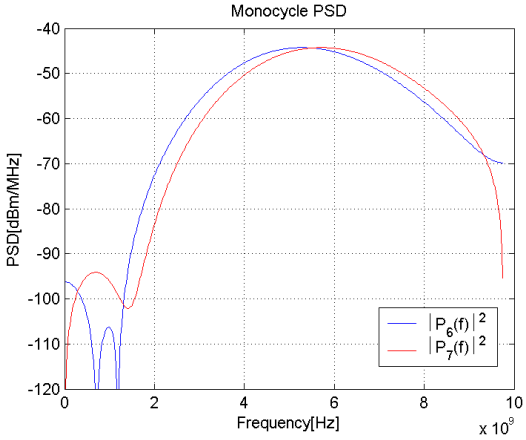

Figure 13: PSD of the used monocycle.

3.3.1 Simulation

To confirm the theoretical PSD expressions given by (3.43), (3.44) and (3.45) a simulation of the PSD of a TH and DS signal have been performed as shown in Figure 14. All system parameters are as mentioned above and the spectrum estimation of the modulated signals are carried out over a total of 10000 bits and is measured in a 1MHz bandwidth.

The TH system uses a hyperbolic hopping code with $SF = 16$ and as expected the hopping code has no influence on the PSD of the signal. In the case of the DS system an m-sequence with $SF = 15$ is used as spreading code and it is seen how the code spectrum directly influences the PSD of the signal. This has the unwanted consequence that peaks in the code spectrum causes the PSD of the transmitted signal to exceed the PSD of the monocycle. Careful selection of the code is therefore needed to maximize transmission power in the DS system.

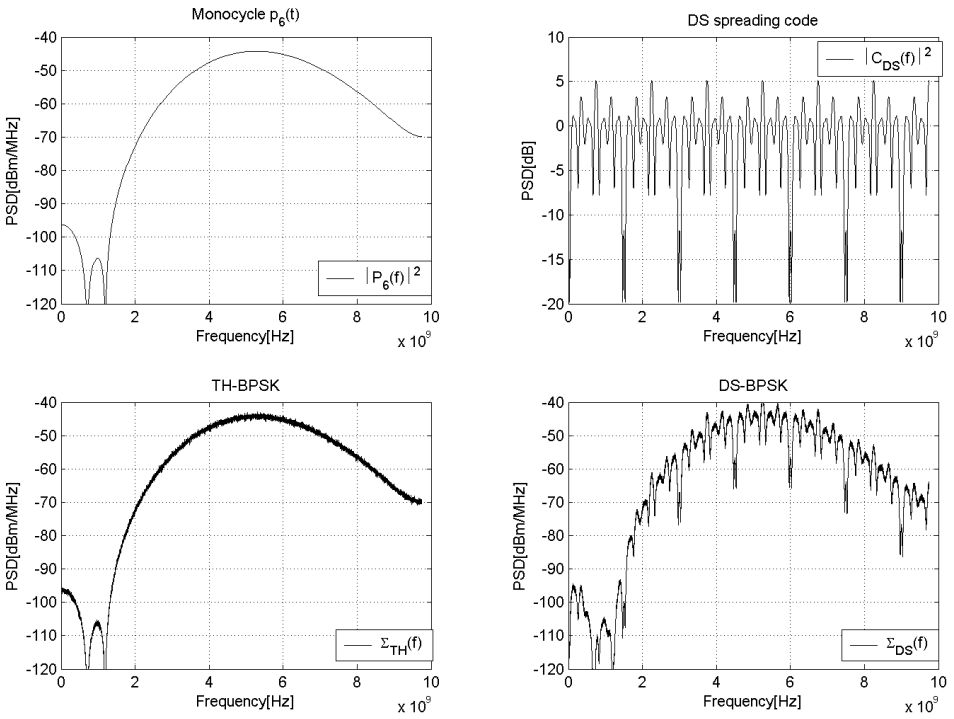

Figure 14: PSDs of TH and DS systems.

3.4 Summary

In this chapter the TH and DS multiple-access schemes are outlined in detail. Appropriate codes for the two systems are discussed and the use of random and hyperbolic hopping codes is described and simulated in a TH system. For the DS system the maximal-length, Kasami and random sequences are introduced and a simulation using Kasami codes is performed. The simulations of the multi-user TH and DS systems clearly show how important the impact of the multiple-access interference is on the BER performance. The PSD of both TH and DS schemes are furthermore derived and a simulation is performed to verify the results.

4 Radio Channel Modelling for UWB

Until now only the AWGN channel has been considered for UWB communications, but unfortunately this is not realistic in many cases and therefore a more realistic channel model must be constructed. This is due to the fact that not only one path exists between transmitter and receiver in a typical environment as illustrated in Figure 15. In this theoretical propagation example pulses transmitted from the transmitter on the left to the receiver on the right has a direct path, a reflected path coming from the floor and one coming from the wall. At the receiver all the electromagnetic pulses will add up creating the electromagnetic field imposed on the receiving antenna. An environment such as this that give rise to more than one path from transmitter to receiver is termed a multipath environment with each path known as a multipath component. Scattering of the electromagnetic pulse, which happens when the wavelength considered is comparable to the structures illuminated by the wave, also causes multiple paths from transmitter to receiver. Finally, diffraction of the electromagnetic wave around obstacles is another source of multipath components especially under Non Line Of Sight (NLOS) conditions.

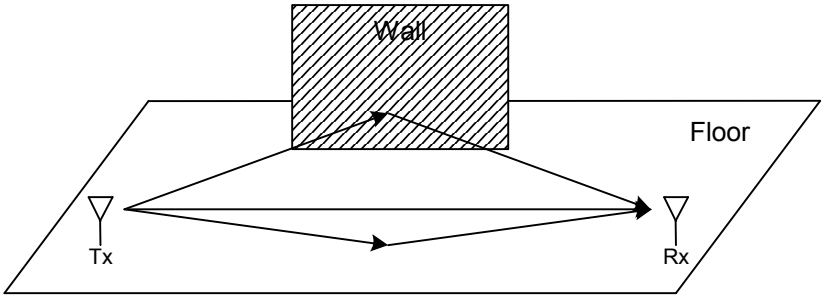

Figure 15: Illustration of multipath components.

A realistic environment will therefore often be so complex that deterministic calculation of the radio channel impulse response will be impossible and also of little practical interest as the impulse response will change dramatically from location to location. The channel model is therefore very often modeled as being stochastic. One common way of modelling the radio channel is to represent the channel as a tapped delay line filter with stochastic tap weights as shown in Figure 16.

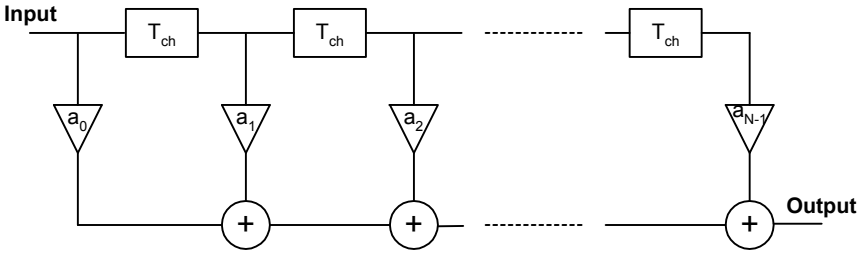

Figure 16: Radio channel modeled as a FIR filter with stochastic tap weights.

The delay T_{ch} between the N tabs in the filter is fixed for a given model and the filter coefficients are random variables drawn from a suitable distribution given by the channel model. The impulse response then becomes

$$h^{(k)}(t) = \sum_{i=0}^{N-1} a_i^{(k)}(t)\,\delta(t - iT_{ch}) \quad (4.1)$$

where k indicates the channel model belonging to the k^{th} user being active in the radio channel. The filter coefficients can often be assumed slowly varying as moving objects in the physical channel will alter the impulse response. The coefficients are therefore assumed constant within the observed time window as represented by $a_i^{(k)}(t) = a_i^{(k)}$. The contribution from each user $r^{(k)}(t)$ onto the received

signal $r(t)$ can now be found by convolving the transmitted signal of each user $s^{(k)}(t)$ with the associated channel model for that user given by (4.1) yielding

$$r^{(k)}(t) = \int_{-\infty}^{\infty} s^{(k)}(t-\tau) h^{(k)}(\tau) d\tau = \sum_{i=0}^{N-1} a_i^{(k)} s^{(k)}(t - iT_{ch}) \quad (4.2)$$

and the received signal is therefore

$$r(t) = \sum_{k=1}^{K} r^{(k)}(t) + z(t) = \sum_{k=1}^{K} \sum_{i=0}^{N-1} a_i^{(k)} s^{(k)}(t - iT_{ch}) + z(t) \quad (4.3)$$

with $z(t)$ being white Gaussian noise with variance σ^2. In the receiver $r(t)$ is sampled with an appropriate sample spacing T_s to provide a column vector $\mathbf{r}(n)$ holding the received values given by

$$r_j(n) = r(nT_f + jT_s) \quad , j = 0,1,...,P-1 \quad (4.4)$$

with the length P determined by the integer $P = \left\lceil \frac{T_f + (N-1)T_{ch}}{T_s} \right\rceil$ so that all the energy from the bit of interest can be collected after being spread out by the radio channel. Although the expression of (4.3) includes all bits transmitted, only a finite number of bits, N_1 before and N_2 after the information bit, will contribute energy to $\mathbf{r}(n)$. The contribution to the received signal from a single bit with no modulation $v^{(k)}(t)$, is found as

$$v^{(k)}(t) = \sum_{i=0}^{N-1} a_i^{(k)} \varphi^{(k)}(t - iT_{ch} - \tau^{(k)}) \quad (4.5)$$

with the length P column vector of samples $\mathbf{v}^{(k)}(m)$, found in a similar way as in (4.4), given by

$$v_j^{(k)}(m) = v^{(k)}(mT_f + jT_s) \quad , j = 0,1,...,P-1 \quad (4.6)$$

It is now possible to express $\mathbf{r}(n)$ using only the relevant bits as

$$\mathbf{r}(n) = \sum_{k=1}^{K} \sum_{m=-N_1}^{N_2} \alpha^{(k)}(n+m) \mathbf{v}^{(k)}(m) + \mathbf{z}(n) \quad (4.7)$$

with $\mathbf{z}(n)$ containing the noise samples. The maximum bit offset that contribute energy to $\mathbf{r}(n)$ is therefore

$$N_1 = N_2 = \left\lceil \frac{(N-1)T_{ch}}{T_f} \right\rceil \quad (4.8)$$

One of the interesting properties of UWB transmission is the very short pulse width of the transmitted monocycles being in the order of a nanosecond. This translates into a spatial length of the monocycle of approximately 30 cm, making it possible to separate the multipath components much more accurately than in narrowband systems. As a result of this it is believed that the multipath components can be separated down to approximately 10 cm of path difference, making it possible to determine the position with a precision in the area of 10 cm. Another side effect is that no more than a few multipath components add up within the pulse width of the monocycle, making constructive and destructive combining of multipath components much less critical [17]. It is therefore not necessary to use large fading margins to counter fading which is the case in ordinary narrowband communications.

An assumption often made in UWB channel modelling is that the multipath components are discrete [16, 20, 21] meaning that the multipath components do not overlap in time. This assumption is based on the short spatial length of the monocycle and has the consequence that the received signal will consist of a set of delayed replicas of the transmitted monocycle each multiplied with their respective amplitude $a_i^{(k)}$. This assumption may be too optimistic, but current state-of-the-art models are generally based on this concept.

There seems to be a general consensus that the amplitude distribution of the filter coefficients $a_i^{(k)}$ in a wideband channel model is best modeled by a Power Delay Profile (PDP) relating the received power of the multipath component to the path delay experienced by the component. The PDP is then split into bins of a given duration and the received power is integrated over this period of time. On average the received power decays exponentially as a function of the bin number and the actual received power is Gamma distributed with the average received power as mean. The reason that the Gamma distribution tends to fit the measured data well can be understood by the short duration of the monocycles giving rise to a low number of multipath components being combined. If the number of combining multipaths were large, one would expect a Rayleigh or Rician distribution, as is normally the case in narrowband communications [11].
In Figure 17 an example of the PDP model described is shown with the bars indicating the average received power decaying exponentially and the line showing the actual received power as a Gamma distribution with the average received power as mean.

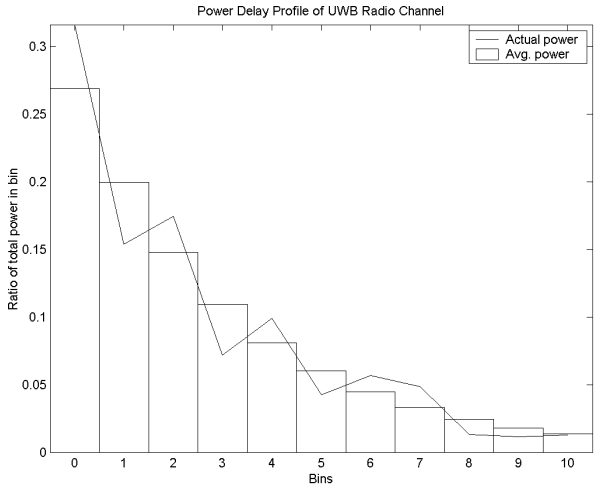

Figure 17: Illustration of UWB Power Delay Profile modelling.

In order to use the PDP to find the filter coefficients, the properties of the sign of the coefficients must be known. This is relevant as reflections tend to change the sign of the electromagnetic pulse and this must be included in the model. In an ordinary narrowband system this phenomenon is usually accounted for by letting the phase of the complex baseband signal be uniformly distributed from 0 to 2π, but as UWB signals are inherently baseband introducing such a complex baseband notation is not justified. Therefore it makes more sense to let the sign of the multipath components be +1 or –1 with equal probability as supported by [22].

One thing not taken into consideration by the PDP model described until now is the fact that bins may be correlated in other ways than just by the exponential power decay. This can happen when waves are scattered around an obstacle, producing several correlated waves being only slightly delayed with respect to each other. This phenomenon is commonly referred to as clustering, because a cluster of paths will arrive with little time difference. One common way of modelling this is to allocate stochastic exponential power decays to each cluster and the PDP will therefore consist of an overall exponential power decay and one for each cluster.

4.1 The Used Channel Model

The radio channel model selected in this thesis for simulations is specified by [16]. The model is based on a series of channel measurements in a given building. It should not be considered the ultimate truth within UWB radio channel modelling and more advanced and accurate models will certainly be produced over time. Nonetheless it is quite representative of the physical nature of the UWB radio channel and is considered state-of-the-art at the present time of writing. As the channel model is somewhat comprehensive, it is recommended that [16] is consulted for a complete description as information provided here only gives an overview of the model.

The model is split into two parts

- Large-scale statistics
- Small-scale statistics

The first part accounts for the changes in the channel when moving the receiver a significant distance compared to the transmitter-receiver separation, e.g. moving from one room to the other. The small-scale statistics on the other hand determines the changes in the channel experienced when moving the receiver a small distance without changing the overall environment.

4.1.1 Large-scale Statistics

The large-scale statistics are modeled by a PDP using bins with exponentially decaying received powers as described earlier in more general terms although the clustering effect is not included. The bin width used is $\Delta\tau = 2ns$ and the model includes $N_{bins} = 100$ bins. The delay incurred in the k^{th} bin is therefore $\tau_k = (k-1)\Delta\tau$ with $k \in \{1, 2, \ldots, N_{bins}\}$, making it possible for the impulse response to have up to 198ns duration.

The average energy gain of the k^{th} bin $\overline{G_k}$ is then defined as the ratio between the average received energy in the k^{th} bin and the total received energy received at a distance of 1m. The total average energy gain $\overline{G_{tot}}$ is therefore defined as [16]

$$\overline{G_{tot}} = \sum_{k=1}^{N_{bins}} \overline{G_k} = \overline{G_1} + \sum_{k=2}^{N_{bins}} \overline{G_2} \exp\left[\frac{-(\tau_k - \tau_2)}{\varepsilon}\right] \quad (4.9)$$

with the right-most part being based on the assumption that the average energy gain of the k^{th} bin $\overline{G_k}$ can be modeled as being exponentially decaying with the decay constant ε. Introducing the power ratio $r = \frac{\overline{G_2}}{\overline{G_1}}$ in (4.9) gives

$$\overline{G_{tot}} = \overline{G_1}\left(1 + r\sum_{k=2}^{N_{bins}} \exp\left[\frac{-(\tau_k - \tau_2)}{\varepsilon}\right]\right) = \overline{G_1}\left(1 + rF(\varepsilon)\right) \quad (4.10)$$

with $F(\varepsilon)$ found by summing the geometric series as

$$F(\varepsilon) = \frac{1 - \exp\left[-(N_{bins} - 1)\frac{\Delta\tau}{\varepsilon}\right]}{1 - \exp\left[-\frac{\Delta\tau}{\varepsilon}\right]} \simeq \frac{1}{1 - \exp\left[-\frac{\Delta\tau}{\varepsilon}\right]} \quad (4.11)$$

The values of $\overline{G_k}$ are now completely specified by $\overline{G_{tot}}$, r, and ε and can be found as

$$\overline{G_k} = \begin{cases} \dfrac{\overline{G_{tot}}}{1+rF(\varepsilon)} & ,k=1 \\ \dfrac{\overline{G_{tot}}}{1+rF(\varepsilon)} r \cdot \exp\left[\dfrac{-(\tau_k - \tau_2)}{\varepsilon}\right] & ,k>1 \end{cases} \quad (4.12)$$

The channel model specifies the total average energy gain $\overline{G_{tot}}$ as having a lognormal distributed around the path-loss PL at a given distance, as shadowing of the transmitter will influence the total power received at a given location. The total average energy gain is distributed according to [16]

$$\overline{G_{tot}} \sim L_N(-PL, 4.3) \quad (4.13)$$

where $L_N(\mu_{LN}, \sigma_{LN})$ indicates that the random variable is normally distributed in dB with mean μ_{LN} and standard deviation σ_{LN} both in dBs. The path-loss is found by a dual slope model given by [16]

$$PL = \begin{cases} 20.4 \cdot \log_{10}\left(\dfrac{d}{1m}\right) & ,d \le 11m \\ -56 + 74 \cdot \log_{10}\left(\dfrac{d}{1m}\right) & ,d > 11m \end{cases} \quad (4.14)$$

where the distance d is measured in meters and PL in dBs. The power ratio and decay constant is given by [16]

$$r \sim L_N(-4, 3) \quad (4.15)$$

$$\varepsilon \sim L_N(16.1, 1.27) \quad (4.16)$$

where the values of ε in dB is found as

$$\varepsilon_{dB} = 10\log_{10}\left(\frac{\varepsilon}{1ns}\right) \quad (4.17)$$

4.1.2 Small-scale Statistics

As mentioned before the actual received energy in the k^{th} bin G_k is Gamma distributed with $\overline{G_k}$ as mean and with shape-factor m_k

$$G_k \sim \Gamma(\overline{G_k}, m_k) \quad (4.18)$$

In the model, the measured channels are used to find a suitable distribution of m_k and the authors decide on a truncated Gaussian, that is the pdf of m_k is given by [16]

$$f_m(x) = \begin{cases} K_m \exp\left[\dfrac{-(x-\mu_m)^2}{2\sigma_m^{\,2}}\right] & ,x \ge 0.5 \\ 0 & ,otherwise \end{cases} \quad (4.19)$$

with K_m being a normalization constant. As the value of μ_m and σ_m depend on the delay, the following regression lines are found [16]

$$\mu_m(\tau_k) = 3.5 - \frac{\tau_k}{73} \quad (4.20)$$

$$\sigma_m^{\,2}(\tau_k) = 1.84 - \frac{\tau_k}{160} \quad (4.21)$$

with τ_k being in nanoseconds.

4.1.3 Using the Model

The model is now complete and the stochastic filter coefficients of the l^{th} user can be found by generating G_k and finding the coefficients by

$$a_k^{(l)} = b\sqrt{G_k} \quad (4.22)$$

where b is a Bernoulli random variable with equal probability +1 and –1 to account for inversions of the pulse by reflections. Notice that when the filter coefficients are specified using (4.22) the channel model will output the signal power relative to a distance of 1m from the transmitter. An example of such an impulse response at a distance of $d = 10m$ is given in Figure 18.

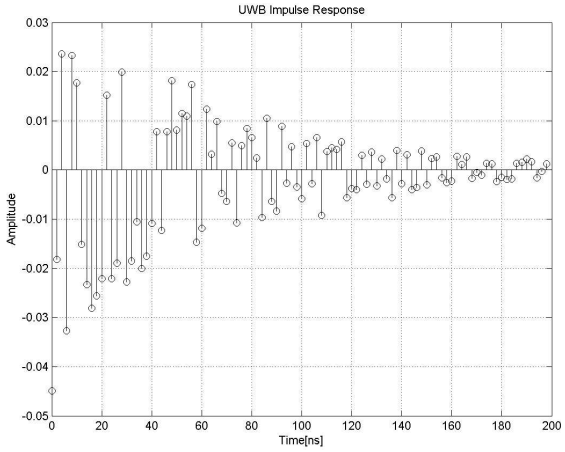

Figure 18: A typical UWB impulse response at d=10m.

When using the channel model for simulations, each user is convolved with an impulse response drawn from the distribution specified by the channel model. Ideally a new impulse response would be drawn from the channel model for each new data block simulated to estimate the average BER over the channel. The problem with this approach is, that as the difference from one channel to the other can be quite large, it requires a huge number of simulations to make the result reproducible with only a small variance. This is needed in order to compare simulations where parameters other than those specified in the channel model are varied.
As this is computationally unattractive it has instead been decided to use 10 fixed channels for each user making it possible to directly compare simulations. The drawback of this solution is that there is no guarantee that these fixed channels will result in the true mean BER in simulations or describe the channel model fully. On the other hand tests have shown that the true BER is within ±0.5dB, when using the 10 fixed channels for simulations.

A major drawback of the outlined channel model is that the spacing between the channel taps is larger than the pulse duration. This can have the unwanted effect that multipath components introduced by the model, arrive only at a few predetermined offsets with respect to the timing of the receiver. However, reducing the channel tap spacing is not easily done. First, measuring the channel impulse response with such high temporal precision is currently not possible, because of the extremely short pulse durations. Second, decreasing the tap spacing will make the complexity of convolving the channel impulse response with the transmitted pulses grow by at least an order of magnitude. This is not realistic as convolving the channel impulse responses with the users transmitted pulses is already the most computational demanding task of simulating multi-user UWB systems. A different approach to this problem will therefore be introduced next.

4.1.4 Channel Model Artifacts

Usually when simulating a communication system operating over a channel the channel model is designed specifically to the system, but in this case the model is fixed. It is therefore important to take notice of potential conflicts between the system and the channel model. One thing of specific interest is the spacing between the taps, which is set in the model to $\Delta\tau = 2ns$. It is therefore necessary to select a monocycle pulse width PW and spreading-factor SF that ensures that the multipath components will be distributed over the possible offsets, as the level of ISI otherwise may be either too high or too low. The relative offset between the first arriving path and the i^{th} multipath component $\Delta t_o(i)$ is given by

$$\Delta t_o(i) = i \cdot \Delta\tau \pmod{SF \cdot PW} \quad (4.23)$$

It is desired that $\Delta t_o(i)$ is uniformly distributed in the interval $[0; SF \cdot PW[$ in order for the channel model to be as realistic as possible by averaging out the offsets introduced by the taps over all possible offsets. This is however not possible to achieve using a tapped delay line model, but an approximation can be made by selecting $PW = \frac{N\Delta\tau}{SF}$, where $N = 100$ is the number of taps in the channel model having a spacing of $\Delta\tau$. Selecting this pulse width insures that $\Delta t_o(i) = 0$ only in the case when $i = 0 \pmod{N}$. This means that $\Delta t_o(i)$ can attain N different values as i goes from zero to $N-1$ providing a good linear averaging over the possible offsets. The problem with this solution is that the value of PW will be too large for UWB usage. Furthermore, the average received energy is dependent on i, making it desirable to average in the interval $[0; SF \cdot PW[$ without i having to cover the entire range from zero to $N-1$. One way to handle these problems is to define the pulse width by

$$PW = \frac{N\Delta\tau}{k_o SF} \quad (4.24)$$

where k_o is an integer used to adjust the pulse width down to the desired range. This results in

$$SF \cdot PW = \frac{N\Delta\tau}{k_o} \quad (4.25)$$

and by observing (4.23) the number of times that $\Delta t_o(i) = 0$ can then be found to be

$$N_{\Delta t_o(i)=0} = \gcd(N, k_o) \quad (4.26)$$

where $\gcd(x, y)$ is the greatest common divisor between x and y. The number of different offsets possible can then be found as

$$N_o = \frac{N}{\gcd(N, k_o)} \quad (4.27)$$

It can therefore be seen that $\Delta t_o(i) = 0$ when $i = 0 \pmod{N_o}$. The challenge is now to select a value of k_o that satisfies the range in which PW is desired. At the same time the value of k_o must ensure that N_o is "large enough" compared to the number of possible offsets needed to get a reliable averaging over the offsets introduced in the physical channel being modeled. Another issue is the fact that the average received power is dependent on i and in order to properly include this in the model a large value of k_o is preferred, but this may conflict with the wanted range of PW.

Examples of this procedure for selecting the pulse width will now be presented. As mentioned in Section 3.3 it is desired that the pulse width should be in the area of $PW = \frac{2}{3}ns$ in order for the signal bandwidth

to be around 3GHz. Considering a UWB system with $SF = 15$ and using $PW = \frac{2}{3}ns$, inserting in (4.24) yields $k_o = 20$ and the number of different offsets over which averaging is performed, can then be found from (4.27) as $N_o = 5$, which is too low too give adequate averaging. Instead selecting $k_o = 19 \vee k_o = 21$ will yield a PW very close to the desired value and $N_o = N = 100$, which is the highest possible value. Selecting $k_o = 21$ results in $PW = \frac{N\Delta\tau}{k_o SF} = \frac{100 \cdot 2ns}{21 \cdot 15} = \frac{40}{63}ns \simeq 0.6349ns$.

A similar approach can be followed when $SF = 16$ giving an initial value of $k_o = 18.75$. Rounding up to $k_o = 19$ will provide $N_o = N = 100$ and $PW = \frac{N\Delta\tau}{k_o SF} = \frac{100 \cdot 2ns}{19 \cdot 16} = \frac{25}{38}ns \simeq 0.6579ns$.

To illustrate how important this averaging over the different offsets is for the performance, a simulation of a single-user DS system using $PW = \frac{2}{3}ns$ and one using $PW = \frac{40}{63}ns$ has been performed. The result shown in Figure 19 clearly shows how this property has immense impact on the performance by producing a lot less ISI when $N_o = N$. Shown in the figure are also a simulation using $PW = \frac{2}{3}ns$ and a modified UWB channel model with a lower tap spacing of $\Delta\tau = 0.1ns$ using the same tap distribution.

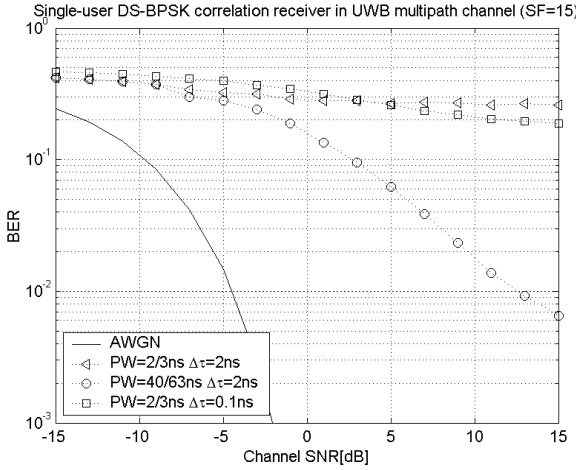

Figure 19: Simulation showing the importance of the pulse width when using the UWB channel model.

Having seen how the choice of pulse width affects the averaging properties of the channel model, one must consider how realistic the channel model really is. Basically what was done above was not adjusting the system to the physical channel, only to the model. It can therefore be considered a kind of reverse engineering from the model. Especially the situation where the pulse width is adjusted so that $N_o = N$ is very questionable as this results in none of the multipath components having a relative offset of zero and none of them will therefore have maximum impact on the decision. This is of course not realistic as some real multipath components will arrive at relative offsets equal to zero. On the other hand the situation where N_o is very low, such as when $PW = \frac{2}{3}ns$ and $SF = 15$, is also not very likely as too many multipath components will arrive at relative offsets of zero and the averaging is performed using only a small number of different offsets. A more realistic channel model will probably place itself somewhere between these two examples, but exactly where is not easily answered.

The heart of this problem with the channel model is the fact that $PW < \Delta\tau$, which has the unwanted consequence that the model does not support offsets being fractions of a pulse width at all delays. This is of course not reasonable, as realistic multipath components are not restricted to arrive at predetermined offsets dependent on their delay as outlined in (4.23). In order to produce a more realistic channel model, the value of $\Delta\tau$ must be small compared to the pulse width used in the system and the number of channel taps N will consequently have to be larger to keep $N\Delta\tau$ constant. Such models have not yet been produced, as the equipment to measure the impulse response with such high temporal precision does not yet exist and secondly it would be computationally more difficult to perform simulations using the model. The simulation depicted in Figure 19 tries to overcome this difficulty by lowering the tap spacing to $\Delta\tau = 0.1ns$. The problem with this approach is that measurements do not exist to support this extension. It is therefore not known if the distribution of the original model is valid if the tap spacing is lowered. Secondly the computational complexity goes up by a factor of 20, which makes simulations using this approach unrealistic.

The result is that the level of ISI introduced by the used channel model is highly dependent on the choice of the pulse width and a pulse width must therefore be agreed upon that produces a level of ISI that seem reasonable.

As a consequence of this, it has been decided to use $PW = \frac{2}{3}ns$ throughout the rest of this thesis, as this will resemble a situation closer to reality than if the pulse width is selected to fit the channel model. The results in the case of a DS system can therefore be considered conservative in the sense that the system will be influenced by more ISI than would most likely be incurred in a real physical channel.

4.2 Impact of ISI on TH and DS Systems

To see the effect of Inter Symbol Interference (ISI) introduced by the radio channel, the interference power from a single-user system at the output of a correlator receiver will now be calculated. By doing this it is possible to compare the performance of a TH and a DS system with each other and thus have an idea about if one is better than the other.

It is shown in Appendix G how an approximation to the ISI variance at the output of the matched filter is upper-bounded by the average of the squared code auto-correlation multiplied by a constant

$$\sigma_{ISI}^2 = E\left[e_{ISI}^2(n)\right] < E\left[\left|\int_{\tau_m}^{T_f} c^{(k)}(t-\tau_m)c^{(k)}(t)\,dt\right|^2\right] \cdot 2\sum_{i=1}^{N-1}\left(a_i^{(k)}\right)^2 \quad (4.28)$$

with the expectation taken with respect to the code $c^{(k)}(t)$ and the offset $0 \le \tau_m \le T_f$, which is modeled as a uniformly distributed random variable. The code $c^{(k)}(t)$ is assumed normalized so that

$$\int_0^{T_f} c^{(k)}(t)c^{(k)}(t)\,dt = SF \quad (4.29)$$

Under the assumption that the channel impulse response has a much longer duration than the pulse width, it is possible to model $e_{ISI}(n)$ as being a Gaussian random variable with zero mean and variance σ_{ISI}^2. It is therefore possible to directly find the resultant BER by viewing the ISI components as an increase in the noise level

$$BER_{ISI} = \frac{1}{2}erfc\left(\sqrt{\frac{SF \cdot N_s}{2}\left(\frac{1}{\sigma^2 + \sigma_{ISI}^2}\right)}\right) \quad (4.30)$$

with σ^2 being the noise variance given by

$$\sigma^2 = \frac{1}{SNR} \quad (4.31)$$

In the case of a TH system using random hopping codes, the spreading code may be modeled as

$$c^{(k)}(t) = \begin{cases} \dfrac{SF}{\sqrt{T_f}} & ,0 \le t \le \dfrac{T_f}{SF} \\ 0 & ,\dfrac{T_f}{SF} < \tau_m \le T_f \end{cases} \quad (4.32)$$

as there will on average be a period of T_f between each transmission. Applying the expectation of the right-hand expression of (4.28) over the code results in

$$TH: \quad E\left[\left|\int_{\tau_m}^{T_f} c^{(k)}(t-\tau_m)c^{(k)}(t)\,dt\right|^2\right] = \begin{cases} 2\left|\int_{\tau_m}^{\frac{T_f}{SF}} \frac{SF^2}{T_f}\,dt\right| = 2SF^2\left(1-\tau_m\frac{SF}{T_f}\right)^2 & , \quad 0 \le \tau_m \le \dfrac{T_f}{SF} \\ 0 & , \quad otherwise \end{cases} \quad (4.33)$$

The reason for multiplying the integral by two in the interval $0 \le \tau_m \le \dfrac{T_f}{SF}$, is that a similar amount of interference will come from the next bit in the interval $T_f - \dfrac{T_f}{SF} \le \tau_m \le T_f$. Integrating over τ_m gives

$$TH: \quad E\left[\left|\int_{\tau_m}^{T_f} c^{(k)}(t-\tau_m)c^{(k)}(t)\,dt\right|^2\right] = \frac{2SF^2}{T_f}\int_0^{\frac{T_f}{SF}}\left(1-\tau_m\frac{SF}{T_f}\right)^2 d\tau_m = \frac{2SF}{3} \quad (4.34)$$

If instead a DS system using random codes is used, the spreading code will have amplitude

$$\left|c^{(k)}(t)\right| = \sqrt{\frac{SF}{T_f}} \quad ,0 \le t \le T_f \quad (4.35)$$

with the sign being random from chip to chip. Computing the ensample average of (4.28) over the code gives

$$DS: \quad E\left[\left|\int_{\tau_m}^{T_f} c^{(k)}(t-\tau_m)c^{(k)}(t)\,dt\right|^2\right] = E\left[\left|\sum_{i=0}^{SF-1}\left(\int_{\tau_m}^{\frac{T_f}{SF}}\frac{SF}{T_f}\,dt + x_b x_b'\int_{\frac{T_f}{SF}}^{\frac{T_f}{SF}+\tau_m}\frac{SF}{T_f}\,dt\right)\right|^2\right] = \quad (4.36)$$

$$SF^2\left(1-\tau_m\frac{SF}{T_f}\right)^2 + \tau_m^2\frac{SF^2}{T_f^2}E\left[\left|\sum_{i=0}^{SF-1}\left(x_b x_b'\right)\right|^2\right] = SF^2\left(1-\tau_m\frac{SF}{T_f}\right)^2 + \tau_m^2\frac{SF^3}{T_f^2} \qquad ,0 \le \tau_m \le \frac{T_f}{SF}$$

and

$$DS: \quad E\left[\left|\int_{\tau_m}^{T_f} c^{(k)}\left(t-\tau_m\right)c^{(k)}\left(t\right)dt\right|^2\right] = E\left[\left|\sum_{i=0}^{SF-1} x_b x_b' \int_{\tau_m}^{\frac{T_f}{SF}+\tau_m} \frac{SF}{T_f}dt\right|^2\right] = E\left[\left|\sum_{i=0}^{SF-1}\left(x_b x_b'\right)\right|^2\right] = SF \qquad , \frac{T_f}{SF} < \tau_m \le T_f \tag{4.37}$$

The two independent Bernoulli random variables x_b and x_b' both have probability ½ causing neighboring chips to give a contribution to the variance equal the amount found in (3.42).

By integrating the results of (4.36) and (4.37) over τ_m in their respective intervals results in

$$DS: \quad E\left[\left|\int_{\tau_m}^{T_f} c^{(k)}\left(t-\tau_m\right)c^{(k)}\left(t\right)dt\right|^2\right] = \frac{SF^2}{T_f}\int_0^{\frac{T_f}{SF}}\left(1-\tau_m\frac{SF}{T_f}\right)^2 + \tau_m^2\frac{SF}{T_f^2}d\tau_m + \frac{SF}{T_f}\int_{\frac{T_f}{SF}}^{T_f} d\tau_m = \frac{4SF}{3}-\frac{2}{3} \tag{4.38}$$

Comparing the result of (4.34) with (4.38), it is seen that the DS system experiences approximately twice the interference power of that of the TH system. One system is therefore not to be preferred over the other when operating in a multipath radio channel. It is of most importance to be able to use a receiver that makes good use of the ISI components to increase the received energy pr. bit.

An interesting property of CDMA systems can also be seen directly from (4.34) and (4.38): As the desired signal at the output of the matched filter scales proportional to SF, letting $SF \to \infty$ will eliminate the ISI components completely, as it is the variance of the ISI that scales with SF. This rejection of the ISI makes CDMA systems interesting for wireless communication.

4.3 Summary

In this chapter the concept of multipath radio channel models have been introduced and a model usable for simulations is outlined. It is also discussed how the channel tap spacing is important in order to model the physical radio channel well and as the used channel model has a tap spacing being larger than the pulse width, an analysis of this effect is presented. Additionally a method to control the ISI offsets is offered, but it is argued that this method is based on reverse engineering from the channel model. This technique therefore does not give a realistic interaction between the channel model and the communication system, and the result can therefore not be considered realistic. Further, it is shown how TH and DS systems using correlator receivers are almost equally resistant to the ISI introduced by the radio channel. The factor determining the performance will consequently be the design of the receiver.

5 Advanced Receivers

In this chapter more advanced receivers will be evaluated capable of improving performance in the case of a multipath channel and/or multiple users communicating simultaneously on the channel. This is of great interest for UWB systems, as many users can potentially be within in range of each other and the receiver must consequently be able to operate under such conditions.

5.1 Optimal Receiver

The optimal receiver is the receiver that by definition achieves the best possible performance independent of the system and communication channel. To do this the receiver must be able to use all the information in the received signal concerning the bit of interest and then find the maximum-likelihood estimate of the bit. However to do this the receiver will also need complete knowledge about all other bits transmitted in the system as they will generally impact the decision on the bit of interest. The complexity of the optimal receiver is therefore extremely high and is consequently often of no practical interest, but may be used as a bound on the achievable performance. Often the performance achieved by the optimal receiver is so close to the single-user AWGN bound that this will be a very good approximation to the optimal receiver, except in situations with extremely good SNR. In this case the optimal receiver will eventually no longer provide lower BER as the SNR increases, because of the finite probability that users will completely cancel each other.

The straightforward way of doing optimal reception is to do an exhaustive search over all possible combinations of transmitted bits, but doing so will have a complexity that will be exponential in the number of bits. A system with K active users transmitting M bits in each frame will therefore have a complexity of $O\left(2^{KM}\right)$ operations to detect all bits. As the number of bits in a frame typically is in the order of 100-1000, this approach is nowhere near realistic. It is however possible to take advantage of system specific properties, making it possible to reduce the complexity to approximately $O\left(2^{KL}\right)$ pr. bit with L being the number of taps in the channel model. This lower complexity is achievable by using for instance Viterbi decoding of the transmitted sequence [15], but even this may become impractical for large KL.

5.2 RAKE Receiver

The RAKE receiver is a single-user detector that is designed to collect as much power as possible from the multipath components and then combine the contributions to estimate the transmitted symbol. Conceptually this can be done by using a correlator for each known multipath component and multiplying the outputs of each correlator by an appropriate weight. Another way to do exactly the same is shown in Figure 20 with the block termed MF being a matched filter and with T_l being the delay and w_l the weight of the l^{th} multipath component. Each of the L taps in the receiver is commonly known as a finger.

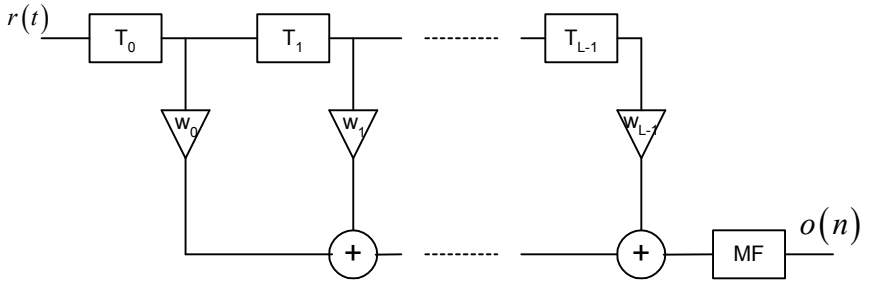

Figure 20: The structure of the RAKE receiver.

The optimal single-user weights in a multipath channel with Gaussian noise will now be derived. Assuming that the receiver has perfect knowledge of the channel impulse response of the desired user, the number of fingers is equal to the number of taps in the channel model meaning that $L = N$. In this case the delays T_l will exactly match those introduced by the channel such that

$$iT_{ch} + T_l = T_{RAKE} \quad , when \quad i = L - l - 1 \quad (5.1)$$

with T_{RAKE} being a fixed delay through the channel and the RAKE fingers. The output of the RAKE receiver is then

$$o(n) = \int_{nT_f + T_{RAKE}}^{(n+1)T_f + T_{RAKE}} \varphi^{(1)}\left(t - nT_f - T_{RAKE}\right)\left(\sum_{l=0}^{L-1} r\left(t - T_l\right) w_l\right) dt \quad (5.2)$$

and using (4.3) in (5.2) gives

$$o(n) = \sum_{l=0}^{N-1}\sum_{i=0}^{N-1} w_l a_i^{(1)} \int_{nT_f + T_{RAKE}}^{(n+1)T_f + T_{RAKE}} \varphi^{(1)}\left(t - nT_f - T_{RAKE}\right) s^{(1)}\left(t - iT_s - T_l\right) dt + z_{RAKE}(n) \quad (5.3)$$

with $z_{RAKE}(n)$ being the noise contribution at the output of the matched filter given by

$$z_{RAKE}(n) = \sum_{l=0}^{N-1} \int_{nT_f + T_{RAKE}}^{(n+1)T_f + T_{RAKE}} z(t) w_l \varphi^{(1)}\left(t - nT_f - T_{RAKE}\right) dt \Rightarrow z_{RAKE}(n) \sim N\left(0, SF\sigma^2 \sum_{l=0}^{N-1} w_l^2\right) \quad (5.4)$$

The reasoning behind the noise variance is, that when DS is used SF independent noise sources with variance σ^2 are added in the matched filter. In the case of TH, the value of the single noise source present is multiplied by $\sqrt{SF}$ in the matched filter.

Next the integral of (5.3) can be evaluated by using the auto-correlation of the template waveform and in order to simplify the expression, it is assumed that the auto-correlation function is zero except at lag zero

$$\rho(\tau) = \int_0^{T_f} \varphi^{(1)}(t)\varphi^{(1)}(t - \tau) dt = \begin{cases} SF & , \tau = 0 \\ 0 & , \tau \neq 0 \end{cases} \quad (5.5)$$

This simplification is made possible by the use of spreading codes, which would ideally have this property and often it is a good approximation. The expression in (5.3) can then be simplified as

$$o(n) \simeq SF\alpha_n^{(1)} \sum_{l=0}^{N-1} w_l a_{N-l-1}^{(1)} + z_{RAKE}(n) \quad (5.6)$$

and the SNR at the output of the matched filter is therefore

$$SNR_{RAKE} = \frac{SF\left(\sum_{l=0}^{N-1} w_l a_{N-l-1}^{(1)}\right)^2}{\sigma^2 \sum_{l=0}^{N-1} w_l^2} \quad (5.7)$$

Using Schwarz's inequality on (5.7) yields

$$SNR_{RAKE} = \frac{SF\left(\sum_{l=0}^{N-1} w_l a_{N-l-1}^{(1)}\right)^2}{\sigma^2 \sum_{l=0}^{N-1} w_l^2} \leq \frac{SF\sum_{l=0}^{N-1} w_l^2 \sum_{l=0}^{N-1}\left(a_{N-l-1}^{(1)}\right)^2}{\sigma^2 \sum_{l=0}^{N-1} w_l^2} = \frac{SF\sum_{l=0}^{N-1}\left(a_{N-l-1}^{(1)}\right)^2}{\sigma^2} \quad (5.8)$$

and in order to achieve equality in (5.8) the weights should be chosen as

$$w_l = c_{MRC} a^{(1)}_{N-l-1} \quad (5.9)$$

where c_{MRC} is an arbitrary constant. This type of combining is known as the Maximal Ratio Combiner (MRC) as it maximizes the SNR at the detection. Using MRC the resulting SNR is therefore given by the rightmost equation of (5.8) with

$$P_{tot} = \sum_{l=0}^{N-1} \left(a^{(1)}_{N-l-1} \right)^2 \quad (5.10)$$

being the total received power, which is the sum of the power received from each finger. A RAKE receiver with weights determined by MRC is therefore optimal as it achieves the same performance as a correlator would do in an AWGN channel with the same received power. It is also possible to use only a limited number of the multipath components available for combining although this will reduce performance proportional to the captured power as can be seen from (5.8). The multipath components with the highest received power should then be selected and MRC should then be applied only for these fingers. The computational complexity of the RAKE receiver is therefore $O(2L + 2M)$ pr. pulse where L is the number of fingers and M is the length of the matched filter.

In order to have a realistic view on the performance of the RAKE receiver, it must be remembered that the derivation is based on the assumption that the auto-correlation function is ideal. This is not realistic and therefore the performance of the RAKE will ultimately be limited by ISI, especially when the duration of the channel impulse response is long compared with the bit duration. Another issue not considered here is the channel estimation, as the channel model is assumed known in the derivation. Finally more than one user may be active in the channel and MAI will then also influence the performance.

5.2.1 Simulation

A number of simulations have been performed to observe the performance of the asynchronous RAKE receiver in the UWB multipath channel described in Chapter 4. The results can be seen in Figure 21 and Figure 22. The RAKE receiver simulated performs MRC with the channel coefficients specified by the channel model and the RAKE therefore has perfect knowledge of the channel. When the number of fingers used is less than the number of taps present in the channel model, fingers are allocated to the taps in order of decreasing power. The TH system uses hyperbolic hopping codes and the DS system uses large-set Kasami codes.

In Figure 21 the BER of the TH and DS single-user systems are plotted. This shows how the energy is spread out in time by the channel and how increasing the number of fingers can lower the BER. It should be noted that only a small gain is achieved by increasing the number of fingers beyond approximately 50 fingers as nearly all energy is already collected. It is also interesting to note how the ISI level is much lower in the TH system than in the DS system, which is a direct result of the DS system being exposed to a lot more ISI, as the number of different offsets that ISI components arrive at can be found by (4.27) to be $N_o = 5$.

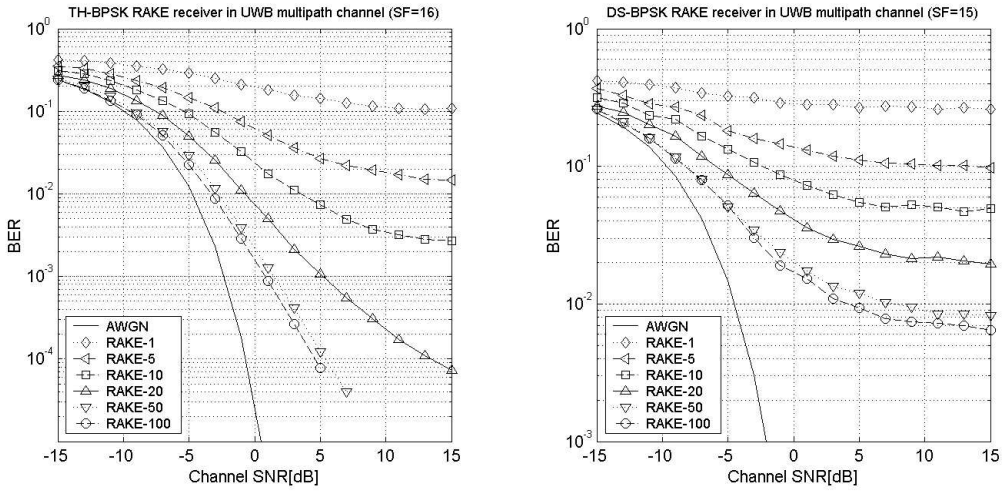

Figure 21: Single-user asynchronous systems in UWB multipath channel with varying number of fingers.

The performance of the RAKE receiver with 100 fingers (RAKE-100) when multiple users are active in the multipath channel is shown in Figure 22. The TH system quickly gets saturated by collisions, as was also the case in the AWGN channel. By contrast the DS system experiences less performance penalty from the added users, but the overall higher level of interference severely limits the performance of the DS system as the RAKE receiver is not capable of rejecting the interference.

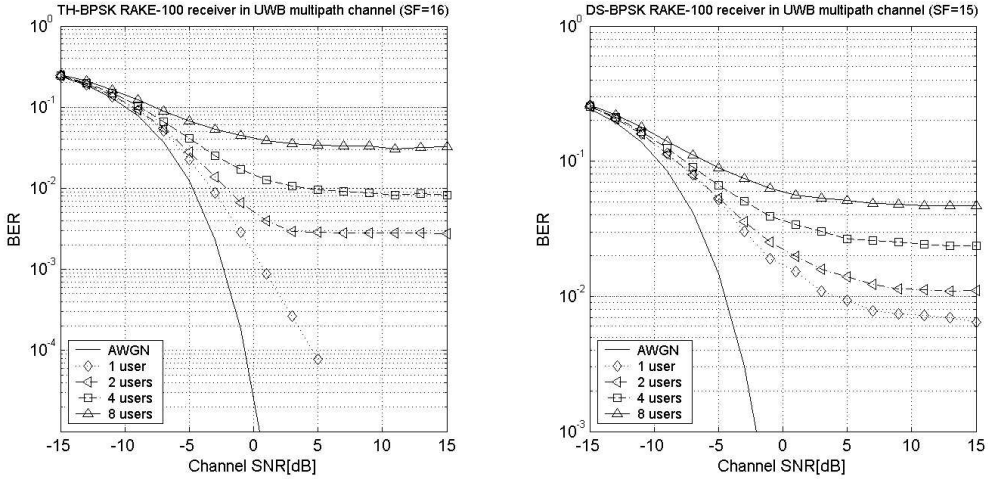

Figure 22: Asynchronous systems in UWB multipath channel with varying number of users.

5.3 Decorrelator Receiver

The decorrelator is a multi-user detector based on eliminating the correlation between users, hence the name. Using synchronous DS as an example the decorrelator achieves this by multiplying the matched filter outputs by the inverse of the correlation matrix

$$\hat{\boldsymbol{\alpha}}(n) = \operatorname{sgn}\left(\mathbf{R}^{-1}\mathbf{y}(n)\right) = \operatorname{sgn}\left(\overbrace{\mathbf{A}\boldsymbol{\alpha}(n)}^{Desired} + \overbrace{\mathbf{R}^{-1}\mathbf{z}(n)}^{Noise}\right) \quad (5.11)$$

It is easily seen from (5.11) that when $\sigma^2 \to 0$ the decorrelator is the optimal receiver, as it can separate the users completely from each other given that $\mathbf{R}$ is known. In order for this to work $\mathbf{R}^{-1}$ must of course

exist, but it can be proven that $\mathbf{R}$ is positive definite if the spreading codes used are linearly independent and therefore $\mathbf{R}^{-1}$ must also exist [15].

The covariance matrix of the noise term of (5.11) is

$$E\left[\left(\mathbf{R}^{-1}\mathbf{z}(n)\right)\left(\mathbf{R}^{-1}\mathbf{z}(n)\right)^{T}\right] = E\left[\mathbf{R}^{-1}\mathbf{z}(n)\mathbf{z}^{T}(n)\mathbf{R}^{-1}\right] = \mathbf{R}^{-1}\sigma^{2}\mathbf{R}\mathbf{R}^{-1} = \sigma^{2}\mathbf{R}^{-1} \quad (5.12)$$

and the noise variance for each user is therefore $\sigma^2 R_{kk}^{-1}$ with R_{kk}^{-1} being the k^{th} diagonal element of $\mathbf{R}^{-1}$ .The BER of each user can then be found as

$$BER_k = \frac{1}{2}erfc\left(\sqrt{\frac{A_k^{\,2}}{2\sigma^2 R_{kk}^{-1}}}\right) \quad (5.13)$$

It is shown in Appendix H that

$$tr\left(\mathbf{R}^{-1}\right) \geq \frac{K}{SF} \quad (5.14)$$

and R_{kk}^{-1} averaged over the K different codes will therefore be

$$E\left[R_{kk}^{-1}\right] \geq \frac{1}{SF} \quad (5.15)$$

Using (5.15) in (5.13) and comparing with the BER of the correlator receiver given in (3.22), it can be seen that the decorrelator will perform worse when performance is limited by noise instead of by MAI. This phenomenon is often referred to as noise enhancement, as the receiver raises the noise variance. This is of course not very desirable, but the decorrelator still represents a major improvement over the correlator receiver in the situation where the correlator is overcome by interference from other users. It is shown in [15 p. 236] that the decorrelator can be implemented as a modified matched filter and the increase in computational complexity over the correlator is therefore determined by the inversion of $\mathbf{R}$. The computational complexity is $O\left(K^3 + 2K^2 + 2M\right)$ pr. pulse when $\mathbf{R}^{-1}$ must be recalculated as it requires $O\left(K^3\right)$ finding the new $\mathbf{R}^{-1}$ and $O\left(2K^2\right)$ to get the new modified matched filter coefficients from this. If $\mathbf{R}^{-1}$ is unchanged, only $O\left(2M\right)$ operations are required for the filter.

A similar approach as described here for DS also apply to TH, but as the hopping pattern changes for every pulse, so does $\mathbf{R}$ and with it $\mathbf{R}^{-1}$. This means that a matrix inversion would have to be carried out for every pulse, which is far too complex. Instead it is possible to take advantage of the fact that it is only when users collide that a decorrelator should be used and a matrix inversion involving only these users should then be performed. The complexity is therefore $O\left(K_c^{\,3} + 2K_c^{\,2} + 2M\right)$ pr. pulse with K_c being the number of users colliding with the user of interest at this specific pulse. What makes this approach interesting is that K_c is very low on average, as more than one colliding user is a very rare event.

5.4 Linear Minimum Mean-Square Error Receiver

In order to improve on the performance of the decorrelator in the noise-limited scenario, the Linear Minimum Mean Square Error (LMMSE) receiver will now be introduced using a synchronous DS system as a baseline.
The LMMSE receiver is the multi-user detector that minimizes the error between the linear transform $\mathbf{M}$ of the matched filter outputs $\mathbf{y}(n)$ and the transmitted bits $\boldsymbol{\alpha}(n)$

$$\arg\min_{\mathbf{M}\in\mathbb{R}^{K\times K}} E\left[\left\|\boldsymbol{\alpha}(n)-\mathbf{M}\mathbf{y}(n)\right\|^2\right] \quad (5.16)$$

and the estimated bits can be found by hard-decision as

$$\hat{\boldsymbol{\alpha}}(n)=\operatorname{sgn}\left(\mathbf{M}\mathbf{y}(n)\right) \quad (5.17)$$

It is shown in [15 p. 293-294] that the linear transform $\mathbf{M}$ that satisfies (5.16) is

$$\mathbf{M}=\left(\mathbf{R}+\sigma^2\mathbf{A}^{-2}\right)^{-1} \quad (5.18)$$

where $\mathbf{A}$ contains all users amplitudes in the diagonal. The result given in (5.18) is very interesting, as it shows that the LMMSE achieves the same performance as the decorrelator when $\sigma\to 0$ and also converges to the correlator receiver as $\sigma\to\infty$. The LMMSE is therefore the optimal linear detector in both of these extremes. It is interesting to note however that the LMMSE receiver is not necessarily the receiver that minimizes the BER of the system, as there is no direct relationship between the mean-square error and the BER. On the other hand it is intuitive that the mean-square error is a reasonable metric to use. This is especially true if the soft-decision values obtained from the linear transform are utilized in a scheme to correct errors such as Viterbi decoding of convolutional codes.

As the case was for the decorrelator, the LMMSE receiver can be implemented using a modified matched filter and the computational complexity is therefore the same as the decorrelator. As the optimal filter coefficients $\mathbf{w}_o$ must minimize the mean-square error, they can be found by the Wiener-Hopf solution

$$\mathbf{R}\mathbf{w}_o=\mathbf{p}\Leftrightarrow\mathbf{w}_o=\mathbf{R}^{-1}\mathbf{p} \quad (5.19)$$

$\mathbf{R}$ is the covariance matrix of the received signal and $\mathbf{p}$ is the cross-correlation vector between the received signal and the information bit $\alpha^{(l)}(n)$ found by

$$\begin{aligned}\mathbf{R}&=E\left[\mathbf{r}(n)\mathbf{r}^T(n)\right]\\ \mathbf{p}&=E\left[\alpha^{(l)}(n)\mathbf{r}(n)\right]\end{aligned} \quad (5.20)$$

where $\mathbf{r}(n)$ contains samples of the received signal. Considering the general case of an asynchronous DS system operating in a multipath channel the covariance matrix can be found from (4.7) as

$$\mathbf{R}=\sum_{k=1}^{K}\sum_{m=-N_1}^{N_2}\mathbf{v}^{(k)}(m)\mathbf{v}^{(k)^T}(m)+\sigma^2\mathbf{I} \quad (5.21)$$

In a similar way the cross-correlation vector can be found as

$$\mathbf{p}=\mathbf{v}^{(l)}(0) \quad (5.22)$$

The optimal filter coefficients can now be found and the performance can be evaluated using these coefficients. The output of the filter is

$$\nu(n)=\mathbf{w}_o{}^T\mathbf{r}(n)=\overbrace{\mathbf{w}_o{}^T\mathbf{v}^{(l)}(0)}^{Desired}+\overbrace{e_{ISI}(n)+e_{MAI}(n)}^{Interference}+\overbrace{e_n(n)}^{Noise} \quad (5.23)$$

where $e_{ISI}(n)$, $e_{MAI}(n)$ and $e_n(n)$ are the contributions on the output from ISI, MAI and noise respectively. Both $e_{ISI}(n)$ and $e_{MAI}(n)$ are approximately Gaussian as shown in [25] and $e_n(n)$ is also Gaussian as the filter is linear. The BER of the LMMSE receiver can therefore be approximated by

$$BER_{LMMSE} = \frac{1}{2} erfc\left(\sqrt{\frac{\left| \mathbf{w}_o{}^T \mathbf{v}^{(l)}(0) \right|^2}{2\left(\sigma_{ISI}{}^2 + \sigma_{MAI}{}^2 + \sigma^2 \right)}} \right) \quad (5.24)$$

with σ^2 being the noise variance and

$$\sigma_{ISI}{}^2 = \sum_{m \neq 0} \left| \mathbf{w}_o{}^T \mathbf{v}^{(l)}(m) \right|^2 \quad (5.25)$$

$$\sigma_{MAI}{}^2 = \sum_{k \neq l} \sum_{m=-N_1}^{N_2} \left| \mathbf{w}_o{}^T \mathbf{v}^{(k)}(m) \right|^2 \quad (5.26)$$

Although the LMMSE receiver generally achieves somewhat better performance than the decorrelator [15 p. 302], it may not be worth the extra effort of estimating the noise variance and the amplitudes of all users. The perhaps biggest advantage of the LMMSE is the possibility of implementing it adaptively, as will be discussed next.

5.5 Adaptive LMMSE Receiver

An equivalent way of satisfying (5.16) is to minimize the mean-square error of the l^{th} user by adaptively estimating the modified matched filter instead of calculating it by means of a matrix inversion. The big advantage of this is that knowledge about users amplitudes, correlations, and the noise variance is not needed. Also knowledge about the shape of the monocycle is not needed, but instead a training sequence is required for the adaptation. This is the basic principle behind the adaptive LMMSE receiver that will be described in this section. Unfortunately the principle described in this section only applies to DS systems, as TH based systems do not have stationary multiple-access interference as required for convergence of the adaptive filter.

5.5.1 AWGN Channel

For a DS system operating in an AWGN channel the received signal is given by

$$r(t) = \sum_{k=1}^{K} A_k s^{(k)}(t) + z(t) \quad (5.27)$$

The length P column vector $\mathbf{w}^{(l)}(n)$ contains the coefficients of the transversal filter and the a posteriori error $e^{(l)}(n)$ between the transmitted bit $\alpha^{(l)}(n)$ and the filter output is therefore

$$e^{(l)}(n) = \alpha^{(l)}(n) - \mathbf{w}^{(l)^T}(n)\mathbf{r}(n) \quad (5.28)$$

In order to minimize the mean-square error, a vector $\mathbf{w}^{(l)}(n)$ must be found that minimizes the cost function

$$J(n) = E\left[\left(e^{(l)}(n) \right)^2 \right] = E\left[\left(\alpha^{(l)}(n) - \mathbf{w}^{(l)^T}(n)\mathbf{r}(n) \right)^2 \right] \quad (5.29)$$

It is readily seen from (5.29) that the mean-square error $J(n)$ is quadratic in $\mathbf{w}^{(l)}(n)$, meaning that no local minima exist. Two of the most used algorithms to minimize $J(n)$ are the Normalized Least Mean-Square (NLMS) and Recursive Least-Squares (RLS) algorithms [18].

The NLMS algorithm uses the fact that the global minimum of the mean-square error can be found by moving in the direction of steepest descent on the cost function as given by the opposite direction of the gradient $\nabla J(n)$ with respect to $\mathbf{w}^{(l)}(n)$. A recursion for updating the filter coefficients can now be written as

$$\mathbf{w}^{(l)}(n+1) = \mathbf{w}^{(l)}(n) - \frac{1}{2}\gamma(n)\nabla J(n) \quad (5.30)$$

with $\gamma(n)$ determining the step-size of the update in the n^{th} iteration. The gradient can be found from (5.29) as

$$\nabla J(n) = 2E\left[\left(\mathbf{w}^{(l)}(n)^T \mathbf{r}(n) - \alpha^{(l)}(n)\right)\mathbf{r}(n)\right] = -2E\left[\mathbf{r}(n)e^{(l)}(n)\right] \quad (5.31)$$

By skipping the expectation from (5.31) the gradient can be approximated by

$$\nabla J(n) \simeq \hat{\nabla} J(n) = -2\mathbf{r}(n)e^{(l)}(n) \quad (5.32)$$

Using this in (5.30) results in

$$\mathbf{w}^{(l)}(n+1) = \mathbf{w}^{(l)}(n) + \gamma(n)\mathbf{r}(n)e^{(l)}(n) \quad (5.33)$$

The step-size is given as

$$\gamma(n) = \frac{\mu}{a + \mathbf{r}(n)^T \mathbf{r}(n)} \quad , a << E\left[\mathbf{r}(n)^T \mathbf{r}(n)\right] \quad (5.34)$$

where μ is the step-size bound to the interval $0 < \mu < 2$ by stability [18]. The constant a is introduced to reduce the impact of gradient noise when $\mathbf{r}(n)^T \mathbf{r}(n)$ attains a small value. The complexity of the NLMS update is $O(4P)$ pr. bit pr. user.
Selection of the step-size parameter μ is a trade-off between convergence speed and the residual interference. Generally a "small" value of μ is desired to minimize residual interference, but this will cause very slow convergence, which again will require a lot of training bits. As this is not desired a somewhat larger value of μ should be selected as a trade-off between convergence speed and the level of residual interference that can be tolerated by the system.

The Recursive Least-Squares (RLS) algorithm minimizes the mean-square error $J(n)$ by recursively estimating the filter coefficients $\mathbf{w}^{(l)}(n)$ that give rise to the smallest least-square residue at each iteration

$$\arg\min_{\mathbf{w}^{(l)}(n)} \left(\sum_{i=1}^{n} e^{(l)}(i)^2 \right) \quad (5.35)$$

It is therefore expected that the performance of the RLS algorithm will always be better than that of the NLMS algorithm when used in the adaptive LMMSE receiver, but as mentioned before there is no direct link between minimizing the mean-square error and minimizing the BER. The RLS update can be written as

$$\mathbf{w}^{(l)}(n) = \mathbf{w}^{(l)}(n-1) + \overbrace{\mathbf{\Phi}^{-1}(n)\mathbf{r}(n)}^{\mathbf{k}(n)} \varepsilon^{(l)}(n) \quad (5.36)$$

with $\mathbf{\Phi}(n)$ being the correlation matrix defined by

$$\mathbf{\Phi}(n) = E\left[\mathbf{r}(n)\mathbf{r}(n)^T\right] \quad (5.37)$$

and $\varepsilon^{(l)}(n)$ being the a priori estimation error given by

$$\varepsilon^{(l)}(n) = \alpha^{(l)}(n) - \mathbf{w}^{(l)^T}(n-1)\mathbf{r}(n) \quad (5.38)$$

The gain vector $\mathbf{k}(n)$ is simply the received signal $\mathbf{r}(n)$ transformed by the inverse correlation matrix $\mathbf{\Phi}^{-1}(n)$.

As a result of the matrix inversion, the RLS algorithm has a complexity of approximately $O(P^3)$ pr. bit pr. user. To reduce this figure, the following recursion is used to reduce the complexity to approximately $O(2P^2)$ pr. bit pr. user [18]

$$\mathbf{k}(n) = \frac{\mathbf{\Phi}^{-1}(n-1)\mathbf{r}(n)}{1+\mathbf{r}(n)^T\mathbf{\Phi}^{-1}(n-1)\mathbf{r}(n)} \quad (5.39)$$

$$\mathbf{\Phi}^{-1}(n) = \mathbf{\Phi}^{-1}(n-1) - \mathbf{k}(n)\mathbf{r}(n)^T\mathbf{\Phi}^{-1}(n-1) \quad (5.40)$$

The RLS algorithm should be initialized with a priori coefficients, or if such are not available

$$\mathbf{w}^{(l)}(0) = \mathbf{0} \quad (5.41)$$

and $\mathbf{\Phi}^{-1}(0)$ being the inverse of the correlation matrix given by (5.37). This is however often too complex as it requires a matrix inversion and instead it can be initialized as

$$\mathbf{\Phi}^{-1}(0) = \frac{\delta}{E\left[\mathbf{r}(n)^T\mathbf{r}(n)\right]}\mathbf{I} \simeq \frac{\delta}{\mathbf{r}(0)^T\mathbf{r}(0)}\mathbf{I} \quad (5.42)$$

with the constant δ being rather important in determining the properties of the algorithm. The reason is that the algorithm now solves the following optimization problem [18, 24]

$$\arg\min_{\mathbf{w}^{(l)}(n)}\left(\left(\mathbf{w}^{(l)}(n)-\mathbf{w}^{(l)}(0)\right)\mathbf{\Phi}(0)\left(\mathbf{w}^{(l)}(n)-\mathbf{w}^{(l)}(0)\right)^T + \sum_{i=1}^{n} e^{(l)}(i)^2\right) \quad (5.43)$$

and the value of δ is therefore important for both the convergence speed and the converged coefficients. Selecting a small value of δ will constrain the coefficients to stay close to the initial value $\mathbf{w}^{(l)}(0)$, whereas a large value of δ does not. The impact on convergence speed can be seen from the fact that when $\delta < 1$ the relative update to $\mathbf{\Phi}^{-1}(n)$ in (5.40) will be smaller than if $\delta > 1$. This is especially true if no forgetting factor [18] is introduced in the RLS algorithm, which will be assumed here as the channel is thought stationary.

The effect of this choice is, that if a small value of δ is selected and assuming that the coefficients are initialized to zero, the resulting filter closely assembles the matched filter solution, which is optimal when only a single user is present. On the other hand selecting a large value of δ, the regularization effect of δ on $\mathbf{\Phi}(n)$ will diminish and an amplification of the noise, as was the case with the decorrelator, will occur if the eigenvalues of $\mathbf{\Phi}(n)$ are small. This will happen if $\mathbf{\Phi}(n)$ has low rank, as is typically the case when only a few users are active in the system. The resultant filter will potentially eliminate more of the MAI, but the solution will not be optimal with only a single user. The choice of δ is therefore a trade-off between single- and multi-user performance, just like the step-size μ was a similar trade-off in the NLMS algorithm.

5.5.2 UWB Multipath Channel

Another great advantage of this adaptive receiver is the ability to include channel estimation directly in the receiver. When the receiver operates in a multipath channel with N taps, the received signal $r(t)$ is given by (4.3) as

$$r(t) = \sum_{k=1}^{K} \sum_{i=0}^{N-1} a_i^{(k)} s^{(k)} \left(t - iT_{ch}\right) + z(t) \quad (5.44)$$

In order to collect all the bit energy spread out in time by the channel, the length P must be equal to the integer given by

$$P = \left\lceil \frac{T_f + (N-1)T_{ch}}{T_s} \right\rceil \quad (5.45)$$

making the filter length P typically much larger than in the AWGN case when the UWB multipath channel is considered. Using the channel model from Chapter 4, the filter length can be written as

$$P = \frac{N_{sp}(N-1)\Delta\tau}{PW} + N_{sp}SF \quad (5.46)$$

This much larger filter length may make the filter too complex and it may therefore be necessary to shorten the length of the filter. As a consequence only the first part of the bit energy spread out by the channel is collected and performance will therefore suffer. To account for this shortening of the filter compared to P, a new filter length P_{short} is introduced as

$$P_{short} = \lfloor \psi P \rfloor \quad , 0 < \psi \leq 1 \quad (5.47)$$

making it possible to adjust the duration of the window in which the filter can capture energy from a given bit. It is therefore conceptually much like adjusting the number of fingers in the RAKE receiver, but the RAKE fingers are not confined to time intervals adjacent to one another. As adjusting the filter length affects how many bits after the bit of interest will spill energy into the receiver, the value of N_2 used in (5.21) and (5.26) must be changed from the value given in (4.8) to

$$N_2 = \left\lceil \frac{\psi(N-1)T_{ch}}{T_f} \right\rceil \quad (5.48)$$

Using the new filter length P_{short}, the operation of the adaptive LMMSE receiver in the multipath channel works like in the AWGN channel. This makes it an attractive candidate for multi-user DS systems in multipath channels at the expense of computational complexity in the receiver. It is therefore desirable to reduce the complexity as will be discussed next.

5.5.3 Reduced Complexity Adaptive LMMSE

The large computational load in the receiver comes from two parts namely the filtering operation and the adaptation of the filter taps. The filtering has complexity $O\left(2\lfloor \psi P \rfloor\right)$ and the adaptation has complexity $O\left(4\lfloor \psi P \rfloor\right)$ in the case of NLMS and $O\left(2\lfloor \psi P \rfloor^2\right)$ in case of RLS. The most interesting algorithm is the RLS algorithm as it generally has better performance, which will also be confirmed by simulations later. The problem however is its large computational complexity. A large gain can therefore be achieved if the complexity of the adaptation can be lowered, as this is the far most dominating element of the computational complexity in the case of RLS adaptation.

Two different strategies are available to reduce the complexity of the adaptive filter. The first is to simply reduce the number of filter taps, which will also reduce the complexity of both the filtering and

adaptation. This operation is also known as sub-sampling as the Nyquist sampling criterion will no longer be satisfied. Secondly the RLS adaptation algorithm can be improved so that it requires less computation. A special class of RLS algorithms known as fast RLS algorithms is of interest here, as they scale linearly in complexity with ψP. These two methods can also be combined to provide even lower complexity if the situation allows for it.

Sub-sampling

To implement sub-sampling a new spacing between samples must be specified as $T_s' = N_{sub}T_s$, where N_{sub} is the decimation factor. The received signal is now sampled as

$$r_i(n) = r\left(nT_f + iT_s'\right) \quad ,i = 0,1,...,\left\lfloor \frac{\psi P}{N_{sub}} \right\rfloor - 1 \quad (5.49)$$

and the filter length is reduced to $\left\lfloor \frac{\psi P}{N_{sub}} \right\rfloor$. The price paid for this reduction in complexity is increased sensitivity to noise and reduced rejection of interference as the number of taps in the filter becomes smaller.

Fast RLS

It is possible to take advantage of the shifting structure in a transversal filter to reduce the complexity of an RLS update from $O\left(2M^2\right)$ to $O(8M)$ by using a Fast Transversal Filter (FTF), where M is the number of taps in the filter [23]. The problem with this approach is that it is not necessary in the adaptive LMMSE receiver for the filter output to be updated for every single shift throughout the delay line. Instead only one in every $\left\lfloor \frac{T_f}{N_{sub}T_s} \right\rfloor$ shifts are necessary, but to use a FTF all the shifts must be performed. The computational reduction by using a FTF over the traditional RLS for the adaptive LMMSE is therefore

$$G \simeq \frac{2\left(\frac{\psi P}{N_{sub}}\right)^2}{\left(\frac{T_f}{N_{sub}T_s}\right)8\left(\frac{\psi P}{N_{sub}}\right)} = \frac{\psi P}{4N_{sub}N_{sp}SF} \quad (5.50)$$

as $\frac{T_f}{T_s} = N_{sp}SF$, where N_{sp} is the number of samples pr. monocycle. To have an idea about the size of this complexity reduction, evaluating (5.50) using $PW = \frac{2}{3}ns$, $N_{sp} = 13$, $SF = 15$ and using the UWB channel model resulting in $P = \frac{N_{sp}(N-1)\Delta\tau}{PW} + N_{sp}SF = 4056$ gives

$$G \simeq \frac{4056\psi}{4N_{sub}\cdot 13\cdot 15} = 5.2\frac{\psi}{N_{sub}} \quad (5.51)$$

This is a very disappointing result as it shows that only a small reduction in computational complexity is possible in this scenario. When $\frac{\psi}{N_{sub}} < \frac{1}{5.2}$ it even has a higher complexity than using the traditional RLS. Situations do exist where the reduction can be significant such as when the product $N_{sub}N_{sp}SF$ is

low, but such situations are not of interest here. The idea of using an FTF RLS algorithm is therefore not considered further.

5.5.4 Simulation

A series of simulations using the adaptive LMMSE receiver in a DS based UWB system will now be presented. First simulations in the AWGN channel will be performed and afterwards the UWB multipath channel model will be used, making it possible to assess the performance of this receiver type under different circumstances. The system simulated is an asynchronous DS system using large-set Kasami codes with $SF = 15$. All users are received with equal power levels unless otherwise stated, i.e. $N_{SIR} = 1$.

As mentioned before the values of μ and δ greatly influences the performance of the NLMS and RLS algorithms respectively. To see how these two parameters affect the convergence of the two algorithms, the error is plotted in Figure 23 as a function of the iteration normalized to the filter length for both algorithms using different values of μ and δ.

Looking at single-user NLMS it is seen, as expected, that a large value of μ gives faster convergence, but the residual error is the same. The reason is that the SNR is "only" 20dB and the smaller value of μ can therefore not be justified as the theoretically lowest value of the error variance can be found to be $\frac{1}{SF \cdot SNR} = -32dB$ in a single-user AWGN system. Note also that when $\mu = 10^{-3}$ convergence is not achieved within the number of iterations simulated. The NLMS convergence with 15 users is significantly slower as strong interference is present and convergence is no longer achieved within the simulated iterations with $\mu = 0.1$.

Turning to the convergence of the RLS algorithm, it is seen how the value of δ influences the speed of convergence as discussed before. Comparing the speed of convergence with NLMS, it should be noted how the RLS with large δ converges faster than the NLMS when many users are active in the channel. On the other hand NLMS with large μ is potentially faster in the single-user case. The reason for this is that the NLMS algorithm can potentially achieve the matched filter solution in a single iteration and this is the optimal solution for this scenario. The same is not true for RLS, as it must first achieve a reliable estimate of the inverse correlation matrix.

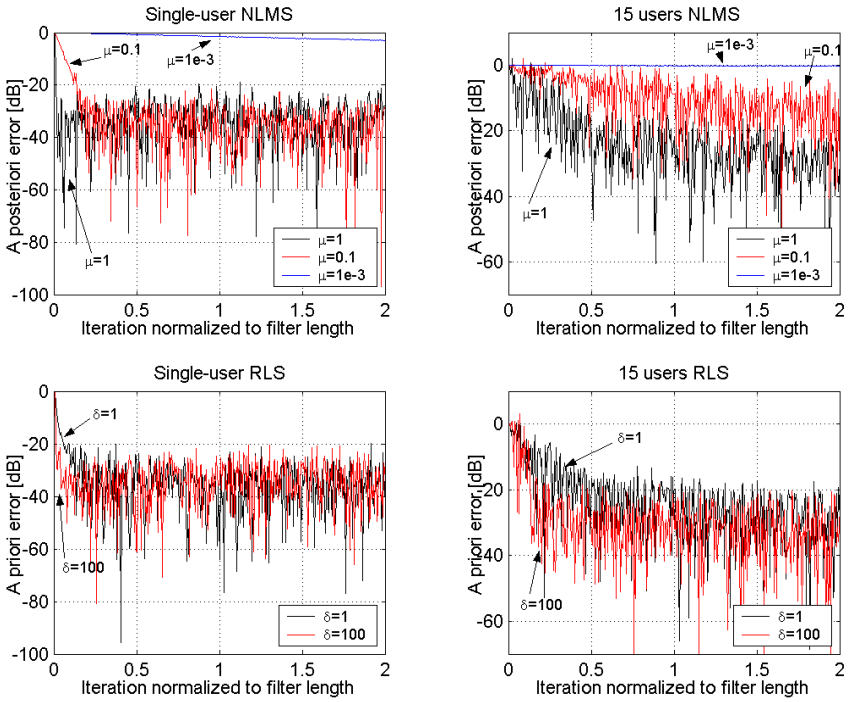

Figure 23: Convergence of the adaptive receiver in the AWGN channel at SNR=20dB.

In Figure 24 the BER of the asynchronous adaptive LMMSE receiver in the AWGN channel is plotted, where all adaptations are done with as many training bits as there are taps in the filter, which is given by $N_{ite} = P$. This is done in order to include the convergence speed in the comparison and an adaptation will therefore not perform well if it does not manage to converge using this number of training bits. Also included in the figure is the BER of the ideal LMMSE receiver with 30 users as given by the expression in (5.24).

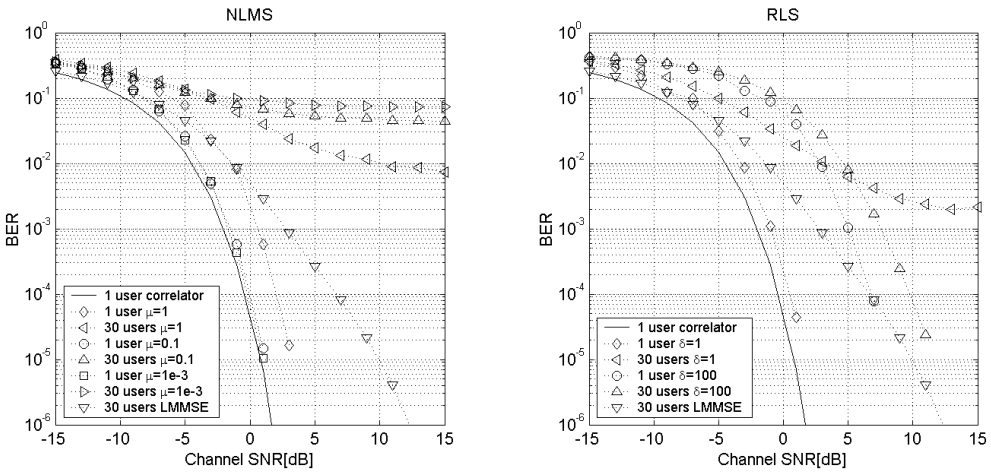

Figure 24: Performance of adaptive LMMSE in the AWGN channel. All users are received with equal power.

Starting with the NLMS adaptation with $\mu = 1$ it is noticeable how performance suffers approximately 2dB in the single-user case compared to the correlator. The reason is the residual error present and the filter coefficients do therefore not closely assemble the matched filter. When lowering the step-size to $\mu = 0.1$ the performance comes closer to the correlator as the residual error is decreased. Interestingly,

the performance with $\mu = 10^{-3}$ is the best. This may come as a surprise as the filter does not reach convergence in this case, but it is explained by the fact that convergence is not needed in the high SNR single-user NLMS case, as the filter will be the matched filter after a single iteration independent of the step-size.
Loading the NLMS system with more users, it is seen how convergence of the filter is now crucial for performance. The case with $\mu = 10^{-3}$, which does not nearly reach convergence, performs only a little bit better than the correlator simulated in Figure 12. Increasing the step-size to $\mu = 0.1$, the filter does not converge completely in the high SNR region, but as can be seen it does improve performance somewhat. Using $\mu = 1$ the NLMS converges fully and a little performance gain is achieved. This shows the trade-off in the NLMS algorithm between convergence speed and residual error dependent on the load of the system.
In the RLS case it is noticed how the single-user performance gets closer to that of the correlator as δ becomes smaller and how a large penalty of about 7dB is incurred when $\delta = 100$. The reason for this is the noise amplification present when the regularization term of $\mathbf{\Phi}(0)$ vanishes. On the other hand the RLS algorithm is substantially better at rejecting the multiple-access interference than NLMS and using a large value of δ help improve performance in the interference-dominated region, as the filter coefficients becomes less constrained by regularization. As a result of this the BER approaches the ideal LMMSE performance at high SNRs when $\delta = 100$. The choice of δ is therefore a similar trade-off between convergence speed and residual error dependent on the load of the system, as is the case with the step-size of the NLMS algorithm.

In order to get an understanding of how the performance of the adaptive LMMSE receiver scales with the number of users, a series of simulations have been performed as shown in Figure 25. This makes it possible to do a direct comparison with the correlator receiver depicted in Figure 12, showing how the LMMSE receiver rejects the multiple-access interference much better. However the result of RLS adaptation using $\delta = 100$ may be a bit surprising as the BER drops when more users are added to the system. This may seem strange, but is actually a result of the fact that almost no regularization is applied to $\mathbf{\Phi}(0)$. This results in the rank of $\mathbf{\Phi}(n)$ being low when only a few users are active in the system and a large noise amplification is observed, but as more users are added the rank of $\mathbf{\Phi}(n)$ increases leading to less noise amplification. Increasing the number of users further, the noise amplification will eventually no longer be the limiting factor and the BER starts to rise when more users are added, because of the increasing multiple-access interference.

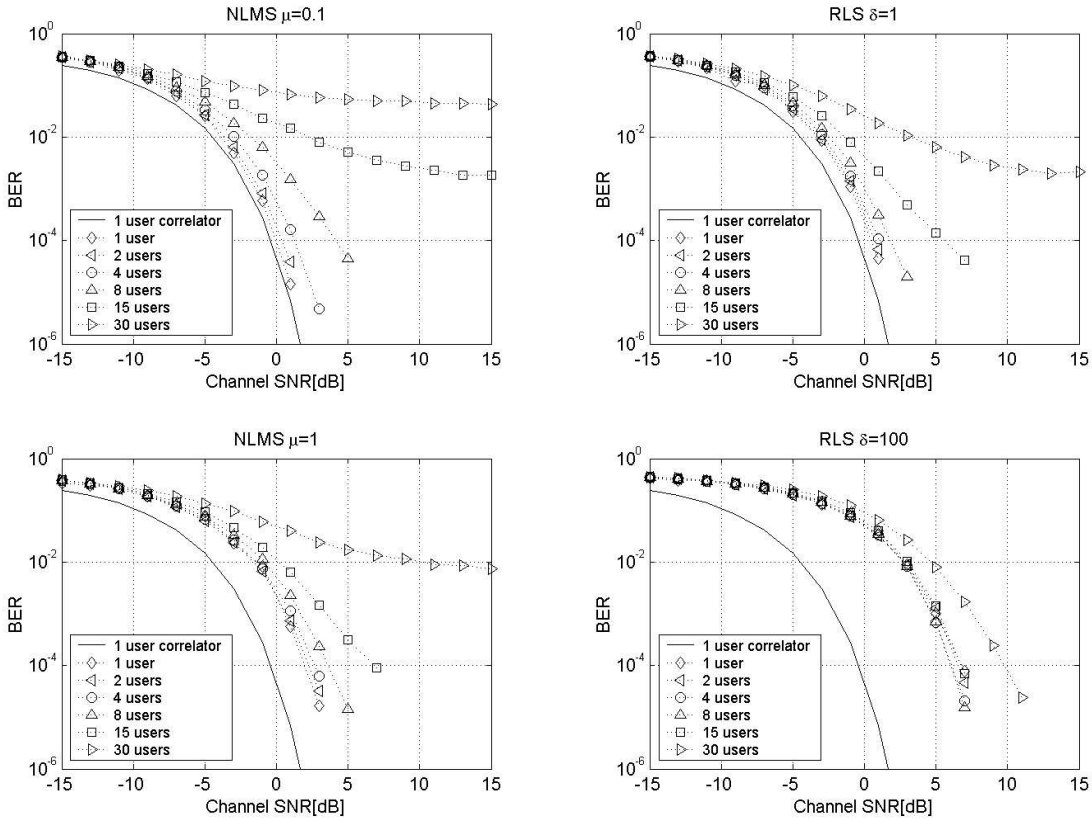

Figure 25: Performance of adaptive LMMSE in the AWGN channel with varying number of users, which are all received with equal power.

To compare the impact of the near-far effect on the adaptive LMMSE receiver with the correlator receiver, the simulations have been rerun, but this time all other users than the desired are received at 10 times higher power, i.e. $N_{SIR} = 10$. The resultant SIR is therefore

$$SIR_{UWB} = \frac{1}{10(K-1)} \quad (5.52)$$

where K is the total number of active users. The results are shown in Figure 26 and it is evident that the adaptive LMMSE receiver performs much better than the correlator and how the increased interference level affects the NLMS adaptation most.

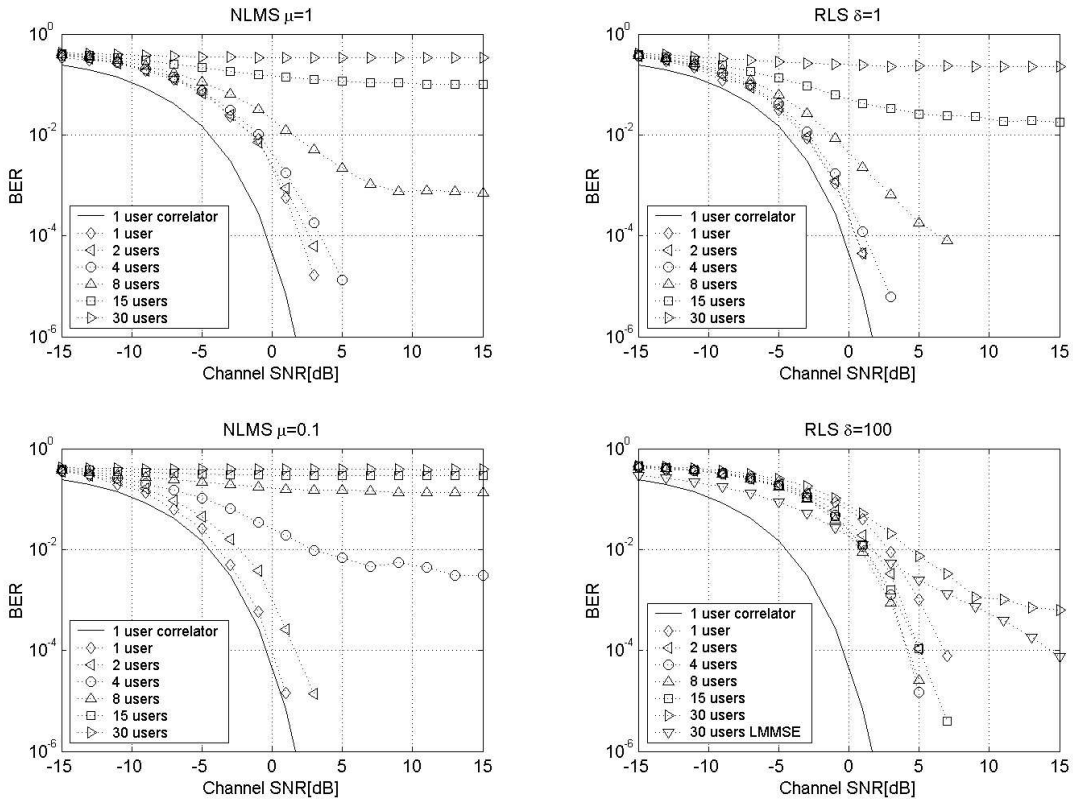

Figure 26: Performance of adaptive LMMSE in the AWGN channel with varying number of users with N_{SIR}=10.

Performance of the adaptive LMMSE receiver in the UWB multipath channel will now be evaluated assuming that all users are received with the same power. The system is the same as in the AWGN channel, but in this case each user is convolved with their respective channel impulse response. As in the AWGN case it is important to study the convergence of the adaptive filter, which is plotted in Figure 27 for $\psi = 1$ and $N_{ite} = P$. Again it is seen how the value of μ and δ affect the speed of convergence. In the single-user case it should be noted how the convergence speed of NLMS has suffered compared to the AWGN channel, as the optimal solution can no longer be reached in a single iteration, whereas the convergence of the RLS is virtually unaffected. Loading the system with more users it is seen, like in the AWGN case, how the RLS is superior in rejecting the multiple-access interference even when operating with this high level of dispersion being introduced by the UWB multipath channel.

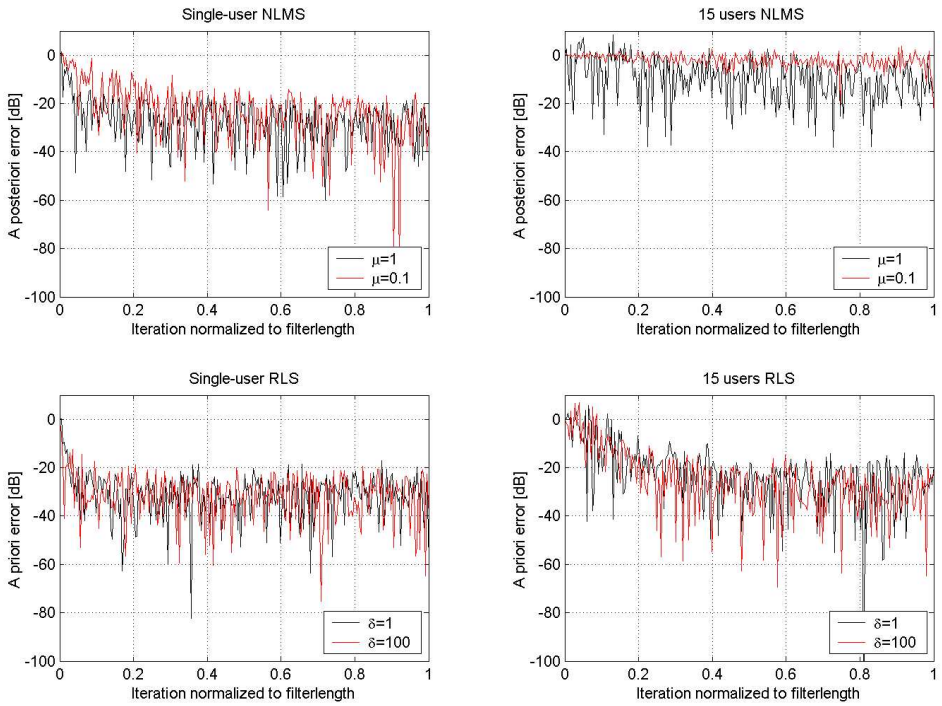

Figure 27: Convergence of the adaptive LMMSE receiver in the UWB multipath channel at SNR=20dB and Ψ=1.

The BER of the NLMS adaptive LMMSE receiver is plotted in Figure 28 for varying values of ψ . First it should be noted how the receiver with $\mu = 0.1$ outperforms that using $\mu = 1$ in the single-user case, as convergence is very fast and a low value of the step-size will thus produce less residue after convergence. Note also how the performance using $\mu = 0.1$ approaches the LMMSE lower bound.
Increasing the number of users to 15, the situation reverses. Using $\mu = 1$ is now best as this receiver will come closer to convergence within the simulated iterations, but it is not able to reach the LMMSE lower bound.

The case of RLS adaptation is plotted next in Figure 29. In the single-user case the value of δ is no longer as important as in the AWGN case, because noise-amplification is no longer the only crucial thing. In this scenario the receiver must also correctly estimate the channel impulse to collect sufficient energy. Turning to the case of 15 users, the value of δ again becomes very significant as in the AWGN channel. This is because performance is strongly influenced by the receiver's ability to reject the multiple-access interference, which is heavily influenced by the choice of δ, as was shown for the AWGN channel in Figure 25 and Figure 26. The case of $\delta = 100$ actually works so well in the multipath channel that the LMMSE lower bound is very close to the observed BER in both single- and multi-user scenarios.

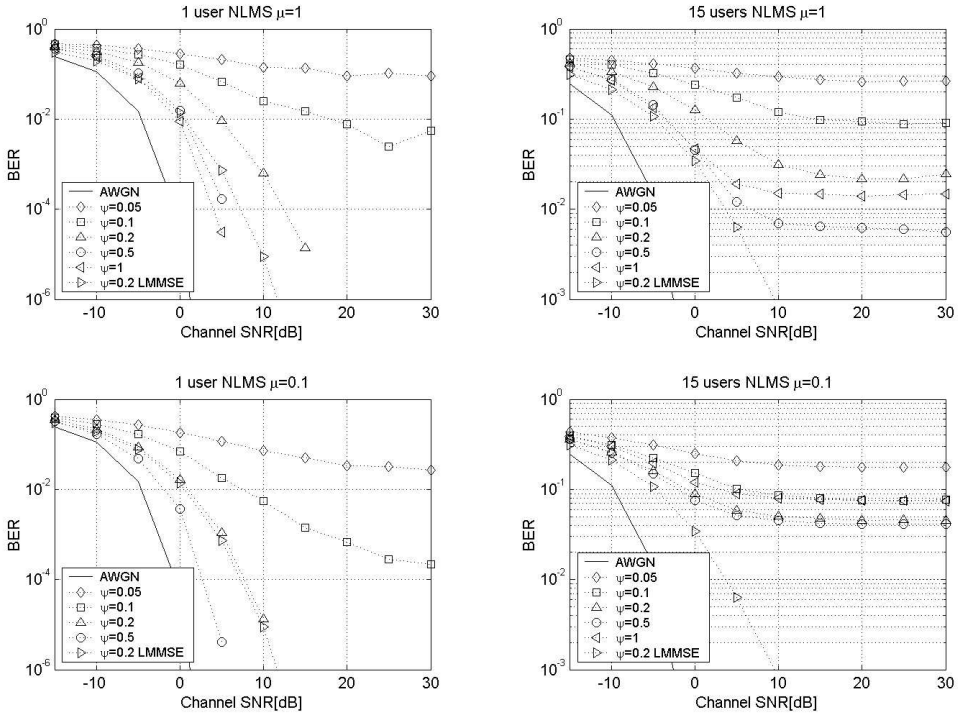

Figure 28: Performance of NLMS adaptive LMMSE in the UWB multipath channel with N_{ite}=P.

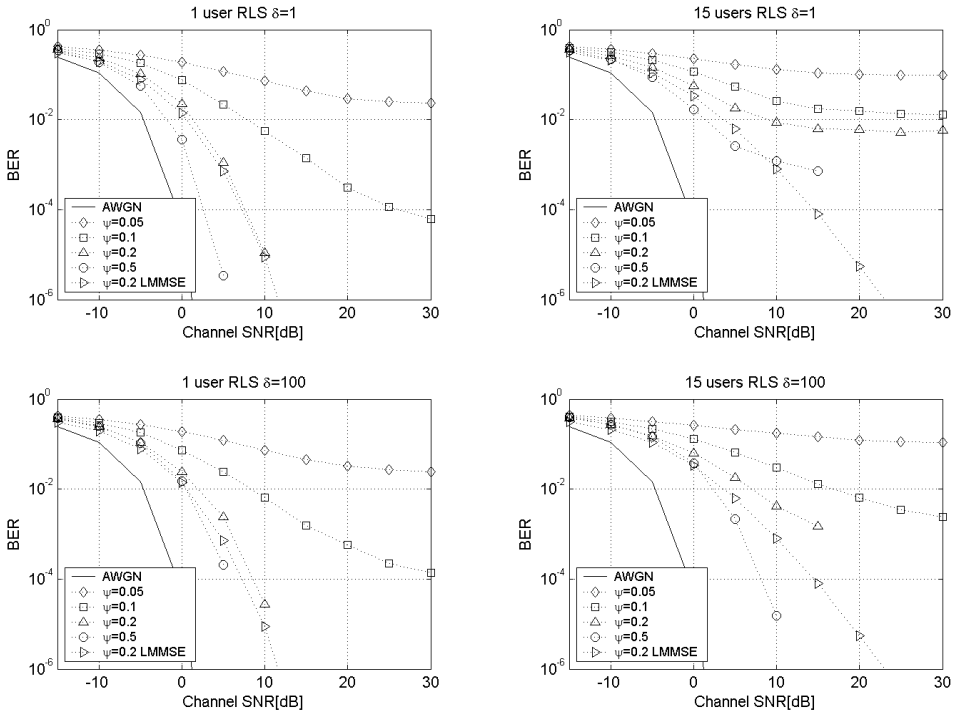

Figure 29: Performance of RLS adaptive LMMSE in the UWB multipath channel with N_{ite}=P.

To illustrate the effect of the subsampling applied to the adaptive RLS LMMSE receiver, simulations have been run using a subsampling of $N_{sub} = 4$ as shown in Figure 30. Systems with a single user and 15

users have been simulated for $\delta = 1$ and $\delta = 100$, but the case of a single-user system using $\delta = 100$ has been left out, as it has been seen from previous simulations that this combination is of little use. As mentioned above, the subsampled system suffers in two ways. The first problem is the increased sensitivity to noise as less noise is rejected from the samples by filtering in the receiver. Secondly the receiver looses some of its ability to effectively eliminate the multiple-access interference as can be seen from the BER floors. This is the price paid for lowering the complexity of the RLS update by $N_{sub}^2 = 16$.

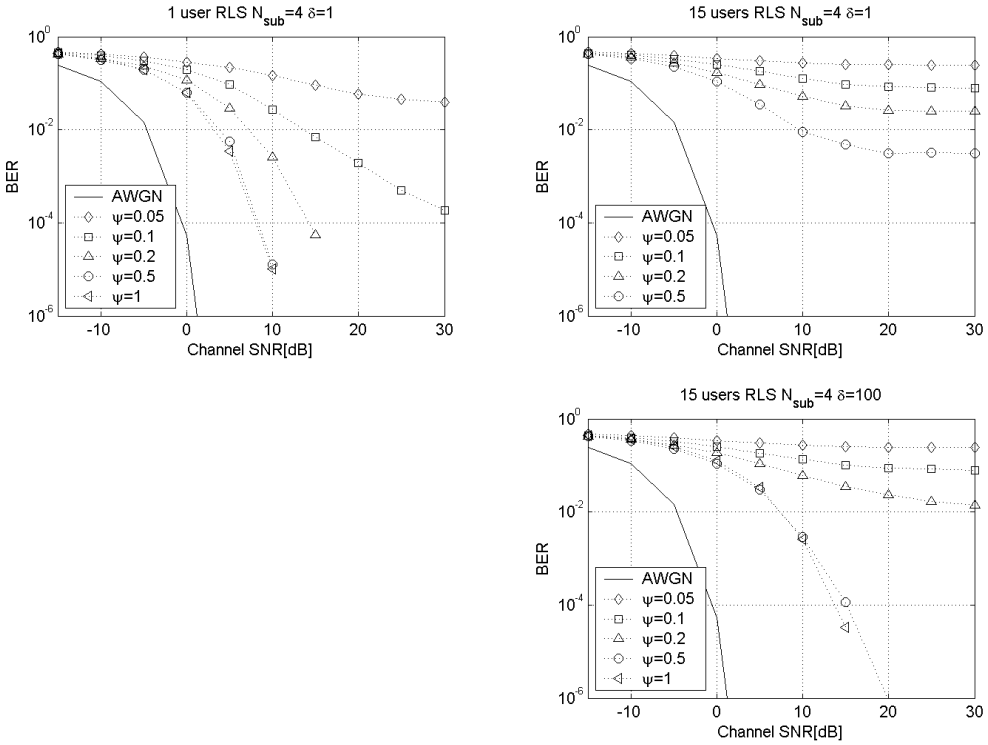

Figure 30: Performance of RLS adaptive LMMSE in the UWB multipath channel with N_{ite}=P and N_{sub}=4.

5.6 Summary

In this chapter some of the most important receiver structures for UWB operation are studied. First, the RAKE receiver is described and the MRC weights are derived and simulations using this receiver are performed. Next, the decorrelator and LMMSE receivers are studied and finally the adaptive version of the LMMSE receiver is examined and extensively simulated. It is shown how the adaptive LMMSE receiver using RLS adaptation is able to yield good performance in both single- and multi-user scenarios over the UWB multipath channel. Furthermore, it is shown how reducing the complexity by sub-sampling has the drawback of reducing the performance of the system.

Much focus in this chapter is on the adaptive LMMSE receiver and for good reason as it provides good performance in both the AWGN and multipath channel scenarios. One of the drawbacks of this solution is that TH system are excluded, as the adaptive nature of the implementation does not work with the non-stationary pseudo-random hopping used in TH systems. This is a major problem for the use of TH systems in which multiple users are active, as no good alternative to the adaptive LMMSE receiver exist today for TH operation.

6 Synchronization

The task of getting the receiver locked onto the transmitted signal and staying synchronized over time is an often-overlooked topic compared to modulation and demodulation. Nonetheless, it is absolutely crucial to the performance of the system that this is done properly; otherwise the performance of the system will drop dramatically compared to the synchronized case.

This chapter is an introduction to some of the synchronization problems typically present in communication systems. For the case of synchronizing UWB communication systems, a novel method based on the adaptive LMMSE receiver is presented and simulated.

6.1 General Synchronization

The job of synchronizing the receiver with the transmitter is usually separated in two. Initially the receiver knows nothing about the timing parameters used by the transmitter and will therefore have to acquire these. This phase is therefore often known as acquisition. Afterwards, when the receiver has locked onto the transmitted signal, it must continuously track the signal as the timing parameters will change over time. This is due to possible changes in the radio channel, but will also stem from the fact that the clocks used in the transmitter and receiver will not be exactly equal and the receiver clock will consequently drift over time compared to the transmitter. To have an idea about this drift, the typical oscillators used today have a precision in the area of 10-100ppm meaning that after 10^4-10^5 samples an offset of up to one sample may be observed between transmitter and receiver.

In order to correct this changing timing offset a correction $\pi(m)$ must be made to the sample timing of sample number m in the receiver such that the transmitter and receiver are running at the same clock. The true value of $\pi(m)$ is of course not available to the receiver and an estimated value $\hat{\pi}(m)$ must be used instead. In Figure 31 a typical transceiver model is shown including the effect of timing offset by controlling the sample time with the estimated correction $\hat{\pi}(m)$. The task of the digital receiver is therefore not only to give the best estimate of the transmitted bit $\hat{\alpha}^{(l)}(n)$ given the samples $r\left(mT_s + \hat{\pi}(m)T_s\right)$, but also to continuously estimate the value of $\hat{\pi}(m)$.

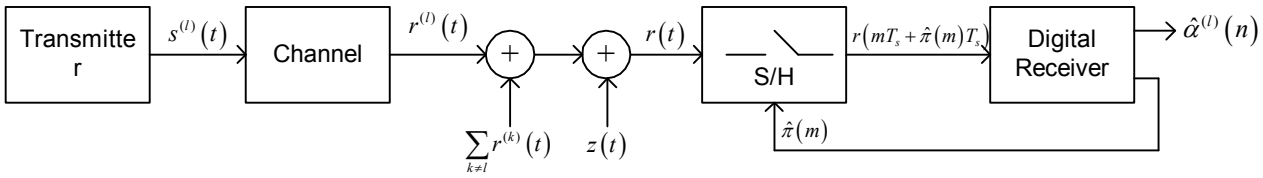

Figure 31: Transceiver model including the effect of timing offset.

Dependent on the chosen implementation, the correction $\hat{\pi}(m)$ can be allowed to operate at the sample-, chip- or bit-level. This basically means that the correction $\hat{\pi}(m)$ will correct errors up to a full sample, chip or bit length, as can be implemented by calculating $\hat{\pi}(m)$ modulo 1, N_{sp} or $N_{sp}SF$ respectively. This will eventually result in a wrap-around and when this happens a sample, chip or bit must be either inserted or removed.

Besides continuously estimating $\hat{\pi}(m)$, the digital receiver must also acquire bit and frame synchronization. The need for bit synchronization can be easily seen in the case of a matched filter receiver in an AWGN channel: If the transmitter and receiver are not synchronized on the bit level, the

signal output after the matched filter coming from the desired bit, will be proportional to the matched filter auto-correlation at the lag that equals the synchronization error between the transmitter and the receiver. This means that if the auto-correlation function is down by $\sqrt{2}$ compared to the synchronized position, a shift of the BER curve of 3dB is experienced as a result of the reduced SNR after the matched filter. It should however be noted, that if $\hat{\pi}(m)$ operates at the bit-level, no further bit synchronization is needed as this is included in the timing offset $\hat{\pi}(m)$.

To get an idea about the synchronization task in UWB systems, the normalized auto-correlation of a typical monocycle is plotted in Figure 32. The monocycle is the 7th derivative of a Gaussian pulse and it has a pulse width of 2/3ns making it rather typical for UWB usage. As can be seen the auto-correlation falls off rapidly around the synchronized position and is even negative in many places leading to bit inversions. In fact, if no more than a 3dB penalty can be accepted, the receiver and transmitter must be synchronized to within ±21ps in this example. This extremely high synchronization precision clearly shows why synchronization is vital to the operation of a UWB communication system.

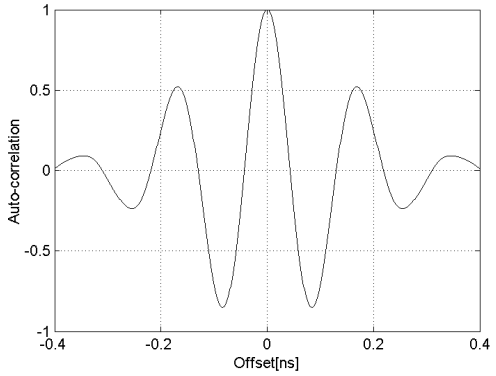

Figure 32: Normalized auto-correlation of $p_7(t)$ with $PW = \frac{2}{3} ns$.

In order to understand the need for frame synchronization, it is first needed to understand what frames are and what they are used for. When communicating from one place to another, not only the data that must be communicated are transmitted. The reason is that the communicating parties need to establish when the data starts and stops, how many bits are transferred, etc. This is usually achieved by transmitting a so-called header before the actual data, informing the receiver about the pending communication as illustrated by Figure 33. The header also holds a special sequence of bits that are known to the receiver, which can be used for synchronization and/or channel estimation purposes.

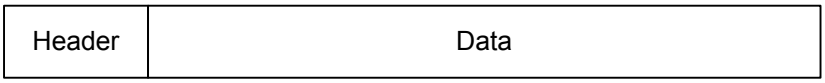

Figure 33: General frame format used for communications.

It is therefore crucial to the receiver to be able to not only be synchronized at the bit level, but also at the frame level in order to correctly decipher the transmission.

The synchronization is usually implemented using one of two different approaches. The first is the feed-forward approach, which finds the maximum-likelihood synchronization parameters by simply trying all possible values and selecting the parameters that give the best match based on some criterion. The other approach is to feed back a low-pass filtered error signal, which will eventually indicate the correct parameters given that the error signal is unbiased. The latter approach has the advantage of having a lower computational complexity, but this comes at the expense of increasing the duration before correct synchronization parameters are achieved. Combinations of the two approaches are of course also possible.

Ideally, all synchronization parameters should be estimated jointly, but this may not be achievable because of the large complexity that this approach may lead to.

In order to understand that synchronization is not a trivial task, it should be mentioned that other communication systems typically use a separate channel with a higher transmitted power for synchronization purposes to make the job easier. This trick is used both in the GSM and WCDMA cellular systems, but this is unfortunately not possible in UWB systems as they are power constrained. The gain coming from the increased power in traditional synchronization systems must therefore instead come from the synchronization sequence itself, as was the case with the processing gain of a TH or DS system. But achieving the gain requires that the systems are already synchronized, which was the whole point of using the synchronization sequence. This chicken and egg problem is the heart of the synchronization problem.

6.2 Synchronizing the Adaptive LMMSE UWB Receiver

As the adaptive LMMSE receiver is the best performing multi-user receiver considered in this thesis, it is natural to focus on the task of synchronization using this receiver. This section will consequently present a method to achieve this and present simulations of its performance.

As mentioned earlier the task of estimating the timing correction $\hat{\pi}(m)$ is important to maintain synchronism, but to simplify the receiver this estimation is not considered further in this chapter. The interested reader is referred to material such as [27] for extensive treatment of this subject. Thus only the task of bit and frame synchronization will be included in this chapter, as it is assumed that only the relative value of $\hat{\pi}(m)$ is known to the receiver. This means that the receiver will still have to perform a full bit and frame synchronization, but the drifting of the clock in the receiver is not considered. As a result of this, it is assumed that the transmitter and receiver uses the exact same clock with the only difference being an unknown offset.

First, bit synchronization must be established. This is possible by taking advantage of the inherent adaptive nature of the receiver. If at first the AWGN channel is observed, it can be noted that if the receiver is not synchronized to the transmitter, extending the filter length by one bit length can capture all energy from a desired bit. The adaptive algorithm will therefore automatically suppress the coefficients outside of the correct bit interval and bit synchronization is therefore automatically achieved, but this comes at the expense of increasing the filter length to twice its original size. Increasing the filter length by a bit length in the UWB multipath channel will, in a similar way as in the AWGN channel, ensure that at least the same energy is captured as if the systems were synchronous.

The increase in the filter length may be modeled by a larger value of ψ given by

$$\psi = \psi_{sync} + \psi_b$$

where ψ_{sync} determines the filter length of the fully synchronized system and ψ_b represent the increase needed to accommodate a full bit length. The value of ψ_b can be found directly from

$$\psi_b P = N_{sp} SF \Rightarrow \quad \psi_b = \frac{N_{sp} SF}{P} = \frac{SF}{SF + \frac{(N-1)\Delta\tau}{PW}} \qquad (6.1)$$

The AWGN channel therefore requires $\psi_b = 1$. Using the same parameters as earlier, i.e. $SF = 15$ and $PW = \frac{2}{3} ns$ and assuming the multipath channel from Chapter 4, the value is found to be $\psi_b = 0.048$ in the multipath channel. This low value of ψ_b directly shows that it is not necessary to lengthen the filter very much in order to extend it by a full bit length and the increase in complexity is therefore minimal.

This is a direct consequence of the fact that the energy spread in the channel is much larger than the bit period.

To see the effect of the increased bit length, the BER performance of the adaptive LMMSE receiver is plotted in Figure 34 showing the performance of both the original and the extended adaptive receiver. These filter lengths correspond to the worst- and best-case synchronization scenarios respectively when using the extended filter length receiver. This is a result of the fact that in the case where the offset is very close to a whole bit length, the performance will be similar to that of the original filter length receiver. The BER is therefore dependent on the actual synchronization offset, but is confined by these two bounds. The result of Figure 34 in the AWGN channel using RLS adaptation is a bit surprising; increasing the filter length improves the performance even though no extra bit energy is collected. The reason is that increasing the filter length will result in effectively lowering the value of δ. This can be seen directly from the expression in (5.42) as the denominator will increase.

The results show that performing bit synchronization in this manner is favorable with the drawback of increasing the complexity as a result of enlarging the filter length. Another drawback is the potential noise and interference introduced outside the correct bit interval, as the adaptive algorithm may be incapable of correctly suppressing these filter coefficients. The increase in complexity may be compensated for by estimating $\hat{\pi}(m)$ from the converged filter coefficients, making it possible to reduce the filter length back to the original length. This estimation is possible as filter coefficients outside the correct bit interval will tend to be smaller than the rest.

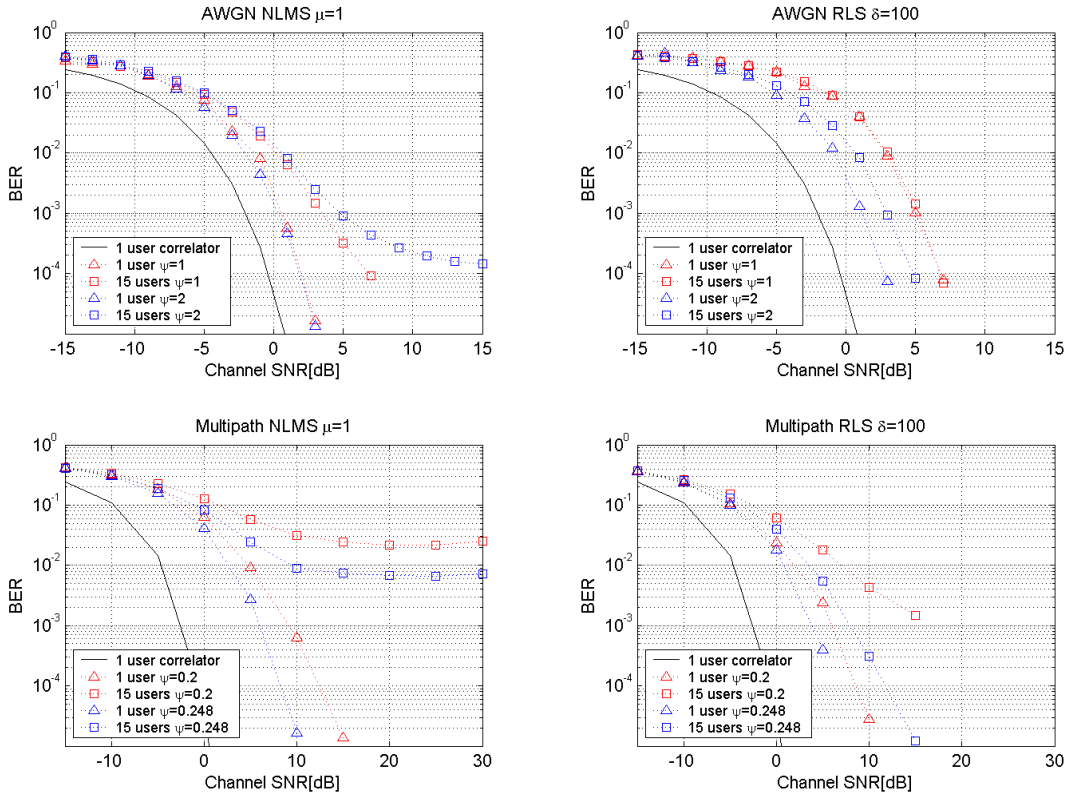

Figure 34: Performance comparison of best- and worst-case energy captures resulting from the synchronization offset. Blue results indicate best-case BERs and red indicate worst-case BERs.

Performing the bit synchronization as mentioned above is only possible if a training sequence is available and the receiver will therefore have to acquire frame synchronization to do this. This is done using the feed-forward method by searching over all possible positions in the frame and selecting the one yielding the smallest mean-square error after performing adaptation using the known training sequence. The complexity of this search is proportional to the length of the frame, and it is therefore often desirable to

perform this search in a serial manner, such that only one position is tested for every received frame. It is also possible to implement this search in a parallel fashion and thus trade acquisition speed for hardware complexity.

For simulation purposes, it is assumed that the frame consists of 500 bits and that the training sequence in the header is a length 31 maximal-length sequence. Furthermore, the search is required to produce the same result over two consecutive frames in order for the receiver to be locked onto the signal. This is done in order to prevent accepting false synchronization positions. A simulation of this system is performed with the results plotted in Figure 35, showing the average number of frames needed before the receiver has correctly acquired frame synchronization. The bit synchronization is performed as mentioned above by increasing the filter length by a full bit length.

It can be seen from the figure how this synchronization method performs well in the AWGN channel in both the single- and multi-user scenarios. Operating in the UWB multipath channel, the single-user system also works rather well loosing only some 5dB compared to the AWGN case. No graph is available for the case of 15 users, as in this case the receiver simply is not able to acquire the correct synchronization independent of the SNR. The problem is that the adaptive filter is not capable of converging within the 31 training bits and a longer synchronization sequence is therefore needed in order to synchronize under these conditions. This will however decrease the number of data bits that can be sent and a tradeoff is therefore needed. However, it is not possible to say exactly how many training bits are required without performing the simulations. The convergence plot in Figure 27 shows that the filter is not fully converged until approximately $0.2P$ iterations have been performed, which is close to 800 training bits, but it may not be necessary for the filter to have converged completely in order to acquire frame synchronization. The problem of synchronizing in the multipath channel with up to 15 active users is therefore a topic of further research.

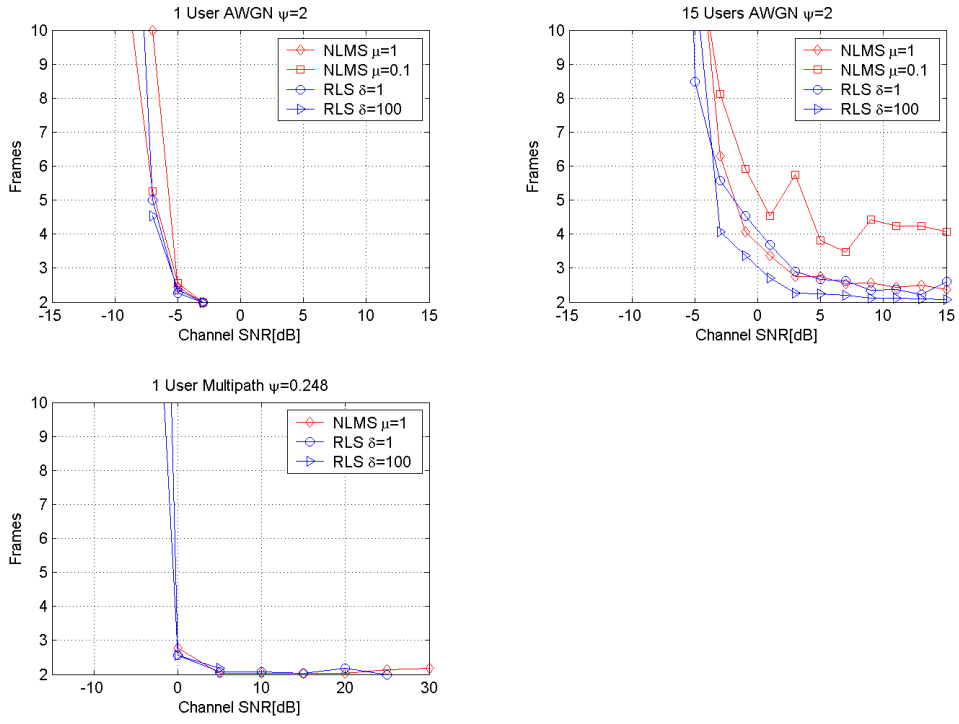

Figure 35: Plot of the average number of frames needed by the receiver to correctly lock onto the transmitted signal.

6.3 Summary

In this chapter an introduction to the topic of synchronization is given. A novel synchronization method taking advantage of the adaptive LMMSE receiver is presented and simulated. The simulations show good results in the AWGN channel and when only a single user is active in the UWB multipath channel. Increasing the number of simultaneous users in the UWB multipath channel to 15, it is no longer possible to achieve correct synchronization. The selected synchronization sequence is simply too short for sufficient convergence of the adaptive receiver and a longer synchronization sequence must be used in order to obtain synchronization in this situation.

7 Coexistence with other Narrowband Communication Systems

Wireless UWB systems are interesting for a number of reasons, but the perhaps most important one is the possibility of reusing already allocated spectrum. As discussed earlier this is feasible because of the low PSD of the radiated UWB signal, but this raises the question of coexistence. As UWB systems for communication purposes are mainly allowed in the 3.1-10.6GHz frequency band, the perhaps most challenging contender in this frequency band are IEEE802.11a/HIPERLAN2 WLAN systems as they operate in the same environment. The interference scenarios outlined in this chapter will therefore focus on this problem, but generally applies to any narrowband technology used in the same environment as UWB systems and using overlapping frequency bands.

The coexistence can be divided into two parts

- Interference from UWB systems to other narrowband systems
- Interference from other narrowband systems to UWB systems

In order to make estimates of the level of interference, it is important to know the propagation conditions under which the systems operate. As the interferers are uncorrelated with the desired signal, the statistical properties of the channel model of the interferers become less important. Instead the received interference power can be used to estimate the impact on the desired signal and it is therefore only necessary to known the Path Loss (PL) of the interferer. The path loss is calculated by [26 p. 143]

$$PL = \frac{c^2}{16\pi^2 r^n f_c^2} \quad (7.1)$$

with c being the speed of light, r being the range and f_c is the center frequency, all in SI units. The path loss exponent n is a function of the environment and is usually in the range 1.5-6. A special case is Free Space Propagation (FSP) where $n = 2$, which is a good approximation when operating in Line Of Sight (LOS) conditions up to approximately 10m [16].

7.1 Interference from UWB Systems

Even though the emitted PSD of a UWB system is low it can potentially interfere with other systems if the systems are placed close together. To see this the separation between interferer and receiver yielding interference power equal to the noise floor will now be calculated.
The interference is most severe when the separation is small and FSP is therefore considered. The thermal noise PSD at the receiver assuming a room temperature of $T = 290K$ is $NF \cdot kT$ with k being Boltzmann's constant and NF being the noise figure of the receiver. A typical value of the noise figure in a narrowband receiver is 2dB [26 p. 214], which is assumed here. The range r, where the UWB signal contributes as much power as the thermal noise does, is therefore

$$\frac{NF \cdot kT}{75nW/MHz \cdot PL} = \frac{NF \cdot kT \cdot 16\pi^2 r^2 f_c^2}{75nW/MHz \cdot c^2} = 1 \Rightarrow r = \frac{c}{4\pi f_c}\sqrt{\frac{75nW/MHz}{NF \cdot kT}} \simeq 18m \quad (7.2)$$

UWB systems can therefore potentially cause interference with other narrowband systems using the same frequency band at a range of up to approximately 18m, but whether it will cause a problem is very dependent on the receiver. If for example the receiver under normal operation is not limited by thermal noise, most likely no problem will occur.

7.2 Interference from Other Systems

In the frequency range used by UWB systems numerous systems operate in already allocated spectrum, but in order to pose a problem for UWB operation, they must be close to the UWB receiver. Secondly

these interfering signals are normally narrowband and therefore only cover a small part of the signal bandwidth.

The most likely interferers are IEEE802.11a/HIPERLAN2 WLAN systems, which operate in the 5GHz Industrial Scientific Medical (ISM) band with a channel bandwidth of 20MHz radiating 200mW. The system is OFDM-based with 52 sub-carriers each modulated using BPSK, QPSK, 16-QAM or 64-QAM delivering up to 54 Mbps pr. channel. The Signal-to-Interference Ratio (SIR) at the UWB receiver is calculated as

$$SIR_{WLAN} = \frac{75nW/MHz \cdot B \cdot PL_{UWB}}{200mW \cdot PL_{WLAN}} \quad (7.3)$$

with B being the UWB bandwidth in MHz and PL_{UWB} and PL_{WLAN} being the path loss of the UWB and WLAN systems respectively.
Assuming a UWB system operating in the frequency band from 3.1-6.1GHz, the resultant bandwidth is 3GHz and the center frequency of both systems will be nearly the same. It is therefore possible to reduce (7.3) to

$$SIR_{WLAN} \simeq 10^{-3} \frac{r_{WLAN}^{n_{WLAN}}}{r_{UWB}^{n_{UWB}}} \Leftrightarrow SIR_{WLAN}[dB] \simeq -30 + 10n_{WLAN}\log_{10}(r_{WLAN}) - 10n_{UWB}\log_{10}(r_{UWB}) \quad (7.4)$$

where n_{WLAN} is the path loss exponent and r_{WLAN} is the range of the WLAN system and likewise for the UWB system. The rightmost part of (7.4) clearly shows the problem: If the WLAN and UWB system both operate using the same distance and path loss exponent, the UWB receiver will experience $SIR = -30dB$. An actual scenario may then be either better or worse than this, but generally the UWB receiver must be able to deal with this large amount of interference. A realistic worst-case scenario is $SIR_{WLAN} = -50dB$, as this will correspond to the WLAN system being 10 times closer to the UWB receiver than the UWB transmitter, assuming $n_{WLAN} = n_{UWB} = 2$. If the UWB system has a range of up to 10m, this will mean that the WLAN transmitter is within 1m of the UWB receiver and it should therefore be possible to reposition the WLAN transmitter or the UWB receiver so that interference is reduced.

In order for the receiver to operate, the interference level must be reduced. This can be done using two different strategies

- Suppress interference by not using the frequency band in which the interferer operates
- Interference cancellation in the receiver

The use of interference suppression may be implemented in various ways. One way is to design the monocycle such that little or no energy is placed in the frequency band where the interference is present. Another way is to have an adaptive notch filter track the interference and then filter out the bands being influenced by interference.

Another option is to use various interference cancellation techniques in the receiver, but at the cost of increased complexity. The adaptive LMMSE receiver as outlined in Section 5.5 is an interesting receiver in this context, because of its ability to reject various forms of interference. It is therefore natural to evaluate this receiver under the influence of IEEE802.11a/HIPERLAN2 interference, as no change to the receiver is necessary and therefore no increase in complexity is produced.

7.2.1 Simulation

To see the influence of the WLAN interference on the adaptive LMMSE receiver, a number of simulations have been performed. The UWB and WLAN systems are assumed operating in an AWGN channel. The OFDM signal can be modeled as being Gaussian, as it consists of a large number of independent sub-carriers added together. The WLAN interference is therefore modeled as Gaussian noise being passed through a band pass filter with a bandwidth of 20MHz centered on 5.2GHz, as specified by

the standard. The performance of the true LMMSE receiver is found by (5.24), but the total noise variance must be increased by an amount equal to filtering the interference and squaring the result. This procedure is completely similar to the one done in (5.25) and (5.26).

The results shown in Figure 36 indicate that dealing with WLAN interference is not trivial; The NLMS adaptation is completely overcome by the interference and is unable to reject the interferer sufficiently. The RLS algorithm is much better at rejecting the interference, as it uses knowledge of the covariance matrix of the signal to eliminate the interferer and is therefore virtually unaffected even when $SIR_{WLAN} = -50dB$.

As a result it must be concluded that unless other measures are taken to reduce the interference to the UWB system, NLMS adaptation of the LMSSE receiver is not an option, as it is not capable of acceptable rejection of the interference.

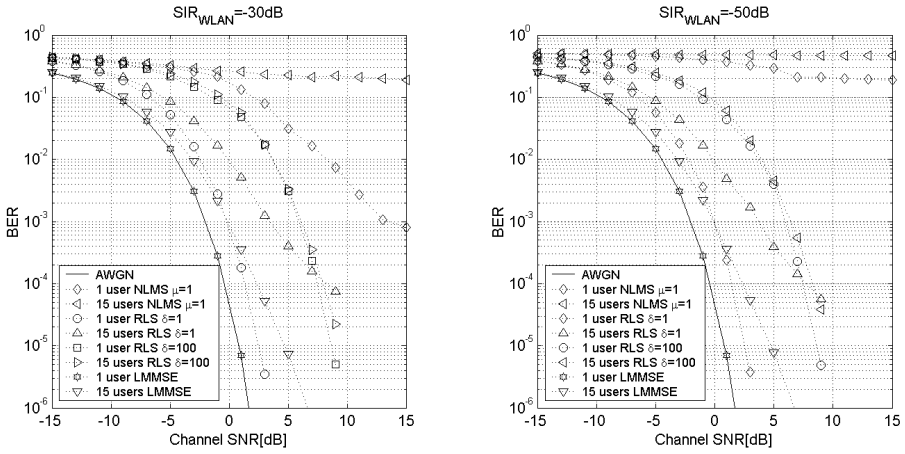

Figure 36: Performance of the adaptive LMMSE receiver in the AWGN channel under WLAN interference.

7.3 Summary

In this chapter it has been illustrated how coexistence with other narrowband systems is crucial to the successful use of UWB systems and possible ways of eliminating the interference from the UWB signal are briefly discussed. In particular the adaptive LMMSE receiver is shown to have good interference rejection, especially when using RLS adaptation.

8 UWB Usage Scenarios

In this chapter the different uses of UWB will be more closely examined and evaluated for a given theoretical scenario in order to quantify the feasibility and maximize performance. Solutions presented in this chapter are the results of the previous chapters, where basic configurations have been examined in order to select system parameters.

The different uses of UWB systems can be divided into roughly three scenarios dependent on their maximum operating range $r_{\max}$

1. ***Very-short distance operation with r_{max}<1 m.*** Operating over such short distances, a UWB system will be capable of delivering a very-high bitrate service such as a wireless USB2 or FireWire connection with a bitrate in excess of 500Mbps. Such a system might be used for wireless access to a portable storage media.
2. ***Short distance operation with r_{max}<10 m.*** This type of UWB operation is of principal interest when UWB is used as a WPAN/WLAN system with a bitrate in the area of 100Mbps.
3. ***Medium to long distance operation with r_{max}<10-1000 m.*** In this case a UWB system is used that gives a low bitrate in the 10-100kbps range coupled with the possibility to determine the position of the UWB terminals making multi-hop routing possible. Such a system could be used for wireless sensors in connection with example fire detectors or factory automation, where a lot of cable installations can be avoided.

UWB systems from all scenarios are assumed to comply with current emission limitations set out by the FCC [1] limiting the radiated power to –41.3 dBm/MHz or 75nW/MHz in the 3.1-10.6 GHz band. All systems considered here will use maximum transmission power in the 3.1-6.1 GHz band with center frequency $f_c = 4.6GHz$ and a 3dB bandwidth of $B = 3GHz$. This has been chosen as it is a reasonable choice for first generation UWB systems and because the attenuation is lowest in this part of the band. A monocycle that fits these specifications well is the 7th derivative of a Gaussian pulse having a pulse width of 2/3 ns making the maximum aggregate throughput 1.5Gbps using binary modulation schemes.
Another thing assumed common to all categories is the antenna and RF front-end performance. The antenna is assumed isotropic having an efficiency of 80% [36] and the RF front-end has a noise figure of approximately 6 dB [28]. The temperature used for thermal noise floor calculations is 290 K. The path loss is calculated by

$$PL = \frac{c^2}{16\pi^2 r^n f_c^2} \quad (8.1)$$

with c being the speed of light and r being the range. The path loss exponent n is a function of the environment and is usually in the range 1.5-6. A special case is Free Space Propagation (FSP) where $n = 2$.
The three different types of UWB system will now be investigated in more detail.

8.1 Scenario 1: Very-short Distance Operation with r_{max}<1 m

As this system only covers a range of up to 1m, the number of simultaneous users that must be supported by the system is very low and in the order of 1-4 dependent on specific use. For this evaluation it has been chosen to only support one simultaneous user. This is valid in the case of wireless USB2/FireWire, as each of these systems has their own multiple-access scheme.
Having the wireless USB2/FireWire application in mind, the spreading factor is selected as $SF = 2$ such that the bitrate will be 750Mbps, which should be sufficient to support these services.
A very important issue in the performance evaluation of any wireless system is the communication channel model as this is often the most limiting factor. It is therefore critical to obtain a reasonable channel model, if a realistic performance measure of the system is to be determined by simulations. Such

a channel model for distances under 1 m is currently not available and a qualified guess must therefore be used.
It is expected that the channel model will converge to the AWGN channel as the distance goes to zero and this is therefore a natural choice in order to keep the model simple. It has therefore been decided that the model used for this scenario is the AWGN channel with Line Of Sight (LOS) FSP path loss.

A link budget for the system can now be constructed and it is shown in Table 1. It can be seen from the table that the SNR on the channel is very good and no further processing gain is therefore needed. It may even be possible to use a higher-order modulation scheme to increase the bitrate further.
Even though it is assumed that the channel is AWGN, it is reasonable to believe that some amount of multipath will be present and it is therefore preferable to use a Time-Hopping scheme to help minimize ISI and keep the receiver simple. A binary modulation scheme, preferably BPSK, should most likely also be preferred because of the expected energy spread in the channel.
As the system is assumed to be operating over an AWGN channel, the receiver used is simply the correlator receiver.

r_{max}<1m, f=3.1-6.1 GHz, f_c=4.6 GHz, K=1, AWGN channel with LOS FSP path loss	
Tx power $B \cdot 75nW / MHz$	-6.5 dBm
Thermal noise floor kTB	-79.2 dBm
FSP path loss $\frac{c^2}{16\pi^2 r^2 f_c^2}$, $r = 1m$	-45.7 dB
Rx antenna efficiency $\eta = 80\%$	-1 dB
RF front-end (Noise Figure 6 dB)	-6 dB
Channel SNR	20 dB
Processing Gain $P_G = 2$	3 dB
Detection SNR	23 dB

Table 1: Link budget for very-short range UWB communications.

A simulation of a system with parameters described as above has been simulated and the result is shown in Figure 37 as a function of both the channel SNR and the corresponding distance when FSP is considered. It can therefore be concluded that given the assumptions used for this simulation, this type of UWB system is definitely feasible.

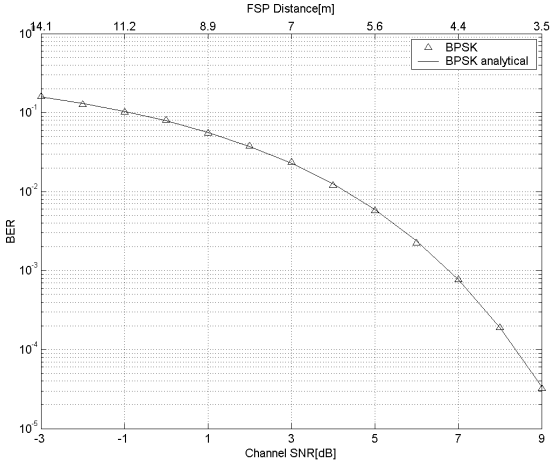

Figure 37: Simulation of very-short range UWB system over AWGN channel.

It may be argued that the assumptions used are overly optimistic, as the ideal FSP impulse response in the UWB case is not the delta pulse assumed when using an AWGN channel. The first reason is that that the FSP path loss is frequency dependent and because of the large bandwidth of UWB systems this should be included. The result will be that the pulse will be smeared out over a larger time interval meaning that the received power will be decreased. Secondly, the assumption of no multipath in the channel may be too simple and the received energy with a correlator receiver will diminish further. If this is the case, a RAKE receiver with a few arms should be used instead, but overall the very-short distance UWB scenario is promising because of the high channel SNR available and the relatively simple transceiver design.

8.2 Scenario 2: Short Distance Operation with r_{max}<10 m

Serving the needs typically set out by a WLAN/WPAN scenario, the maximum number of users that should be able to simultaneously use the system within a 10m radius will typically be in the range of 8-32 users. For this evaluation it has been decided to fix it at a maximum of approximately 16 users by using a spreading factor of $SF \approx 16$ giving a bitrate of close to 100Mbps.
The environment that such a system will typically operate in will be a home/office environment with both LOS and NLOS conditions as the signal path may be obstructed and the used channel model must reflect this. The channel model described in Section 4.1 will be used as basis for this evaluation as it deals with exactly this scenario.
In order to get an idea of the potential of the system, a link budget is set up as shown in Table 2. The path loss used is based on the channel model stating that the path loss is close to the FSP path loss up to approximately 10m. One thing that must be remembered when observing the link budget is that the SNR depicted is the total power when all multipath components are collected. If this is not the case, the collected signal power will be smaller. The link budget also does not take into account the performance penalty incurred when multiple users are active in the system and the link budget should therefore only be seen as an indication that such a system is feasible if sufficient multipath energy can be captured and if the multiple-access interference can be sufficiently rejected. As can be seen, the SNR is relatively poor and this is the reason why only BPSK should be used as will be assumed here.

r_{max}<10m, f=3.1-6.1 GHz, f_c=4.6 GHz, K≤16, NLOS multipath channel with FSP path loss	
Tx power $B \cdot 75nW / MHz$	-6.5 dBm
Thermal noise floor kTB	-79.2 dBm
FSP path loss $\frac{c^2}{16\pi^2 r^2 f_c^2}$, $r = 10m$	-65.7 dB
Rx antenna efficiency $\eta = 80\%$	-1 dB
RF front-end (Noise Figure 6 dB)	-6 dB
Channel SNR	0 dB
Processing Gain $P_G = 15$	12 dB
Detection SNR	12 dB

Table 2: Link budget for short distance UWB communications.

As mentioned in Chapter 5 no good multi user detector is available for TH systems and in order to deal with up to 16 simultaneous users, a DS system using an adaptive LMMSE receiver will be the best choice. It is then possible to use either NLMS or RLS adaptation for the receiver, but as shown in Chapter 7 the NLMS adaptation is not well suited to the low Signal-to-Interference Ratio (SIR) that may arise from a IEEE802.11a/HIPERLAN2 WLAN interferer being close to the receiver. This leaves RLS adaptation as the only truly viable choice. From Figure 29 it is seen that using a large value of δ generally gives better performance in the multipath channel and δ is therefore selected as $\delta = 100$. The spreading factor is selected to be $SF = 15$ in order to use large-set Kasami codes to separate the users, making it possible for up to approximately 15 users to use the radio channel simultaneously.

The selected setup makes it possible to directly translate the results of Section 5.5 into an estimate of the BER performance of the actual system at a given distance between transmitter and receiver as specified by the FSP path loss. This is what is done in Figure 38, where the BER performance of the receiver is depicted as a function of the SNR on the lower horizontal axis and as a function of the FSP distance in meters on the upper horizontal axis. The translation to distance is done according to the link budget given in Table 2.
From this result it can be seen, that selecting $\psi = 0.2$ is the perhaps best trade-off between complexity and performance, as a raw BER of 10^{-1}-10^{-2} will usually be sufficient to correct most errors if an error correcting code is used, which is almost always the case. This coding will however lower the effective bitrate by a factor equal to the used code rate.
As a consequence it may not be possible to achieve the goal of 100 Mbps for all 15 simultaneous users, if the distance is 10m. A solution to this problem is to use a variable spreading factor dependent of the signal condition leading to different bitrates in different situations.

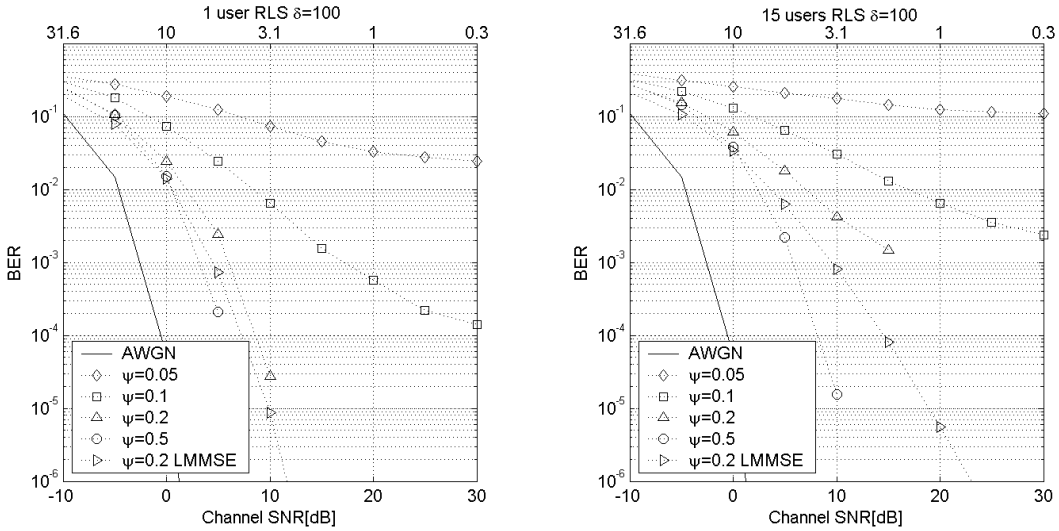

Figure 38: Performance of the adaptive LMMSE receiver in realistic multipath channel.

Generally this communication scenario is also very promising, as it directly shows that high bitrate links are possible up to approximately 10m. The major problem of this scenario is the large complexity of the receiver.

8.3 Scenario 3: Medium to Long Distance Operation with r_{max}<10-1000 m

A system like this having a bitrate in the area of 10-100kbps will typically be used for voice communication and other types of low bitrate communication. Its main strength is the multi-hop capability made possible by determining the approximate position by observing pulse delays and the inherent low power usage.

The number of simultaneous users in the system can be large because of the longer range and it has therefore be selected that the system must support up to 1000 users. This amount may be bit overly pessimistic, but as can be seen from the link budget in Table 3, the number of users is not the limiting factor as much the thermal noise is. The table depicts two scenarios: One with a path loss exponent of 2 and one of 4. The reason for this is that no recognized channel model has been found for this medium to long distance system and one can consequently only try and give a qualified guess as to the value of the path loss exponent. But as the nature of this channel will be pure NLOS and the distances are relatively large, it is reasonable to assume that the path loss exponent will be closer to 4 than to 2. The table

includes a processing gain of 10^5, which will result in a raw bitrate of 10kbps if transmissions occur using 1Gpulses/s and therefore represents the highest possible processing gain.

Distances over 100m are therefore clearly not possible and even at 100m the detection SNR will be close to 0dB. Further it is expected that the energy spread in the channel will be even worse than specified by the model in Section 4.1 because of the longer distance and to reach the SNR depicted in Table 3 all of this energy must be collected.

It must therefore be concluded that this type of UWB system is not feasible given the current power constraints, as the receiver will have to be very complex to capture enough energy. This type of system is much better constructed using ordinary narrowband systems working at a much higher Power Spectral Density (PSD) to give a good SNR and possibly combined with a positioning device.

r_{max}<10-1000m, f=3.1-6.1 GHz, f_c=4.6 GHz, K≤1000, NLOS multipath channel		
	n=2	***n=4***
Tx power $B \cdot 75nW / MHz$	-6.5 dBm	
Thermal noise floor kTB	-79.2 dBm	
Path loss $\frac{c^2}{16\pi^2 r^n f_c^2}$, $r = 100m$	-85.7 dB	-125.7 dB
Rx antenna efficiency $\eta = 80\%$	-1 dB	
RF front-end (Noise Figure 6 dB)	-6 dB	
Channel SNR	-20 dB	-60 dB
Processing Gain $P_G = 10^5$	50 dB	
Detection SNR	30 dB	-10 dB

Table 3: Link budget for medium to long distance UWB communications.

8.4 Summary

In this chapter different usage scenarios are introduced along with their possible UWB implementation and simulations are presented to quantify their performance. The conclusion is that UWB systems are a very strong candidate for the very-short and short distance scenarios. Operation up to a distance of 1m is seen to be possible using simple transceivers at the same time as bitrates in excess of 500Mbps are possible for wireless USB2/FireWire applications.
Increasing the distance up to 10m, operation is possible by means of an adaptive LMMSE receiver using RLS adaptation with a filter length found by $\psi = 0.2$. This allows for a bitrate of up to 100Mbps pr. user allowing up to 15 simultaneous users. The price paid for increasing the operational distance to 10m, is increased complexity needed to capture the multipath energy and reject multiple-access interference.
The last scenario considered involves a maximum range of up 1000m with a low bitrate, but allowing for many simultaneous users. From the link budget of this scenario, it is concluded that this scenario is unrealistic, as the receiver will have to be extremely complex in order to capture sufficient power. A narrowband system will most likely be a better candidate for this type of system.

9 Conclusion and Topics for Further Research

In this thesis UWB communication systems have been introduced and analyzed from a signal processing point of view, with the purpose of using the technology for wireless short-range communication involving high bitrate links.

9.1 Project Summary

Chapter 1 gives an introduction to the physical nature of UWB signals and what such signals can be used for. Furthermore, a comparison with other competing narrowband communication systems is performed.
A mathematical model of the UWB communication system is presented in Chapter 2 along with the optimal single-user receiver in the AWGN channel. It is shown how the performance of the optimal receiver using QPPAM modulation can be approximated and verified and it is shown that to maximize performance, BPSK modulation should always be used.
Adding more users to the system is investigated in Chapter 3 showing how it is possible to separate users by the use of spreading codes in a CDMA system. In addition, it is shown how the simple correlator receiver is not well suited for multi-user detection, as it is overcome by the strong interference. The PSD of the transmitted signal is derived and verified, showing that the spreading code of a DS-CDMA system impacts the PSD of the signal and that it is therefore potentially possible to control the used spectrum.
The problem of multipath signal propagation is studied in Chapter 4 and a radio channel model suitable for multipath simulation purposes is described. A novel proof of the ISI rejection of both TH- and DS-CDMA systems is given, showing that both systems achieve nearly the same ISI rejection. It is consequently more important that the receiver is capable of making good use of the ISI energy, than what spreading scheme is used.
Advanced receivers capable of capturing the ISI energy is examined in Chapter 5. The RAKE receiver with the Maximal Ratio Combiner (MRC) is derived and simulated and results clearly show its great ability to capture multipath energy. Loading the system with more users performance drops dramatically, as the RAKE receiver is a single-user receiver. The decorrelator and LMMSE receivers are therefore introduced along with the adaptive LMMSE receiver. Simulations show that the adaptive LMMSE receiver performs very well in both AWGN and multipath scenarios and this receiver is therefore selected for the remaining studies. Especially the adaptive LMMSE receiver using RLS adaptation is shown to exhibit good performance.
The problem of synchronizing the receiver to the transmitter is briefly discussed in Chapter 6 and a new way of obtaining synchronism using the adaptive LMMSE receiver is introduced and simulated. Results show that this type of synchronization performs well, but that long training sequences may be needed for the adaptive receiver to properly converge if many users are active in the UWB multipath radio channel.
Next, coexistence with overlapping narrowband systems is discussed, as the UWB and narrowband systems may pose an interference threat to each other. Calculations show that UWB systems can potentially interfere with overlapping narrowband systems, but the overlapping narrowband systems will introduce massive interference into the UWB receiver. It is further shown how the adaptive LMMSE receiver can reject this strong interference, given that RLS adaptation is used.
Finally, in order to analyze the suitability of UWB communication systems for different scenarios, the possible usage scenarios are split into three and analyzed separately. From these analyses it is concluded that the very-short and short-range scenarios are viable, making it possible to deliver high bitrate links over short distances. Using UWB signals for long distance links, is however not practical, as the energy spread introduced by the radio channel becomes too great.

9.2 The future of UWB

It has been shown in this thesis that wireless UWB communication systems achieving high bitrate connections are possible and configurations suitable for short-range communication are outlined in Chapter 8. However, to make UWB systems commercially feasible some signal processing hurdles still have to be overcome. The large computational complexity of the adaptive LMMSE receiver must be reduced; from a practical point of view, it is unrealistic to use RLS adaptation with today's technology, as

its complexity grows with the squared filter length. Adaptive algorithms yielding similar performance, but with complexity scaling proportional to the filter length, is a topic for further studies. A candidate for such further research is the conjugate gradient algorithm [32]. Another area that requires more research is that of sub-Nyquist sampling [33], as this can potentially reduce the necessary samples pr. monocycle and thus the filter length.
The chapter on synchronization gives an introduction to the synchronization topic, but it does not include all realistic synchronization problems. It is shown that the adaptive LMMSE receiver is capable of performing synchronization, but in worst-case scenarios synchronization is not achieved. A more comprehensive research must be performed to solve this problem and should include all realistic synchronization effects, including the clock drift between transmitter and receiver.

In an attempt to solve the before mentioned problems, the standardization body IEEE802.15.3a responsible for high bitrate WPAN standards, may move towards a multiband UWB solution in favor of the singleband solution presented in this thesis. The reason is that the multiband solution is scalable, as the number of bands used can be adjusted dependent on market demand and cost. Furthermore it is possible to dynamically disable bands where strong interferers are present. The drawback of using the multiband solution is that the bandwidth of each band will consequently be much smaller and the advantages relating to the ultra-wideband nature of the singleband UWB signal will therefore be reduced. It is therefore possible that the first commercial UWB devices will be of the multiband type, but performance will not equal that of the singleband system. Performance may then be scaled as production costs drop over time by increasing the number of used frequency bands.
The topics discussed in this thesis are still being actively researched and can be expected to evolve significantly in the future, making the approach taken in this thesis ever more competitive with existing technologies. Another fact supporting this view is the continuous scaling of the CMOS production process according to Moore's Law [34], stating that the number of transistors double every 18 months. This scaling will eventually make it feasible to implement radio architectures based solely on software, as described in this thesis for the special case of UWB communications. This concept of Software Defined Radios (SDR) is believed to become commercially available within the next decade [35], enabling the removal of classical RF building blocks such as mixers and Phase Locked Loops (PLL). Another benefit of this approach is the possible support of multiple standards in a single physical layer design as only the software changes.

10 References

The references used in this thesis are outlined here. If possible they have been included on the CD accompanying this thesis named by their respective reference number.

[1] Federal Communications Commission (FCC), "Revision of Part 15 of the Commission's Rules Regarding Ultra-Wideband Transmission Systems", First Report and Order, ET Docket 98-153, FCC 02-48; Adopted: February 14, 2002; Released: April 22, 2002.

[2] C. E. Shannon, "A Mathematical Theory of Communications", The Bell System Technical Journal, Vol. 27 July-October 1948, p.379-423, p. 623-656.

[3] S. Zeisberg, C. Müller, J. Siemes, A. Finger, "PPM based UWB System Throughput Optimisation"

[4] M. Z. Win, R. A. Scholtz. "Impulse Radio: How it works" IEEE Communications letters, Vol. 2, No. 1, January 1998.

[5] Q. Li & L. A. Rusch. "Multiuser receivers for DS-CDMA UWB" 2002 IEEE Conference on UWB Systems and Technologies.

[6] J. Foerster, "The Effects of Multipath Interference on the Performance of UWB Systems in an Indoor Wireless Channel", VTC2001, presented May 2001.

[7] H. F. Engler, Jr. "Technical Issues in Ultra-Wideband Radar Systems", Introduction to Ultra-Wideband Radar Systems, ch. 2, Edited by J. D. Taylor, CRC Press 1994, ISBN:0849344409

[8] S. V. Maric & E. L. Titlebaum, "A Class of Frequency Hop Codes with Nearly Ideal Characteristics for Use in Multiple-Access Spread-Spectrum Communications and Radar and Sonar Systems", IEEE Transactions on Communications, Vol. 40, No. 9, September 1992, p. 1442-1447.

[9] E. L. Titlebaum, "Frequency- and Time-Hop Coded Signals for use in Radar and Sonar Systems and Multiple Access Communications Systems" IEEE 1993, p. 1096-1100.

[10] S. W. Golomb & H. Taylor, "Construction and properties of Costas arrays", Proc. IEEE, Vol. 72, p. 996-1009, August 1984.

[11] J. G. Proakis, "Digital Communications", 3rd edition, ISBN 0-07-051726-6.

[12] P. Parkway, "Ultra-Wideband Data Transmission System", World Intellectual Property Organization WO 01/93520 A2, 2001.

[13] J. W. McCorkle, "Ultra Wideband Communication System: Method and Device with Low Noise Pulse Formation", World Intellectual Property Organization Wo 01/93520 A2, 2001.

[14] I. Vajda, "Code Sequences for Frequency-Hopping Multiple-Access Systems", IEEE Transactions on Communications, Vol. 43, No. 10, p. 2553-2554, October 1995.

[15] S. Verdú, "Multiuser Detection", 1st edition, Cambridge University Press 1998, ISBN 0-521-59373-5.

[16] D. Cassioli, M. Z. Win & A. F. Molisch, "The Ultra-Wide Bandwidth Indoor Channel: From statistical Model to Simulations" IEEE Journal on Selected Areas in Communications, Vol. 20, No. 6, p. 1247-1257, August 2002.

[17] M. Z. Win & R. A. Scholtz, "On the Robustness of Ultra-Wide Bandwidth Signals in Dense Multipath Environments", IEEE Communications Letters, Vol. 2, No. 2, February 1998, p. 10-11.

[18] S. Haykin, "Adaptive Filter Theory", 3rd edition, Prentice Hall 1996, ISBN 0-13-322760-X.

[19] E. H. Dinan & B. Jabbari, "Spreading Codes for Direct Sequence CDMA and Wideband CDMA Cellular Networks", IEEE Communications Magazine, September 1998.

[20] A. A. M. Saleh & R. A. Valenzuela, “A Statistical Model for Indoor Multipath Propagation”, IEEE Journal on Selected Areas in Communications, Vol. SAC-5, No. 2, p. 128-137, February 1987.

[21] F. Zhu, Z. Wu & C. R. Nassar, “Generalized Fading Channel Model with Application to UWB”, IEEE Conference on Ultra Wideband Systems and Technologies, p. 13-17, 2002.

[22] G. Durisi & G. Romano, “Simulation Analysis and Performance Evaluation of an UWB system in indoor multipath channel”, IEEE Conference on Ultra Wideband Systems and Technologies, p. 255-258, 2002.

[23] D. T. M. Slock & T. Kailath, “Numerically Stable Fast Transversal Filters for Recursive Least Squares Adaptive Filtering”, IEEE Transactions on Signal Processing, Vol. 39, No. 1, January 1991, p. 92-114.

[24] A. H. Sayed & T. Kailath, “A State-Space Approach to Adaptive RLS Filtering”, IEEE Signal Processing Magazine, July 1994, p. 18-60.

[25] H. V. Poor & S. Verdú, “Probability of Error in MMSE Multiuser Detection”, IEEE Transactions on Information Theory, Vol. 43, No. 3, May 1997, p. 858-871.

[26] D. M. Pozar, “Microwave and RF Design of Wireless Systems”, 1st edition, John Wiley & Sons 2001, ISBN 0-471-32282-2.

[27] H. Meyr, M. Moeneclaey & S. A. Fechtel, “Digital Communication Receivers”, John Wiley & Sons 1998, ISBN 0-471-50275-8.

[28] General Atomics, “Overview of General Atomics PHY Proposal to IEEE 802.15.3a”.

[29] J. Romme & L. Piazzo, ”On the Power Spectral Density of Time-Hopping Impulse Radio”, UWBST 2002, Baltimore, May 2002.

[30] IEEE 802.15 WPAN Working Group homepage, http://ieee802.org/15/index.html.

[31] XtremeSpectrum, “Mono-Phase and Bi-Phase Ultra-Wideband”, White Paper.

[32] C. M. Bishop, “Neural Networks for Pattern Recognition”, Oxford University Press, ISBN 0-19-853864-2.

[33] J. G. Proakis & D. G. Manolakis, “Digital Signal Processing, Principles, Algorithms, and Applications”, 3rd edition, Prentice Hall, ISBN 0-13-373762-4.

[34] Information from Intel about Moore’s Law, http://www.intel.com/research/silicon/mooreslaw.htm

[35] General Information about Software Defined Radios (SDRs), http://www.sdrforum.org.

[36] H. G. Schantz, “Ultra Wideband Technology Gains A Boost from New Antennas”, Time Domain Corp.

[37] Xtremespectrum FAQ, http://www.xtremespectrum.com/products/faq.html.

[38] Information concerning future Bluetooth development, http://bluetooth.weblogs.com/discuss/msgReader$523?mode=day

Appendix A Selection of the Monocycle Pulse Width

In order to use the approximation that monocycles can be modeled as derivatives of Gaussian pulses, a finite pulse width must be assigned to the pulse. A common way of truncating pulses is to alter the duration of the pulse so that a specified part k_t of the pulse power is included in the truncated version and this is also the definition used here. Assuming that the monocycle $p_x(t)$ can be modeled as the x^{th} derivative of a Gaussian pulse $p(t)$ the definitions are

$$p(t) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(\frac{-t^2}{2\sigma^2}\right) \quad \text{(A.1)}$$

$$p_x(t) = \frac{d^x p(t)}{dt^x} \quad \text{(A.2)}$$

where σ^2 is the variance of $p(t)$. The definition of k_t can now be written as

$$k_t = \frac{\int_{-\frac{k_w\sigma}{2}}^{\frac{k_w\sigma}{2}} p_x^{\,2}(t)\,dt}{\int_{-\infty}^{\infty} p_x^{\,2}(t)\,dt} \quad \text{(A.3)}$$

with the pulse width $PW = k_w\sigma$ being the unknown. As the top integral of (A.3) is analytically unsolvable another more relaxed definition can be made

$$\frac{\int_{-\frac{k'_w\sigma}{2}}^{\frac{k'_w\sigma}{2}} p_x^{\,2}(t)\,dt}{\int_{-\infty}^{\infty} p_x^{\,2}(t)\,dt} \geq k_t \quad \text{(A.4)}$$

with k'_w being the smallest integer that satisfies (A.4) and thus $k'_w = \lceil k_w \rceil$. The pulse width will therefore be defined as $PW = k'_w\sigma$.

This makes it possible to perform simple iterations in which k'_w can be found given that the bottom integral of (A.4) can be calculated. In Table 4 shown below, the values of the bottom integral along with values of k'_w are shown for $x = 0..7$ when $k_t = 99.9\%$ with $\Gamma(n) = \int_0^{\infty} t^{n-1}e^{-t}dt$ being the Gamma-function. The proof of the relationship between the integral and the Gamma-function shown at the top of the table is left out because of space considerations.

x	$\int_{-\infty}^{\infty} p_x{}^2(t)\,dt = \frac{1}{2\pi\sigma^{(2x+1)}}\Gamma\left(x+\frac{1}{2}\right)$	k'_w
0	$\frac{1}{2\pi\sigma}\sqrt{\pi}$	5
1	$\frac{1}{2\pi\sigma^3}\frac{\sqrt{\pi}}{2}$	6
2	$\frac{1}{2\pi\sigma^5}\frac{3}{4}\sqrt{\pi}$	7
3	$\frac{1}{2\pi\sigma^7}\frac{15}{8}\sqrt{\pi}$	7
4	$\frac{1}{2\pi\sigma^9}\Gamma\left(\frac{9}{2}\right)$	7
5	$\frac{1}{2\pi\sigma^{11}}\Gamma\left(\frac{11}{2}\right)$	8
6	$\frac{1}{2\pi\sigma^{13}}\Gamma\left(\frac{13}{2}\right)$	6
7	$\frac{1}{2\pi\sigma^{15}}\Gamma\left(\frac{15}{2}\right)$	7

Table 4: Values of k'_w for x=0..7 with $k_t = 99.9\%$.

Appendix B Sampling of Monocycles

In this appendix the number of samples necessary to represent a given monocycle in an acceptable way will be investigated. This is an interesting question as the needed signal processing capability increases with the number of samples necessary.

Given that a monocycle $p_x(t)$ can be modeled as the x^{th} derivative of a Gaussian pulse $p(t)$ the definitions are given in (2.1) and (2.2). The PSD is then given by

$$\Sigma_p(f) = \left| \int_{-\infty}^{\infty} p(t) \exp(-j2\pi ft)\, dt \right|^2 = \exp\left(-(2\pi\sigma f)^2\right) \quad \text{(B.1)}$$

$$\Sigma_{p_x}(f) = (2\pi f)^{2x} \Sigma_p(f) \quad \text{(B.2)}$$

In order to determine the number of samples pr. pulse N_{sp} needed to reduce aliasing to at least k_{alias} dB under the maximum of the PSD, the following must be satisfied

$$k_{alias} \le -10\log_{10}\left(\frac{\Sigma_{p_x}\left(\frac{f_s}{2}\right)}{\Sigma_{p_x}(f_{\max})}\right) = -10\log_{10}\left(\left(\frac{f_s^2}{4f_{\max}^2}\right)^x \exp\left(-4\pi^2\sigma^2\left(\frac{f_s^2}{4} - f_{\max}^2\right)\right)\right) \quad \text{(B.3)}$$

where $f_{\max}$ is the frequency where the PSD is maximum and f_s is the sample frequency. In order to solve (B.3) the maximum of $\Sigma_{p_x}(f)$ with respect to f must be found

$$\frac{d\Sigma_{p_x}(f)}{df} = (2\pi)^{2x} \frac{d}{df}\left(f^{2x} \exp\left(-4\pi^2\sigma^2 f^2\right)\right) = 0 \Leftrightarrow$$

$$(2\pi)^{2x}\left(2xf^{2x-1}\exp\left(-4\pi^2\sigma^2 f^2\right) - 4\pi^2\sigma^2 2f \cdot f^{2x}\exp\left(-4\pi^2\sigma^2 f^2\right)\right) = 0 \Rightarrow \quad \text{(B.4)}$$

$$2xf^{2x-1} - 8\pi^2\sigma^2 f^{2x+1} = 2x - 8\pi^2\sigma^2 f^2 = 0 \Rightarrow f_{\max} = \sqrt{\frac{x}{4\pi^2\sigma^2}}$$

Using the fact that the pulse width $PW = k_w'\sigma = \frac{N_{sp}}{f_s}$ in (B.4) yields

$$f_{\max} = \sqrt{\frac{xk_w'^2 f_s^2}{4\pi^2 N_{sp}^2}} \quad \text{(B.5)}$$

Substituting (B.5) in (B.3) gives

$$k_{alias} \le -10\log_{10}\left(\left(\frac{\pi^2 N_{sp}^2}{xk_w'^2}\right)^x \exp\left(x - \frac{\pi^2 N_{sp}^2}{k_w'^2}\right)\right) \quad \text{(B.6)}$$

As isolating N_{sp} in the above equation is not possible the solution must instead be found by iteration.

An example of this procedure is presented next. Using the monocycle $p_2(t)$ as an example and requiring $k_{alias} = 50dB$ and setting $k_w' = 7$ as found in Table 4 yields

$$50dB \le -10\log_{10}\left(\left(\frac{\pi^2 N_{sp}^2}{2\cdot 7^2}\right)^2 \exp\left(2 - \frac{\pi^2 N_{sp}^2}{7^2}\right)\right) \overset{iteration}{\Rightarrow} N_{sp} = 10 \quad \text{(B.7)}$$

Appendix C Optimal UWB Single-user Receivers in AWGN

In this appendix the optimal receiver in an AWGN channel is given and the BER of selected UWB modulation schemes will be presented. It is also shown that an analytical expression for the BER of QPPAM can not be given.

C.1 General AWGN Detector Design

When the transmitted signal $s_m(t)$ having $m = 1...M$ different waveforms and $0 \le t \le T$ is subjected to a memory-less AWGN channel, the received signal $r(t)$ can be modeled as

$$r(t) = s_m(t) + n(t) \quad \text{(C.1)}$$

where the noise term $n(t)$ is Gaussian with zero mean and variance σ^2. It is shown in [11 p.233-249] that a bank of N correlators yields sufficient statistics to perform optimal detection with N being the dimension of the signal. The output of these N correlators are then given by

$$r_k = \int_0^T r(t) f_k(t)\, dt = \int_0^T \left(s_m(t) + n(t)\right) f_k(t)\, dt \Leftrightarrow$$

$$r_k = s_{mk} + n_k = \int_0^T s_m(t) f_k(t)\, dt + \int_0^T n(t) f_k(t)\, dt \quad , k = 1..N \quad \text{(C.2)}$$

$$\mathbf{r} = [r_1, ..., r_N] \qquad \mathbf{s}_m = [s_1, ..., s_N] \qquad \mathbf{n} = [n_1, ..., n_N]$$

with $f_k(t)$ being the N orthonormal basis functions that makes it possible to represent $s_m(t)$ as a linear combination of $f_k(t)$, that is $s_m(t)$ can be represented as

$$s_m(t) = \sum_{k=1}^{N} a_{mk} f_k(t) \quad \text{(C.3)}$$

where a_{mk} are the k constants for the m^{th} waveform.

Given that the prior distribution of the transmitted signal is uniform, the optimal detector is the Maximum-Likelihood (ML) detector and it can be implemented by selecting the value of m that minimizes the Euclidian distance metric

$$D(\mathbf{r}, \mathbf{s}_m) = \sum_{k=1}^{N} (r_k - s_{mk})^2 = |\mathbf{r} - \mathbf{s}_m|^2 \quad \text{(C.4)}$$

This means that the ML detector should choose the waveform $s_m(t)$ that gives rise to the smallest error in terms of (C.4).

C.2 BPSK Detection and Performance

When using BPSK to modulate the information onto the UWB pulses only two waveforms are possible namely $s_1(t) = s(t)$ when the information bit b is +1 and $s_2(t) = -s(t)$ when it is -1. This results in $N = 1$ and only a single correlator is needed to demodulate the signal optimally. In order to minimize (C.4) one simply has to choose the sign of s_m to be the same as r, because $s_1 = -s_2$. An optimal receiver for BPSK is therefore given by

$$\hat{b} = \mathrm{sgn}\left(\int_0^T r(t)s(t)dt \right) = \mathrm{sgn}\left(s_1 + n_1 \right) \quad \text{(C.5)}$$

where $\mathrm{sgn}(\cdot)$ is the sign function. The BER of BPSK can now be found by observing the statistics of s_1 and n_1. It is shown in [11 p. 236] that n_1 is a Gaussian variable with zero mean and variance σ^2 and given that the waveform is normalized so that $\int_0^T s^2(t)\,dt = 1$, the BER can be found as

$$BER = \frac{1}{2} p\left(s_1 + n_1 < 0 \mid s_1 \right) + \frac{1}{2} p\left(s_2 + n_2 > 0 \mid s_2 \right) \quad \text{(C.6)}$$

and because of symmetry this can be written as

$$BER = p\left(s_1 + n_1 < 0 \mid s_1 \right) = p\left(n_1 > 1 \right) = \frac{1}{\sqrt{2\pi}\sigma} \int_1^\infty e^{\frac{-\alpha^2}{2\sigma^2}} d\alpha = \frac{1}{2} erfc\left(\sqrt{\frac{1}{2\sigma^2}} \right) \quad \text{(C.7)}$$

where $erfc(\cdot)$ is the complementary error function defined by

$$erfc(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-z^2} dz \quad \text{(C.8)}$$

Using the fact that the Signal-to-Noise Ratio is given by $SNR = \frac{1}{\sigma^2}$ and inserting into (C.7) gives

$$BER = \frac{1}{2} erfc\left(\sqrt{\frac{SNR}{2}} \right) \quad \text{(C.9)}$$

This result can easily be generalized to the case where more than one pulse is transmitted pr. bit. Assuming that N_s pulses are transmitted pr. bit and that the receiver sums up the output of the correlator after each pulse the probability of bit error can be written as

$$BER = p\left(N_s s_1 + n_1' < 0 \mid s_1 \right) \quad \text{(C.10)}$$

and the noise variable n_1' will now be Gaussian with zero mean and variance $N_s\sigma^2$ and the BER can then be found directly as

$$BER = \frac{1}{2} erfc\left(\sqrt{\frac{N_s \cdot SNR}{2}} \right) \quad \text{(C.11)}$$

C.3 BPPM Detection and Performance

Similar to BPSK there are only two different waveforms when BPPM is used. When the information bit $b = +1$ then $s_1(t) = s(t - \delta)$ is transmitted, where δ is the modulation index. If on the other hand $b = -1$ then $s_2(t) = s(t)$ is transmitted. As it is generally not possible to create $s_2(t)$ as a linear combination of $s_1(t)$ then $N = 2$ and two correlators are therefore needed, one matched to $s_1(t)$ and one matched to $s_2(t)$.

In order to minimize (C.4) the waveform having the largest value of $\mathbf{r}$ should be selected and its index indicates the value of m. This is the same as finding b by

$$\hat{b} = \operatorname{sgn}\left(r_1 - r_2\right) = \operatorname{sgn}\left(\int_{\delta}^{T+\delta} r(t)\, s_1(t)\, dt - \int_{0}^{T} r(t)\, s_2(t) dt\right) \quad \text{(C.12)}$$

The probability of bit error can now be found by observing the statistics of $r_1 - r_2$

$$BER = \frac{1}{2} p\left(r_1 - r_2 < 0 \mid s_1\right) + \frac{1}{2} p\left(r_1 - r_2 > 0 \mid s_2\right) \quad \text{(C.13)}$$

and by symmetry

$$BER = p\left(r_1 - r_2 < 0 \mid s_1\right) = p\left(s_1 - s_2 + n_1 - n_2 < 0 \mid s_1\right) \quad \text{(C.14)}$$

The problem is that $s_1(t)$ and $s_2(t)$ may overlap in time depending on the value of δ and r_1 and r_2 are therefore potentially correlated. Assuming $s_1(t)$ is the transmitted waveform and realizing that s_2 will then be equal to the value $\rho(\delta)$ of the autocorrelation function of $s(t)$ for lag δ, (C.14) becomes

$$BER = p\left(1 - \rho(\delta) + n_1 - n_2 < 0\right) \quad \text{(C.15)}$$

Next the statistics of the noise contribution $n_1 - n_2$ must be calculated. This has been done in [3] and the result is given by

$$n_1 - n_2 \sim N\left(0, 2\sigma^2\left(1 - \rho(\delta)\right)\right) \quad \text{(C.16)}$$

meaning that $n_1 - n_2$ is a Gaussian random variable with zero mean and variance $2\sigma^2\left(1 - \rho(\delta)\right)$. Using (C.16) in (C.15) yields

$$BER = p\left(n_1 - n_2 > 1 - \rho(\delta)\right) =$$

$$\frac{1}{\sqrt{2\pi}\sqrt{2\left(1-\rho(\delta)\right)\sigma^2}} \int_{1-\rho(\delta)}^{\infty} e^{\frac{-\alpha^2}{4(1-\rho(\delta))\sigma^2}} d\alpha = \frac{1}{2} erfc\left(\sqrt{\frac{\left(1-\rho(\delta)\right)}{4\sigma^2}}\right) \quad \text{(C.17)}$$

and by using $SNR = \frac{1}{\sigma^2}$

$$BER = \frac{1}{2} erfc\left(\sqrt{\frac{\left(1-\rho(\delta)\right) SNR}{4}}\right) \quad \text{(C.18)}$$

or if more than one pulse is used pr. bit, the BER can be found in a similar way as done with BPSK as

$$BER = \frac{1}{2} erfc\left(\sqrt{\frac{\left(1-\rho(\delta)\right) N_s \cdot SNR}{4}}\right) \quad \text{(C.19)}$$

This result shows that the BER of BPPM given by (C.19) has a penalty over that of BPSK given by the factor $\frac{2}{\left(1-\rho(\delta)\right)}$. In the case of non-overlapping BPPM $\rho(\delta) = 0$ and the penalty is equal to $\frac{1}{2}$ or 3dB, which may be seen directly from the fact that the two correlators will collect twice the noise power of the BPSK receiver. Whenever an overlap with $\rho(\delta) < 0$ is used, the performance gap between BPPM and BPSK will narrow, but BPPM is never able reach the performance of BPSK.

Another way of doing optimal BPPM detection is to use only one correlator with $s_1(t) = s(t-\delta) - s(t)$ and the decision is then made as

$$b = \operatorname{sgn}(r_1) \quad \text{(C.20)}$$

The performance of this detector will also be given by (C.19), as the same operation is performed only now by a single correlator with a larger integration period.

C.4 QPPAM Detection and Performance

In this case four different waveforms are now possible. The first information bit b_1 determines the position of the pulse as in the case of BPPM and the second information bit b_2 determines the sign of the pulse as in the case of BPSK. It is therefore necessary to have $\rho(\delta) = 0$ as the sign of the autocorrelation function is no longer known and using $\rho(\delta) \neq 0$ will therefore only degrade performance.

As in the case of BPPM the dimension of the signal is $N = 2$ and the signal can therefore be demodulated using two correlators in the same manner as was done with BPPM. In order to perform ML detection, the correlator with the largest numerical value at the output must be selected and used to estimate the transmitted bits

$$\hat{b}_1 = \operatorname{sgn}\left(|r_1| - |r_2|\right) = \operatorname{sgn}\left(\left|\int_{\delta}^{T+\delta} r(t)\, s_1(t)\, dt\right| - \left|\int_{0}^{T} r(t)\, s_2(t) dt\right|\right) \quad \text{(C.21)}$$

$$\hat{b}_2 = \operatorname{sgn}\left(\max\left(|r_1|, |r_2|\right)\right) \quad \text{(C.22)}$$

This effectively means that $\hat{b}_2$ depends on $\hat{b}_1$.

In order to find the BER, the statistics of the decision variable $|r_1| - |r_2| = |s_1 + n_1| - |s_2 + n_2|$ must first be known. It is important to understand that the probability distributions of n_1 and n_2 are independent, as was the case with non-overlapping BPPM, but s_1 and s_2 are coupled as only one of them can be zero at a time.

As r_x with $x \in \{1, 2\}$ is Gaussian with mean $\overline{r_x} \in \{-1, 0, +1\}$ and variance σ^2, which will be indicated by $r_x \sim N\left(\overline{r_x}, \sigma^2\right)$, the pdf of the numerical value $r_x' = |r_x|$ is given by

$$p(r_x') = \frac{1}{\sqrt{2\pi}\sigma}\left(\exp\left(\frac{-\left(r_x' - \overline{r_x}\right)^2}{2\sigma^2}\right) + \exp\left(\frac{-\left(r_x' + \overline{r_x}\right)^2}{2\sigma^2}\right)\right) \quad , r_x' \geq 0 \quad \text{(C.23)}$$

and the term $r_x' \sim A\left(\overline{r_x}, \sigma^2\right)$ will be used for further calculations to indicate that r_x' has a pdf as given by (C.23).

The probability that an error has occurred during reception of b_1 is then given by

$$p\left(\hat{b}_1 \neq b_1\right) = p\left(|s_1 + n_1| > |s_2 + n_2| \,|\, s_1 = 0\right) + p\left(|s_1 + n_1| < |s_2 + n_2| \,|\, s_2 = 0\right) \quad \text{(C.24)}$$

and by using symmetry and the fact that $s_1 = 0$ this reduces to

$$p\left(\hat{b}_1 \neq b_1\right) = p\left(|n_1| > |s_2 + n_2|\right) \quad \text{(C.25)}$$

Now using the substitution

$$\begin{aligned}\Gamma &= |n_1| \\ \Lambda &= |s_2 + n_2|\end{aligned} \quad \text{(C.26)}$$

and that $p(s_2 = +1 \mid s_1 = 0) = \frac{1}{2}$ and $p(s_2 = -1 \mid s_1 = 0) = \frac{1}{2}$ in (C.23) it can be seen that

$$\begin{aligned}\Gamma &\sim A(0,\sigma^2) = 2N(0,\sigma^2) \quad ,\Gamma \geq 0 \\ \Lambda &\sim A(1,\sigma^2) \qquad\qquad\qquad\quad ,\Lambda \geq 0\end{aligned} \quad \text{(C.27)}$$

Rewriting (C.25) using (C.26) yields

$$p(\hat{b}_1 \neq b_1) = p(\Gamma > \Lambda) = \int_0^\infty p(\Gamma > \Lambda = \lambda)\, p(\lambda)\, d\lambda = \int_0^\infty \int_\lambda^\infty p(\gamma)\, d\gamma \cdot p(\lambda)\, d\lambda \quad \text{(C.28)}$$

and using (C.27) gives

$$p(\hat{b}_1 \neq b_1) = \frac{1}{\pi\sigma^2} \int_0^\infty \int_\lambda^\infty \exp\left(\frac{-\gamma^2}{2\sigma^2}\right) d\gamma \cdot \left(\exp\left(\frac{-(\lambda-1)^2}{2\sigma^2}\right) + \exp\left(\frac{-(\lambda+1)^2}{2\sigma^2}\right)\right) d\lambda \quad \text{(C.29)}$$

Solving (C.29) analytically is not possible and an analytical expression for the BER can therefore not be obtained, but a very good approximation can be made if $\sigma^2 >> 1$ as

$$\sigma^2 >> 1: \quad A(1,\sigma^2) \cong N(1,\sigma^2) \quad \text{(C.30)}$$

and (C.28) can then be written as

$$\sigma^2 >> 1: \quad p(\hat{b}_1 \neq b_1) = p(\Lambda - \Gamma < 0) = p(\Delta < 0) \quad \text{(C.31)}$$

where $\Delta = \Lambda - \Gamma$ and $\Delta \sim 2N(1, 2\sigma^2)$ and (C.31) can then be written as

$$\sigma^2 >> 1: \quad p(\hat{b}_1 \neq b_1) = \frac{2}{\sqrt{2\pi}\sqrt{2}\sigma} \int_{\lambda=1}^\infty \exp\left(\frac{-\lambda^2}{4\sigma^2}\right) d\lambda = erfc\left(\sqrt{\frac{SNR}{4}}\right) = 2BER_{BPPM} \quad \text{(C.32)}$$

From (C.32) it can be seen that the bit error probability of the BPPM modulated bit of QPPAM is twice that of pure BPPM, making it unattractive at high SNRs.

The overall BER of QPPAM can now be found using either numerical integration of (C.29) or directly using (C.32), if appropriate, as

$$\begin{aligned}BER &= \frac{1}{2} p(\hat{b}_1 \neq b_1) + \frac{1}{2}\left(\left(1 - p(\hat{b}_1 \neq b_1)\right) BER_{BPSK} + p(\hat{b}_1 \neq b_1)\frac{1}{2}\right) = \\ &\frac{3}{4} p(\hat{b}_1 \neq b_1) + \left(1 - p(\hat{b}_1 \neq b_1)\right)\frac{BER_{BPSK}}{2}\end{aligned} \quad \text{(C.33)}$$

because the decision about $\hat{b}_2$ is analogous to that done in BPSK where BER_{BPSK} is the expression given in (C.9). Using these formulas one must bear in mind that they only apply when $\rho(\delta) = 0$.

Appendix D Derivation of Channel Capacities

In this appendix the channel capacities of 2-PAM, 4-PAM and M-PPM will be given analytically. As shown in Appendix C.4, it is not possible to give an analytical expression for the BER of QPPAM and an analytical expression for the channel capacity can therefore not be given either. For simplicity it is assumed that only a single pulse is transmitted pr. bit and that non-overlapping PPM is used, i.e. $N_s = 1$ and $\delta = 0$.

Given a discrete memory-less AWGN communications channel having J input values x_j where $j = 1...J$ and having output values y_k with $k = 1...K$ the mutual information I is given by Shannon [2] as

$$\begin{aligned} I &= I(X;Y) = H(X) - H(X \mid Y) \\ I &= I(Y;X) = H(Y) - H(Y \mid X) \end{aligned} \quad \text{(D.1)}$$

where $H(Z)$ is the entropy of the random variable Z defined as

$$H(Z) = \sum_i p(z_i) \log_2 \left(\frac{1}{p(z_i)} \right) \quad \text{(D.2)}$$

The channel capacity is now given as the maximum of (D.1) with respect to the distribution of x_j

$$C = \arg\max_{p(x_j)} \left(I(X;Y) \right) \quad \text{(D.3)}$$

It is a reasonable assumption that the UWB channel is symmetric and $p(x_j)$ should therefore be uniformly distributed in order to maximize (D.3). This means that the channel capacity can be found directly from (D.1).

Writing out (D.1) via (D.2) and using the input conditioned on the output yields

$$\begin{aligned} &I(X;Y) = H(Y) - \sum_{k=1}^{K-1} H(X \mid Y = y_k)\, p(y_k) = \\ &H(Y) - \sum_{k=1}^{K-1} \sum_{j=1}^{J-1} p(x_j \mid y_k) \log_2 \left(\frac{1}{p(x_j \mid y_k)} \right) p(y_k) \end{aligned} \quad \text{(D.4)}$$

and in case of the output conditioned on the input

$$\begin{aligned} &I(Y;X) = H(X) - \sum_{j=1}^{J-1} H(Y \mid X = x_j)\, p(x_j) = \\ &H(X) - \sum_{j=1}^{J-1} \sum_{k=1}^{K-1} p(y_k \mid x_j) \log_2 \left(\frac{1}{p(y_k \mid x_j)} \right) p(x_j) \end{aligned} \quad \text{(D.5)}$$

D.1 Channel Capacity of 2-PAM

Because of symmetry both $p(x_j)$ and $p(y_k)$ are uniformly distributed

$$p(x_j) = \frac{1}{2} \qquad p(y_k) = \frac{1}{2} \quad \text{(D.6)}$$

The probability of bit error p can readily be found as in (C.9) and the conditional probabilities are then given by

$$p\left(y_k \mid x_j\right)=\begin{cases} p & ,k\neq j \\ 1-p & ,k=j \end{cases} \qquad p=\frac{1}{2}erfc\left(\sqrt{\frac{SNR}{2}}\right) \quad \text{(D.7)}$$

Inserting into (D.5) gives

$$I\left(Y;X\right)=1+p\log_2\left(p\right)+\left(1-p\right)\log_2\left(1-p\right) \quad \text{(D.8)}$$

D.2 Channel Capacity of 4-PAM

Again the input is uniformly distributed, but now the output distribution is not uniform

$$p\left(x_j\right)=\frac{1}{2} \qquad p\left(y_k\right)=\sum_{j=1}^{J-1}p\left(y_k \mid x_j\right)p\left(x_j\right) \quad \text{(D.9)}$$

By observing the transmitted signal given by

$$x_j=\begin{cases} +1.5 & ,j=0 \\ +0.5 & ,j=1 \\ -0.5 & ,j=2 \\ -1.5 & ,j=3 \end{cases} \qquad E\left[X\right]=0 \qquad E\left[X^2\right]=\frac{5}{4}$$

the conditional probabilities can readily be found as

$$p\left(y_1 \mid x_0\right)=p\left(y_2 \mid x_3\right)=\frac{1}{2}\left(erfc\left(\frac{a}{2}\right)-erfc\left(\frac{3a}{2}\right)\right)$$

$$p\left(y_2 \mid x_0\right)=p\left(y_1 \mid x_3\right)=\frac{1}{2}\left(erfc\left(\frac{3a}{2}\right)-erfc\left(\frac{5a}{2}\right)\right)$$

$$p\left(y_3 \mid x_0\right)=p\left(y_0 \mid x_3\right)=\frac{1}{2}erfc\left(\frac{5a}{2}\right)$$

$$p\left(y_0 \mid x_0\right)=p\left(y_3 \mid x_3\right)=1-p\left(y_1 \mid x_0\right)-p\left(y_2 \mid x_0\right)-p\left(y_3 \mid x_0\right)$$

$$p\left(y_0 \mid x_1\right)=p\left(y_3 \mid x_2\right)=\frac{1}{2}erfc\left(\frac{a}{2}\right)$$

$$p\left(y_2 \mid x_1\right)=p\left(y_1 \mid x_2\right)=\frac{1}{2}\left(erfc\left(\frac{a}{2}\right)-erfc\left(\frac{3a}{2}\right)\right)$$

$$p\left(y_3 \mid x_1\right)=p\left(y_0 \mid x_2\right)=\frac{1}{2}erfc\left(\frac{3a}{2}\right)$$

$$p\left(y_1 \mid x_1\right)=p\left(y_2 \mid x_2\right)=1-p\left(y_0 \mid x_1\right)-p\left(y_2 \mid x_1\right)-p\left(y_3 \mid x_1\right)$$

where $a=\sqrt{\frac{SNR}{2}}$. All probabilities needed to calculate the channel capacity by (D.5) are now known.

D.3 Channel Capacity of M-PPM

Similar to the case of 2-PAM, both $p\left(x_j\right)$ and $p\left(y_k\right)$ are uniformly distributed because of symmetry

$$p\left(x_j\right)=\frac{1}{M} \qquad p\left(y_k\right)=\frac{1}{M} \quad \text{(D.10)}$$

The probability of bit error p can readily be found in a similar way as (C.18) by observing that the energy difference between decision boundaries is inversely proportional to M. The conditional probabilities are then given by

$$p\left(y_k \mid x_j\right)=\begin{cases} p & ,k \neq j \\ 1-\left(M-1\right)p & ,k=j \end{cases} \qquad p=\frac{1}{M}erfc\left(\sqrt{\frac{SNR}{2M}}\right) \quad \text{(D.11)}$$

Inserting into (D.5) gives

$$I\left(Y;X\right)=\log_2\left(M\right)+\left(M-1\right)p\log_2\left(p\right)+\left(1-p\left(M-1\right)\right)\log_2\left(1-p\left(M-1\right)\right) \quad \text{(D.12)}$$

Appendix E Asynchronous Time-Hopping Performance using Random Codes

When evaluating asynchronous Time-Hopping (TH) systems, it is easily seen that the collision probability is given by

$$P_{col} = 1 - \left(1 - \frac{2}{SF}\right)^{K-1} \quad \text{(E.1)}$$

with SF being the spreading factor and K the number of users when a uniformly distributed random hopping code is used. The reason is that two timeslots can potentially collide with the user of interest because of the time difference between users. Assuming that the time offset of all other users relative to the user of interest is uniformly distributed, the probability of having the offset λ that lies in the interval $-T_{mono} < \lambda < T_{mono}$ relative to the timeslot used by the user of interest, is easily seen to be

$$P_c(\lambda) = \frac{P_{col}}{\frac{2}{SF}} = \left(1 - \left(1 - \frac{2}{SF}\right)^{K-1}\right)\frac{SF}{2} \quad \text{(E.2)}$$

with T_{mono} being the pulse width of the monocycle. Defining the auto-correlation of the monocycle as

$$\rho(\lambda) = \int_{-T_{mono}}^{T_{mono}} p_{rx}(t)\, p_{rx}(t-\lambda)\, dt \quad \text{(E.3)}$$

normalized so that

$$\rho(0) = 1 \quad \text{(E.4)}$$

and limiting the modulation scheme to being BPSK and at the same time considering only one collision at a time, the decision statistics of a given offset is $\operatorname{sgn}(1-\rho(\lambda)+n)$ and $\operatorname{sgn}(1+\rho(\lambda)+n)$ each with probability ½ and the noise variable $n \sim N(0,\sigma^2)$ with σ^2 being the noise variance. The assumption of only one collision at a time is reasonable when K<<SF, but will be unjustified if K is comparable to SF. If this is the case more than one user collides with the user of interest and the decision statistic will therefore contain contribution from all these users.

Assuming K<<SF, the BER can then be calculated by averaging over all possible offsets

$$BER \simeq (1 - P_{col})\, BER_{BPSK} + \frac{1}{2}\int_{-T_{mono}}^{T_{mono}} P_c(\lambda)\frac{1}{\sqrt{2\pi\sigma^2}}\left(\int_{1+\rho(\lambda)}^{\infty} e^{\frac{-\alpha^2}{2\sigma^2}} d\alpha + \int_{1-\rho(\lambda)}^{\infty} e^{\frac{-\alpha^2}{2\sigma^2}} d\alpha\right) d\lambda \quad \text{(E.5)}$$

and inserting (E.2) gives

$$BER \simeq (1 - P_{col})\, BER_{BPSK} + \frac{SF}{8}\left(1 - \left(1 - \frac{2}{SF}\right)^{K-1}\right)\int_{-T_{mono}}^{T_{mono}} erfc\left(\frac{1+\rho(\lambda)}{\sqrt{2\sigma^2}}\right) + erfc\left(\frac{1-\rho(\lambda)}{\sqrt{2\sigma^2}}\right) d\lambda \quad \text{(E.6)}$$

Even in this simplest form of collision, where only one user collides with the user of interest, a closed form solution to the integral of (E.6) can not be given. Using the fact that well-behaved monocycles will have $|\rho(\lambda \neq 0)| << 1$, it can be seen that the BER of asynchronous TH will be close that synchronous TH and potentially even lower. But the exact BER can only be determined by numerical integration of (E.6) or by simulations.

Appendix F PSD of BPSK Modulated TH and DS Signals

In this appendix expressions are derived for the PSD of BPSK modulated TH and DS signals after spreading and modulation has been applied. This is important as strict regulations apply to the PSD of UWB signals.

Compared to the signal models outlined in (2.10) and (2.17), it is assumed that only BPSK modulation is supported as this is the best performing modulation scheme. The asynchronous offset $\tau^{(k)}$ is set to zero as it has no effect on the PSD of the signal. Furthermore, in the case of TH it is assumed that only one pulse is transmitted pr. bit meaning $N_s = 1$. In the notation used N_p is the period of the hopping code and $T_f = SF \cdot T_{mono}$ is the frame time with SF being the spreading factor and T_{mono} the duration of the monocycle $p_{tx}(t)$.

F.1 PSD of TH Signals

In order to include the effect of the periodic hopping code used in a realistic TH system, the transmitted signal must be separated into blocks having a length equal to the period of the hopping code. This is done by expressing the transmitted signal $s^{(k)}(t)$ as

$$s^{(k)}(t) = \sum_i s_i^{(k)}\left(t - iN_pT_f\right) \quad \text{(F.1)}$$

where $s_i^{(k)}\left(t - iN_pT_f\right)$ is the modulated signal over the i^{th} period of the hopping code as given by

$$s_i^{(k)}(t) = \sum_{h=0}^{N_p-1} \alpha^{(k)}\left(iN_p + h\right) p_{tx}\left(t - hT_f - c_{TH}^{(k)}(h)T_c\right) \quad \text{(F.2)}$$

To find the PSD of the transmitted signal the autocorrelation function can be found as

$$\begin{aligned} E\left[s^{(k)}(t)s^{(k)}(t-\tau)\right] &= E\left[\sum_{i,l} s_i^{(k)}\left(t - iN_pT_f\right)s_l^{(k)}\left(t - lN_pT_f - \tau\right)\right] \\ &= \sum_{i,l} E\left[s_i^{(k)}\left(t - iN_pT_f\right)s_l^{(k)}\left(t - lN_pT_f - \tau\right)\right] \end{aligned} \quad \text{(F.3)}$$

where the expectation is taken with respect to both the time t and the code $c_{TH}^{(k)}$. The expression of (F.3) may be split into two as

$$\sum_{i=l} E\left[s_i^{(k)}\left(t - iN_pT_f\right)s_l^{(k)}\left(t - lN_pT_f - \tau\right)\right] + \sum_{i\neq l} E\left[s_i^{(k)}\left(t - iN_pT_f\right)s_l^{(k)}\left(t - lN_pT_f - \tau\right)\right] \quad \text{(F.4)}$$

Concentration on the left part of (F.4) first, taking the expectation over t it may be written as

$$\begin{aligned} &\sum_{i=l} E\left[s_i^{(k)}\left(t - iN_pT_f\right)s_l^{(k)}\left(t - lN_pT_f - \tau\right)\right] = \sum_{i=l} E\left[\frac{1}{N_pT_f}\int_0^{N_pT_f} s_i^{(k)}\left(t - iN_pT_f\right)s_l^{(k)}\left(t - lN_pT_f - \tau\right)dt\right] \\ &= \frac{1}{N_pT_f} E\left[\int_{-\infty}^{\infty} s_i^{(k)}(t)s_i^{(k)}(t-\tau)\,dt\right] = \frac{1}{N_pT_f} E\left[R_i(\tau)\right] \end{aligned} \quad \text{(F.5)}$$

with $R_i(\tau)$ being the autocorrelation of $s_i^{(k)}(t)$. Turning to the right part of (F.4) and performing the substitution $d = i - l$ results in

$$\sum_i \sum_{d\neq 0} E\left[s_i^{(k)}\left(t - iN_pT_f\right)s_{i-d}^{(k)}\left(t - (i-d)N_pT_f - \tau\right)\right] \quad \text{(F.6)}$$

Next, by taking the expectation over t, it may be written as

$$\sum_i \sum_{d\neq 0} E\left[\frac{1}{N_p T_f}\int_0^{N_p T_f} s_i^{(k)}\left(t-iN_pT_f\right)s_{i-d}^{(k)}\left(t-(i-d)N_pT_f-\tau\right)dt\right]=$$
$$\frac{1}{N_pT_f}\sum_{d\neq 0} E\left[\int_{-\infty}^{\infty} s_i^{(k)}(t)\,s_{i-d}^{(k)}\left(t+dN_pT_f-\tau\right)dt\right]=\frac{1}{N_pT_f}\sum_{d\neq 0} E\left[R_{i,i-d}\left(\tau+dN_pT_f\right)\right] \quad \text{(F.7)}$$

where $R_{i,i-d}(\tau)$ is the cross correlation between $s_i^{(k)}(t)$ and $s_{i-d}^{(k)}(t)$. However, as $s_i^{(k)}(t)$ is a stationary process $R_{i,i-d}(\tau)$ is independent of d given that $d\neq 0$.

The correlation function of (F.3) can therefore be written using (F.5) and (F.7) as

$$E\left[s^{(k)}(t)\,s^{(k)}(t-\tau)\right]=\frac{1}{N_pT_f}\left(E\left[R_i(\tau)\right]+\sum_{d\neq 0}E\left[R_{i,i-d}\left(\tau+dN_pT_f\right)\right]\right) \quad \text{(F.8)}$$

To find the PSD of $s^{(k)}(t)$, named $\Sigma_{TH}^{(k)}(f)$, the Fourier transform $F\{\bullet\}$ of (F.8) must be calculated. Doing this gives

$$\Sigma_{TH}^{(k)}(f)=\frac{1}{N_pT_f}E\left[F\left\{R_i(\tau)+\sum_{d\neq 0}R_{i,i-d}\left(\tau+dN_pT_f\right)\right\}\right]=$$
$$\frac{1}{N_pT_f}E\left[F\left\{\overbrace{R_i(\tau)-R_{i,i-q}(\tau)}^{Non-periodic}+\overbrace{R_{i,i-q}(\tau)+\sum_{d\neq 0}R_{i,i-d}\left(\tau+dN_pT_f\right)}^{Periodic}\right\}\right] \quad ,q\neq 0 \quad \text{(F.9)}$$

Introducing the variable q makes it possible to separate (F.9) into a periodic and non-periodic correlation function. This is important as the Fourier transform of a periodic function is discrete and $\Sigma_{TH}^{(k)}(f)$ will therefore potentially contain both discrete and continuous components. Performing the Fourier transform in (F.9) yields

$$\Sigma_{TH}^{(k)}(f)=\frac{1}{N_pT_f}\left(\overbrace{E\left[\left|S_i(f)\right|^2\right]-E\left[S_i(f)S_{i-q}^*(f)\right]}^{Continuous}+\overbrace{\Delta(f)E\left[S_i(f)S_{i-d}^*(f)\right]}^{Discrete}\right)= \quad \text{(F.10)}$$
$$\frac{1}{N_pT_f}\left(E\left[\left|S_i(f)\right|^2\right]+\left(\Delta(f)-1\right)E\left[S_i(f)S_{i-d}^*(f)\right]\right)$$

with $S_i(f)$ being the Fourier transform of $s_i^{(k)}(t)$ and $\Delta(f)$ is the Fourier transform of a train of delta impulses spaced N_pT_f apart

$$\Delta(f)=\frac{1}{N_pT_f}\sum_d \delta\left(f-\frac{d}{N_pT_f}\right) \quad \text{(F.11)}$$

Next the expression for $S_i(f)$ can be found directly as

$$S_i(f)=P_{tx}(f)\sum_{h=0}^{N_p-1}\alpha^{(k)}\left(iN_p+h\right)e^{-j2\pi f\left(hT_f+c_{TH}^{(k)}(h)T_c\right)} \quad \text{(F.12)}$$

with $P_{tx}(f)$ being the Fourier transform of $p_{tx}(t)$ and $j=\sqrt{-1}$. The expectations needed in (F.10) can now be written as

$$E\left[S_i(f)S_n^*(f)\right]=\left|P_{tx}(f)\right|^2 E\left[\sum_{h=0}^{N_p-1}\sum_{g=0}^{N_p-1}\alpha^{(k)}\left(iN_p+h\right)e^{-j2\pi f\left(hT_f+c_{TH}^{(k)}(h)T_c\right)}\alpha^{(k)}\left(nN_p+g\right)e^{j2\pi f\left(gT_f+c_{TH}^{(k)}(g)T_c\right)}\right] \quad \text{(F.13)}$$

where n can be any integer. Evaluating (F.13) for the special case of BPSK modulation, it is easily seen that

$$E\left[S_i(f)S_n^*(f)\right]=\begin{cases}0 & ,n\neq i\\ N_p\left|P_{tx}(f)\right|^2 & ,n=i\end{cases} \quad \text{(F.14)}$$

as the expectation of the pulse amplitude is zero when $n\neq i$, assuming that all bits are independent. In the case of $n=i$ the only contribution comes from $g=h$ and in this case the exponentials will cancel each other. The resultant PSD of the BPSK modulated signal is therefore

$$\Sigma_{TH}^{(k)}(f)=\frac{\left|P_{tx}(f)\right|^2}{T_f} \quad \text{(F.15)}$$

This shows how the PSD of a TH system using BPSK modulation and one pulse pr. bit is independent of the applied periodic hopping code and has no discrete components. The only way to modify the PSD of the signal is therefore to modify the monocycle $p_{tx}(t)$.

F.2 PSD of DS Signals

In the case of a BPSK modulated DS system, the signal consists of the coded monocycles $\varphi^{(k)}(t)$ multiplied by the information bits $\alpha^{(k)}(n)$ as

$$s^{(k)}(t)=\sum_n \alpha^{(k)}(n)\varphi^{(k)}\left(t-nT_f\right) \quad \text{(F.16)}$$

and the autocorrelation of the resultant signal is therefore

$$E\left[s^{(k)}(t)s^{(k)}(t-\tau)\right]=E\left[\sum_n \alpha^{(k)}(n)\varphi^{(k)}\left(t-nT_f\right)\sum_l \alpha^{(k)}(l)\varphi^{(k)}\left(t-lT_f-\tau\right)\right] \quad \text{(F.17)}$$

Assuming that the bits are independent, only the term where $n=l$ will contribute to (F.17) and it may therefore be written as

$$E\left[\sum_{n=l}\varphi^{(k)}\left(t-nT_f\right)\varphi^{(k)}\left(t-lT_f-\tau\right)\right]=\frac{1}{T_f}E\left[\int_{-\infty}^{\infty}\varphi^{(k)}(t)\varphi^{(k)}(t-\tau)\,dt\right]=\frac{1}{T_f}E\left[R_\varphi(\tau)\right] \quad \text{(F.18)}$$

by taking the expectation over t and introducing the autocorrelation function of $\varphi^{(k)}(t)$. It is consequently possible to find the PSD of $s^{(k)}(t)$ by taking the Fourier transform $F\{\bullet\}$ of $\varphi^{(k)}(t)$ as

$$\Phi^{(k)}(f)=F\left\{\varphi^{(k)}(t)\right\}=F\left\{\sum_{i=0}^{SF-1}c_{DS}^{(k)}(i)\,p_{tx}\left(t-iT_{mono}\right)\right\}=P_{tx}(f)C_{DS}^{(k)}(f) \quad \text{(F.19)}$$

where $C_{DS}^{(k)}(f)$ is the spectrum of the spreading code as given by

$$C_{DS}^{(k)}(f) = \sum_{i=0}^{SF-1} c_{DS}^{(k)}(i) e^{-j2\pi fiT_{mono}} \quad \text{(F.20)}$$

The PSD can now be found as

$$\Sigma_{DS}^{(k)}(f) = F\left\{E\left[s^{(k)}(t)s^{(k)}(t-\tau)\right]\right\} = \frac{1}{T_f} E\left[F\left\{R_\varphi(\tau)\right\}\right] = \frac{1}{T_f}\Phi^{(k)}(f)\Phi^{(k)^*}(f) = \frac{\left|P_{tx}(f)\right|^2}{T_f}\left|C_{DS}^{(k)}(f)\right|^2 \quad \text{(F.21)}$$

and is thus directly dependent on the spectrum of the spreading code. Still, it has the same direct dependency on the spectrum of the monocycle as the case was for the TH system. It should also be noted that the expression in (F.21) simplifies to (F.15) when the DS spreading code is converted into a TH hopping code, which is possible by letting all but one value of the code be equal to zero.

An important difference between the TH and DS system is therefore that whereas it is not possible to modify the PSD by altering the coding in a TH system, this is possible in a DS system.

Appendix G Impact of ISI at the Matched Filter Output

In this appendix the relationship between Inter Symbol Interference (ISI) at the matched filter output and the spreading code auto-correlation will be given. This is interesting in order to evaluate the performance of a wireless system in a multipath radio channel.

Excluding noise, the output of the receiver after the matched filter can be written as

$$y^{(k)}(n)=\sum_{i=0}^{N-1}a_i^{(k)}\int_{nT_f}^{(n+1)T_f}\sum_{j=0}^{\infty}\alpha^{(k)}(j)\varphi^{(k)}\left(t-jT_f-iT_{ch}\right)\varphi^{(k)}\left(t-nT_f\right)dt \quad \text{(G.1)}$$

By introducing the left and right auto-correlation functions of the coded monocycles $\varphi^{(k)}(t)$ as

$$\rho_l(\tau_m)=\int_0^{\tau_m}\varphi^{(k)}(t-\tau_m)\varphi^{(k)}(t)\,dt \quad \text{(G.2)}$$

$$\rho_r(\tau_m)=\int_{\tau_m}^{T_f}\varphi^{(k)}(t-\tau_m)\varphi^{(k)}(t)\,dt \quad \text{(G.3)}$$

it is possible to rewrite (G.1) to

$$y^{(k)}(n)=\sum_{i=0}^{N-1}a_i^{(k)}\left(\alpha^{(k)}\left(\left\lfloor n-\frac{iT_{ch}}{T_f}\right\rfloor\right)\rho_l\left(iT_{ch}\left(\bmod T_f\right)\right)+\alpha^{(k)}\left(\left\lceil n-\frac{iT_{ch}}{T_f}\right\rceil\right)\rho_r\left(iT_{ch}\left(\bmod T_f\right)\right)\right) \quad \text{(G.4)}$$

The expression in (G.4) basically states that every tap in the channel model will give a contribution from two adjacent bits. If, however, the argument to the left auto-correlation function is zero, then only a single bit will contribute, as is the case for $i=0$. It is therefore possible to separate the desired signal from the ISI as

$$y^{(k)}(n)=\overbrace{SF\cdot a_0^{(k)}\cdot\alpha^{(k)}(n)}^{Desired}+\overbrace{e_{ISI}(n)}^{ISI} \quad \text{(G.5)}$$

with

$$e_{ISI}(n)=\sum_{i=1}^{N-1}a_i^{(k)}\left(\alpha^{(k)}\left(\left\lfloor n-\frac{iT_{ch}}{T_f}\right\rfloor\right)\rho_l\left(iT_{ch}\left(\bmod T_f\right)\right)+\alpha^{(k)}\left(\left\lceil n-\frac{iT_{ch}}{T_f}\right\rceil\right)\rho_r\left(iT_{ch}\left(\bmod T_f\right)\right)\right) \quad \text{(G.6)}$$

As the actual channel is not known, it is assumed that $N>>1$ and $\frac{T_{ch}}{T_f}<<1$, which is realistic in a NLOS channel having a lot of multipaths. With these assumptions, it is possible to model the argument of the auto-correlation functions τ_m as a random variable x_m being uniformly distributed in the interval $\left[0,T_f\right]$.

It is now interesting to observe the statistics of $e_{ISI}(n)$. The mean is

$$\begin{aligned}E\left[e_{ISI}(n)\right]&=\sum_{i=1}^{N-1}a_i^{(k)}\left(E\left[\alpha^{(k)}\left(\left\lfloor n-\frac{iT_{ch}}{T_f}\right\rfloor\right)\rho_l(x_m)+\alpha^{(k)}\left(\left\lceil n-\frac{iT_{ch}}{T_f}\right\rceil\right)\rho_r(x_m)\right]\right)=\\ &\sum_{i=1}^{N-1}a_i^{(k)}\left(E\left[\alpha^{(k)}\left(\left\lfloor n-\frac{iT_{ch}}{T_f}\right\rfloor\right)\rho_l(x_m)\right]+E\left[\alpha^{(k)}\left(\left\lceil n-\frac{iT_{ch}}{T_f}\right\rceil\right)\rho_r(x_m)\right]\right)=0\end{aligned} \quad \text{(G.7)}$$

as x_m is uniformly distributed and $\rho_l(x_m)+\rho_r(x_m)=\int_0^{T_f}\varphi^{(k)}(t-x_m)\varphi^{(k)}(t)\,dt$. The variance is

$$E\left[e_{ISI}^2(n)\right]=\sum_{i=1}^{N-1}\left(a_i^{(k)}\right)^2 E\left[\left(\alpha^{(k)}\left(\left\lfloor n-\frac{iT_{ch}}{T_f}\right\rfloor\right)\rho_l(x_m)+\alpha^{(k)}\left(\left\lceil n-\frac{iT_{ch}}{T_f}\right\rceil\right)\rho_r(x_m)\right)^2\right]=$$
$$\sum_{i=1}^{N-1}\left(a_i^{(k)}\right)^2\left(E\left[\rho_l^2(x_m)\right]+E\left[\rho_r^2(x_m)\right]+2E\left[\rho_l(x_m)\rho_r(x_m)\alpha^{(k)}\left(\left\lfloor n-\frac{iT_{ch}}{T_f}\right\rfloor\right)\alpha^{(k)}\left(\left\lceil n-\frac{iT_{ch}}{T_f}\right\rceil\right)\right]\right) \quad \text{(G.8)}$$

By symmetry this can be reduced to

$$E\left[e_{ISI}^2(n)\right]=2\left(E\left[\rho_r^2(x_m)\right]+E\left[\rho_l(x_m)\rho_r(x_m)x_b\right]\right)\sum_{i=1}^{N-1}\left(a_i^{(k)}\right)^2 \quad \text{(G.9)}$$

where x_b is a Bernoulli random variable with probability ½, but as $E[x_b]=0$ this term will disappear and the result is

$$E\left[e_{ISI}^2(n)\right]=E\left[\rho_r^2(x_m)\right]\cdot 2\sum_{i=1}^{N-1}\left(a_i^{(k)}\right)^2 \quad \text{(G.10)}$$

and the ISI is therefore directly proportional to the variance of the auto-correlation function as can be found from

$$E\left[\rho_r^2(x_m)\right]=\frac{1}{T_f}\int_0^{T_f}\rho_r^2(x_m)\,dx_m \quad \text{(G.11)}$$

To determine this variance, knowledge of the monocycle and spreading code is needed. Given that the monocycle $p_{rx}(t)$ and spreading code $c^{(k)}(t)$ belongs to user k, it is possible to write $\varphi^{(k)}(t)$ as

$$\varphi^{(k)}(t)=c^{(k)}(t)\sum_{j=0}^{SF-1}p_{rx}(t-jT_{mono})=c^{(k)}(t)\,p'_{rx}(t) \quad \text{(G.12)}$$

with $p'_{rx}(t)$ being SF consecutive copies of $p_{rx}(t)$. Using this in (G.3) and squaring yields

$$\rho_r^2(x_m)=\left|\int_{x_m}^{T_f}c^{(k)}(t-x_m)\,p'_{rx}(t-x_m)\cdot c^{(k)}(t)\,p'_{rx}(t)\,dt\right|^2 \quad \text{(G.13)}$$

which directly shows how the auto-correlation is dependent on both the spreading code and the monocycle. It is however not possible to separate the expression in (G.13) into two separate parts; one dependent on the spreading code and one dependent on the monocycle. Therefore in order to show how the spreading code influences (G.13), upper bounding the expression by Schwarz's inequality gives

$$\rho_r^2(x_m)=\left|\int_{x_m}^{T_f}c^{(k)}(t-x_m)\,p'_{rx}(t-x_m)\cdot c^{(k)}(t)\,p'_{rx}(t)\,dt\right|^2<\left|\int_{x_m}^{T_f}p'_{rx}(t-x_m)\,p'_{rx}(t)\,dt\right|^2\left|\int_{x_m}^{T_f}c^{(k)}(t-x_m)\,c^{(k)}(t)\,dt\right|^2 \quad \text{(G.14)}$$

Next, the expectation of the rightmost part of (G.14) must be found with respect to x_m, but as both parts of the expression involves x_m this is not directly possible. Making the approximation that the monocycle autocorrelation function is well modeled as a delta function, it is possible to write the expectation of (G.14) as

$$E\left[\rho_r^2\left(x_m\right)\right] < E\left[\left|\int_{x_m}^{T_f} c^{(k)}\left(t-x_m\right)c^{(k)}\left(t\right)dt\right|^2\right] \quad (7.15)$$

where the expectation is taken with respect to both the random offset x_m and the code $c^{(k)}\left(t\right)$. The ISI component given by (G.10) is thus upper-bounded by the average squared code auto-correlation multiplied by a constant given by the channel impulse response as

$$E\left[e_{ISI}^2\left(n\right)\right] < E\left[\left|\int_{x_m}^{T_f} c^{(k)}\left(t-x_m\right)c^{(k)}\left(t\right)dt\right|^2\right] \cdot 2\sum_{i=1}^{N-1}\left(a_i^{(k)}\right)^2 \quad (G.16)$$

This makes it possible to directly compare a TH system with that of a DS system based on the difference in their spreading codes.

Appendix H A Lower Bound of the Inverse Correlation Matrix

Given a $K \times K$ symmetric positive definite correlation matrix $\mathbf{R}$ with the value SF as diagonal elements, it is shown in this appendix that the trace of $\mathbf{R}^{-1}$ must satisfy

$$tr\left(\mathbf{R}^{-1}\right) \geq \frac{K}{SF} \quad \text{(H.1)}$$

Doing a SVD factorization of $\mathbf{R}$ yields

$$\mathbf{R} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T \quad \text{(H.2)}$$

where $\mathbf{U}$ and $\mathbf{V}$ are orthogonal $K \times K$ matrices holding the left and right singular vectors respectively and $\boldsymbol{\Sigma}$ is diagonal containing the singular values. As $\mathbf{R}$ is symmetric this can be reduced to

$$\mathbf{R} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^T \quad \text{(H.3)}$$

with $\boldsymbol{\Sigma} = diag\left(\lambda_1, \lambda_2, \cdots, \lambda_K\right)$ holding the eigenvalues of $\mathbf{R}$, which must all be positive, and $\mathbf{U}$ being the eigenvectors of $\mathbf{R}$. The inverse can now be easily calculated by

$$\mathbf{R}^{-1} = \mathbf{U}^T\boldsymbol{\Sigma}^{-1}\mathbf{U} \quad \text{(H.4)}$$

resulting in

$$\lambda_i' = \frac{1}{\lambda_i} \quad \text{(H.5)}$$

with λ_i' being the eigenvalues of $\mathbf{R}^{-1}$. Using

$$tr\left(\mathbf{R}\right) = \sum_{i=1}^{K} \lambda_i = K \cdot SF \quad \text{(H.6)}$$

and from (H.5) a similar expression for the inverse can be given

$$tr\left(\mathbf{R}^{-1}\right) = \sum_{i=1}^{K} \lambda_i' = \sum_{i=1}^{K} \frac{1}{\lambda_i} \quad \text{(H.7)}$$

In order to minimize (H.7) under the constraint of (H.6), it can be shown using Lagrange multipliers that λ_i must be independent of i and therefore

$$\lambda_i = SF \quad \text{(H.8)}$$

meaning that

$$tr\left(\mathbf{R}^{-1}\right) \geq \frac{K}{SF} \quad \text{(H.9)}$$