Design of Wireless Communication Systems – Issues on Synchronization, Channel Estimation and Multi-Carrier Systems

Fredrik Tufvesson

Lund 2000

Department of Applied Electronics
Lund University
Box 118
SE-221 00 LUND
Sweden

No. 19
ISSN 1402-8662
ISBN 91-7874-059-2

Printed in Sweden by KFS AB, Lund
August 2000

Abstract

This thesis deals with certain aspects in the design of wireless communications systems. It is focused on problems related to the mobile or wireless channel: synchronization, channel estimation and design of wireless orthogonal frequency division multiplex (OFDM) systems. There is a short introduction to the field of wireless systems and a deeper review of pervious work and the state of the art in each of the research fields. Throughout the thesis the goal has been to analyze the problems analytically, deriving expressions for the resulting bit error rate (BER) or synchronization performance.

The work on channel estimation is focused on the influence of the pilot pattern in OFDM systems and on the possibility of compensating the signal in advance for impairments introduced by the channel. The BER is derived and compared when using different pilot patterns in an OFDM system. For the pre-compensated system the BER, maximum block length, and resulting peak-to-average power ratio are investigated when compensating the power attenuation and/or the phase rotation introduced by the channel.

The work on wireless OFDM system design is directed towards preamble-based synchronization and optimization of the sub-channel bandwidth. A new synchronizer structure is introduced for time and frequency synchronization. In this synchronizer the order of correlation and multiplication has been changed compared to conventional synchronizers, which results in significantly improved detection performance. The sub-channel bandwidth is optimized based on the expected Doppler frequency and excess delay of the channel.

The use of superimposed pilot sequences is introduced for time and frequency synchronization in OFDM systems, as well as for channel estimation in flat fading channels. As opposed to conventional pilot-based systems, the pilot sequence is linearly added to the data sequence and therefore transmitted simultaneously and in the same band as the data sequence. The performance when using superimposed pilot sequences for synchronization and for joint channel estimation and detection is analyzed. The scheme is also compared to other well known pilot-based and differential detection methods.

Key words: Wireless, orthogonal frequency division multiplex, synchronization, channel estimation, detection, pilot pattern, pre-compensation, precoding, sub-channel bandwidth, preambles, superimposed pilot sequences

Preface

When I started as a Ph.D. student my advisor, Torleiv Maseng, explained to me that OFDM had a potential of being an efficient and attractive modulation technique for future wireless and mobile systems. There were, however, still some key components that had to be improved and required further research. Channel estimation was one of these areas, synchronization another. Efficient low-complex channel estimation techniques were required to enable coherent detection and higher order modulation schemes, while robust synchronization techniques were required for operation in highly disturbed environments.

We decided to start with channel estimation and a study of channel estimation techniques. It was often stated that the loss and overhead caused by the pilot symbols in coherent systems were too large. Therefore we looked for, and found, a way to reduce the number of required pilot symbols. When studying the influence of the channel in OFDM systems we realized that the symbols could not be arbitrarily long; the channel changes have to be small within each symbol. On the other hand, the excess delay of the channel called for long symbols. We therefore started to investigate the optimal symbol length (and thereby the optimal sub-channel bandwidth) and found an analytical expression of how to make a compromise between these channel related effects.

When presenting the results on channel estimation I got the question from a guest professor, Mike Faulkner, if I could say something about the time a channel estimate is valid. The answer I gave was that it depends on the Doppler frequency, but I was not able to give any figures. We discussed the question more deeply and Mike suggested using an analysis similar to the channel estimation paper to determine the maximal delay. I changed the analysis accordingly and derived expressions for the maximal block lengths in a system where channel estimation tasks are performed in the base station only.

One day an ASIC designer at the department, Shousheng He, asked me why PN-sequences were not used for synchronization in OFDM systems. He had some problem getting the synchronization unit to work properly and we thought that maybe one could combine a PN-sequence with the OFDM signal in some way. I discussed the problem with Peter Hoeher, Mike Faulkner and Ove Edfors and we decided that the idea was worth further investigations. The resulting superimposed pilot system showed good synchronization performance

also for low SNRs; it really seemed to be a good alternative to conventional synchronization. I started to play around with the parameters of the superimposed pilot sequences and compared the results to that of conventional synchronization. By using only the pilot sequence the detection performance was much better compared to conventional synchronization and therefore we analyzed this preamble-based synchronization system further. We were also thinking about the possibilities of using the superimposed pilot sequences for channel estimation such that no other pilot symbols were required. Peter Hoeher came up with a suitable receiver structure and his simulations showed very promising results. Later, I made an analytical analysis and investigated the system further. After that, Ove and I were discussing how to compare the detection scheme to other schemes. We tried to use the analysis derived for superimposed pilot sequences and it worked well for the most important detection methods. So in this way we were able to compare the behavior for different Doppler rates and SNR values.

As seen above many persons have been involved in my research. All the time Torleiv has been by my side to help me, but mostly as guide for future research areas. The research has been focused on channel related problems in wireless communication systems, mostly synchronization and channel estimation problems. The issues presented reflect the research areas I have been involved in during my years at the Department of Applied Electronics and should therefore not be seen as a complete overview of the system design process.

The thesis consists of an overview of wireless communication systems, a background of the research area and seven papers. There is a short description of the papers and their main results after the background section. The papers have been reformatted to improve readability, but except where explicitly marked the content is the same as originally published. The included papers are:

A F. Tufvesson and T. Maseng. Pilot assisted channel estimation for OFDM in mobile cellular systems. In *Proc. IEEE Vehicular Technology Conference*, pages 1639–1643, Phoenix, USA, May 1997.

B F. Tufvesson and T. Maseng. Optimization of sub-channel bandwidth for mobile OFDM systems. In D. Everitt and M. Rumsewicz, editors, *Multi-access, mobility and teletraffic - advances in wireless networks*, pages 103–114. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.

C F. Tufvesson, M. Faulkner, and T. Maseng. Pre-compensation for Rayleigh fading channels in time division duplex OFDM systems. *Accepted for publication in Wireless Personal Communications*, 2000.

D F. Tufvesson and O. Edfors. Preamble-based time and frequency synchronization for OFDM systems. *Submitted to IEEE Journal on Selected Areas in Communications*, January 2000.

E F. Tufvesson, M. Faulkner, P. Hoeher, and O. Edfors. OFDM time and frequency synchronization by spread spectrum pilot technique. In *Proc. Eighth Communication Theory Mini-Conference in conjunction with IEEE ICC '99*, pages 115–119, Vancouver, Canada, June 1999.

F F. Tufvesson and P. Hoeher. Channel estimation using superimposed pilot sequences. *Submitted to IEEE Transactions on Communications*, April 2000.

G F. Tufvesson and O. Edfors. A comparative analysis between different detection schemes in flat Rayleigh-fading channels. In *Proc. IEEE Vehicular Technology Conference*, pages 1205-1209, Tokyo, Japan, May 2000.

Parts of the contents in the thesis and papers have also been published elsewhere. In principle I have been involved in four different research projects, where I have been the author or co-author of the following publications:

- Channel estimation in OFDM systems [69] [76] [77] [80] [81] [82]
- Wireless OFDM system design [70] [79] [83]
- Preamble-based time and frequency synchronization for OFDM [72] [73] [74]
- Channel estimation and synchronization by superimposed pilot sequences [32] [33] [71] [75] [78].

Acknowledgments

During my time as a Ph.D. student, I have had the opportunity to get to know a lot of interesting and clever people. I would like to thank you all who, in some way or the other, have contributed to this work or to the great social environment.

Especially, I would like to thank Ove Edfors for being such a good friend, for creating such a good atmosphere and for sharing his wide knowledge about almost everything. There is always time for questions and discussions, no matter if they are about details in a signal processing algorithm or about the suitability of sitting in a canoe when chased by a hippopotamus.

I would also like to specially thank Johan Hokfelt. We started at the department at the same time and it has really been a pleasure having such a good colleague being on the same level. It is not often that you have the privilege to work in such close cooperation with a friend and classmate from junior school.

I am also grateful to Torleiv Maseng for introducing me to the world of wireless communications and for starting up many of the projects presented in the thesis. Ever since I joined the department he has been by my side sharing his knowledge and intuition. Without him, there would definitely not be any thesis at all.

There are many other persons worth mentioning in the radio communications group and at the department, both for their help and for contributing to such a good atmosphere. I would especially like to thank Carl Fredrik Leanderson, Daniel Landström and Bengt-Arne Molin, the latter also for proofreading and providing valuable suggestions regarding the manuscript.

I would also like to thank some people outside the department for fruitful cooperation. I have had the privilege to work together with Mike Faulkner for a long time. First, he was a visiting professor at the department and later I visited Victoria University, Melbourne, Australia, for half a year. It is really inspiring to work together with such a person interested in everything related to wireless communication, no matter if it is about linearizing amplifiers or optimizing PN sequences. I also really appreciate the cooperation with Telia Research, Malmö, and especially working together with Peter Karlsson and Christian Bergljung. I have really enjoyed the open atmosphere at Telia Research and there I learned a lot about standardization and wireless systems in practice. I

have also very much enjoyed working together with Peter Hoeher at University of Kiel, Germany. It has really been a pleasure having the opportunity to cooperate with such a clever and competent person.

Finally, I am very grateful to my wife, Petra, who has encouraged me all the time and to my family and friends who have distracted me from work during the same period.

This work has been financed by NUTEK (The Swedish National Board for Industrial and Technical Development), in the telecommunications program (project no. 8422-94-06195). I would like to thank the following organizations and foundations, providing financial support for conferences and visits abroad: *Claes Adelskölds minnesfond*, *Kungliga fysiografiska sällskapet*, *Telefonaktiebolaget L M Ericssons stiftelse för främjande av elektroteknisk forskning* and *Civilingenjör Hakon Hanssons stiftelse.*

Contents

Part I

Overview of the Research Area

Chapter 1

Introduction

Radio transmission has allowed people to communicate without any physical connection for more than hundred years. When Marconi managed to demonstrate a technique for wireless telegraphy more than a century ago, it was a major breakthrough and the start of a completely new industry. Maybe one could not call it a mobile wireless system, but there was no wire! Today, the progress in semiconductor technology has made it possible, not to forget affordable, for millions of people to simultaneously communicate on the move all over the world. The first automatic mobile telephony system in Sweden was introduced in 1955 [14], offering voice services in three major cities. However, it was only after the introduction of the Nordic Mobile Telephone system (NMT) in 1981 that the number of users in Sweden exceeded 20 000. Today, the number of users is increasing dramatically and there is a movement towards more data communication and multimedia services. These services require new technology to handle the data rates and the amount of traffic expected.

The mobile communication systems are often categorized as different generations depending on the services offered. The *first generation* comprises the analog frequency-division multiplex access (FDMA) systems such as the NMT and AMPS (Advanced Mobile Phone Services). The *second generation* consists of the first mobile digital communication systems such as the time-division multiple access (TDMA) based GSM (Global System for Mobile communication), D-AMPS (Digital AMPS), PDC (Pacific Digital Cellular) and the code-division multiple access (CDMA) based system IS-95. These systems offer mainly speech communication, but also data communication limited to rather low transmission rates. The *third generation* is currently under development and the first prototype systems have just been delivered. This group is characterized by enhanced data communication capabilities providing data rates of at least 384 kbit/s over a wide area and contains the Wideband-CDMA system (W-CDMA) and the CDMA-2000 system.

In parallel to the development of the third generation systems, there has

been an increasing interest in high-speed wireless local area networks (W-LANs). These systems are the focus of this thesis, though the techniques presented can be applied to other systems as well. The W-LAN systems do not offer the same wide area coverage as the third generation mobile systems do, but within their limited coverage area they provide much higher data rates. Therefore, the W-LAN systems can be seen as a good complement to the mobile systems. Within the coverage area, for example the office, the wireless systems offer high bit rates with performance close to a wired connection. Outside the building, the wide-area systems offer a similar service at a lower data rate. However, there are of course applications only suitable for the high data rates. The W-LAN systems are, for example, represented by the IEEE 802.11a [3] and IEEE 802.11b [1] systems in the U.S., the Broadband Radio Access Network (BRAN) HiperLAN/2 system in Europe [4] and the MMAC system in Japan.

There is also another group of radio systems that has achieved large interest during the last years. These are the so-called ad-hoc networks, which for example include the Bluetooth system [25]. In these systems, there is no pre-determined access point, but the terminals establish connections between each other and every terminal can have the role as a master or "base station". This changes the conditions to establish and maintain the connection, since there is no dedicated base station to talk to. Nevertheless, the problems of synchronization and possibly channel estimation are still there. Also in these systems the receiver needs to find the start of a packet, the frequency offset, and so on, to be able to demodulate the data.

1.1 Digital Communication Systems

A digital communication system is often divided into several functional units as shown in Figure 1.1 [56]. The task of the source encoder is to represent the digital or analog information by bits in an efficient way. The bits are then fed to the channel encoder, which adds bits in a structured way to enable detection and correction of transmission errors. The bits from the encoder are grouped and transformed to certain symbols, or waveforms, by the modulator, and the waveforms are mixed with a carrier to get a signal suitable to be transmitted through the channel. At the receiver the reverse functions take place. The received symbols are demodulated and soft or hard values of the corresponding bits are passed to the decoder. The decoder analyzes the structure of the received bit pattern and tries to correct or detect errors. Finally, the corrected bits are fed to the source decoder that is used to reconstruct the analog speech signal or digital data input.

This thesis deals with the three blocks to the right in Figure 1.1: the modulator, the channel, and the demodulator. The main question is how to design certain parts of the modulator and demodulator to achieve efficient and robust transmission through a mobile wireless channel. A wireless channel has some

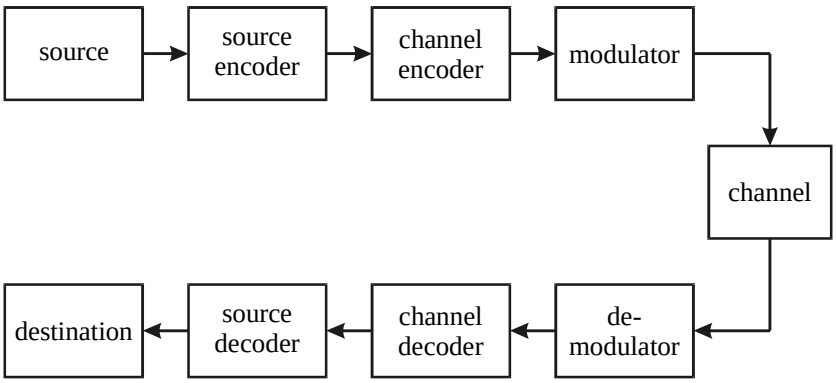

Figure 1.1: Functional blocks in a communication system.

properties that make the design especially challenging: it introduces time varying echoes and phase shifts as well as a time varying attenuation of the amplitude. The thesis focuses on the following parts in the modulator-demodulator chain:

- *OFDM system design.* Orthogonal frequency division multiplex (OFDM) has proven to be a modulation technique well suited for high bit rates on time dispersive channels [52]. There are, however, some specific requirements when designing wireless OFDM systems, for example how to choose the bandwidth of the sub-channels used for transmission and how to achieve reliable synchronization. The latter is especially important in packet-based systems since synchronization has to be achieved within a few symbols.

- *Channel estimation.* In order to achieve good performance the receiver has to know the impact of the channel. The problem is how to extract this information in an efficient way. Conventionally, known bits are multiplexed into the data sequence in order to estimate the channel. In OFDM systems, there are some specific opportunities for channel estimation since several adjacent carriers are used for the transmission as opposed to single-carrier systems.

- *Synchronization.* The receiver has to determine when there is a signal to demodulate and where the packets and symbols start in order to interpret the received signal. When there is no transmission, it is important that the receiver is able to enter some kind of sleep mode to save power. However, as soon as transmission starts the receiver has to achieve synchronization in a very short time in to prevent the loss of data.

As mentioned before, the transmitted information is represented by different waveforms that modulate a carrier of a certain frequency. The transmitted signal can be represented as [56]

$$s(t) = \mathrm{Re}\{x(t)e^{j2\pi f_c t}\}, \tag{1.1}$$

where $\mathrm{Re}\{\cdot\}$ denotes the real valued part, f_c is the carrier frequency and $x(t)$ is the low-pass equivalent, the so called complex envelope, of the transmitted signal. In digital communication systems the complex envelope is often described by a data part and a pulse shape. For pulse amplitude modulated (PAM) schemes or phase modulated, phase shift keying (PSK), schemes the complex envelope can be described by

$$x(t) = \sum_k c_k w(t - kT_s), \tag{1.2}$$

where $w(t)$ is the pulse shape, c_k is the complex-valued data part and T_s is the symbol time.

One of the challenging problems in the design of digital communication systems is how to choose the modulation scheme in order to get robust and efficient systems. This choice is greatly affected by the environment in which the system is supposed to work. The influence of the channel can be described by its impulse response and often, there is also additive white Gaussian noise (AWGN) representing different disturbances in the system. For mobile or wireless applications, the channel is often described as a set of independent multipath components. The time varying impulse response can be described by [58]

$$h(t, \tau) = \sum_{i=0}^{I-1} a_i(t)\delta(\tau - \tau_i(t)), \tag{1.3}$$

where $a_i(t)$ denotes the complex-valued tap gain for path number i, I is the number of taps, $\tau_i(t)$ is the delay of tap i, and δ is the Dirac delta function. Among the most important parameters when choosing the modulation scheme are the delay and the expected received power for different delays. Large delays for strong paths mean that the interference between the different received signal parts can be severe, especially when the symbol rate is high so that the delay exceeds several symbols. In that case one have to introduce an equalizer to mitigate the effect of intersymbol interference (ISI). Another alternative is to use many parallel channels so that the symbol time on each of the channels is long. This means that only a small part of the symbol is affected by ISI and this is the idea behind orthogonal frequency division multiplex, OFDM.

1.2 Wireless Systems

Wireless systems are operating in an environment which has some specific properties compared to fixed wireline systems and these call for special design con-

siderations. In a wired network, there are no fast movements of terminals or reflection points and the channel parameters are changing very slowly. In addition, time dispersion is less severe in a wired system, though it might still be a hard problem due to high data rates. In a mobile system the terminals are moving around; the received signal strength, as well as the phase of the received signal, are changing rapidly. Further, the signal transmitted over the radio channel is reflected by houses, cars etc., leading to different propagation paths to the receiver. If the length of the paths differ, the received signal will contain several delayed versions of the transmitted signal according to the channel impulse response (1.3). As mentioned before, the delays make it necessary to use complex receiver structures. In a mobile wireless system, the terminals are of course intended to be portable. This means that power consumption is important since batteries sometimes will power the equipment. Therefore, low complexity and low power consumption are properties that are even more desirable in a wireless system than in a wired system.

1.2.1 Packet-Based vs. Circuit-Switched Systems

A wireless system can either be packet-based or circuit-switched. In a packet-based system the information bits are grouped and transmitted in packets, and transmission occurs only when there is a need for communication. Therefore, in a pure packet-based system, no capacity is allocated in advance and these systems are suitable for bursty traffic conditions, such as data communication. In circuit-switched systems, a physical or virtual connection is established and occupied as long as communication proceeds. This means that capacity is reserved in advance, and that no other user can use this capacity as long as the connection is established. Circuit-switched systems are best suited for real-time services when delay is a limiting factor.

In packet-based systems, the receiver has to achieve synchronization in a very short time. It is hard to track channel variations between the packets, and therefore fast *acquisition* algorithms are required. After acquisition, initial estimates of the synchronization parameters are available and the estimator enters *tracking* mode. In this mode, the task is to follow changes and continuously update the channel estimates and synchronization parameters. Since the changes in general are relatively slow the requirements on the estimation time can be somewhat lower and therefore other algorithms can be used during tracking [27]. In a circuit-switched system, the receiver needs to enter acquisition mode more seldom due to the continuous transmission. Therefore the requirements on fast acquisition can be loosened in these systems. Today there is a trend towards more and more packet-based systems due to the increased data traffic. For example, both the third-generation mobile systems based on W-CDMA [35] and the HiperLAN/2 system [4] based on OFDM use packet-based communication for data traffic.

1.2.2 Coherent vs. Non-coherent Systems

In general coherent schemes result in better detection performance compared to differential systems, but they require channel estimation in order to form a reference for the decisions. Differential schemes, on the other hand, require no channel estimation, but there is a performance loss compared to coherent detection. Differential schemes also make it hard to apply multilevel modulation.

In coherent schemes the channel estimates are often achieved by multiplexing known, so called, pilot symbols into the data sequence. The technique is called Pilot Symbol Assisted Modulation (PSAM). It was originally proposed by Moher and Lodge in [49] and further analyzed by Cavers in [15]. The receiver observes the influence of the channel on the pilot symbols and uses interpolation to get an estimate of the channel impact on the data symbols. The receiver then removes this impact in order to make decisions. The pilot symbols transmit no data and therefore there is a small overhead causing a bandwidth expansion and an energy loss. Both these losses are dependent on the pilot-to-data symbol ratio. In multi-carrier systems interpolation in both time and frequency can be used and therefore the pilot-to-data symbol ratio can be decreased compared to single-carrier systems [A].

In conventional differential systems, the last received symbol is used as a reference when making the decisions. Since there is no absolute phase reference, differential encoding is required. In these systems, there is no need for channel estimation and therefore the complexity of the receiver can be decreased significantly compared to coherent schemes. There is, however, a performance loss compared to coherent systems since, in conventional differential systems, only one symbol is used to form the reference. This means that the reference symbol and the received data symbol are disturbed by noise samples both having the same variance. One way to get around this problem is to use multiple-symbol differential detection, introduced by Divsalar and Simon in [21]. In this scheme, the receiver uses more than the last received symbol to form a reference and is therefore less affected by noise. Naturally, this scheme is more complex than conventional differential detection.

1.3 OFDM

The first OFDM scheme with overlapping spectra was introduced by Chang in 1966 [17]. As opposed to today's systems, a bank of sinusoidal generators was used to produce the orthogonal channels. In 1971 Weinstein and Ebert [86] proposed to use the Discrete Fourier Transform (DFT) for this purpose and the first "modern" OFDM system was born. There exist many overviews and detailed presentations of OFDM systems in the literature, for example [7] [12] [22] [26] [52]. Here I will review the most fundamental properties and concepts needed for the rest of the thesis.

In OFDM systems many sub-channels are used in parallel. The channels are

overlapping in frequency, but the distance between them is chosen so that the different channels anyhow are orthogonal. In its most simple form the sample at time k of the complex envelope is given by [7]

$$x_k = \sqrt{\frac{E_s}{N}} \sum_{n=0}^{N-1} c_n e^{j2\pi n \frac{k}{N}}, \tag{1.4}$$

where N is the size of the DFT and E_s is the symbol energy. The size of the DFT is equal to the number of sub-channels available for transmission, but all of the channels need not be active. The sub-channel bandwidth is given by

$$f_{sc} = \frac{1}{T_s} = \frac{f_{samp}}{N}, \tag{1.5}$$

where f_{samp} is the sample rate and T_s is the symbol time. Normally, a cyclic extension of the symbol [54] is used to mitigate the effect of ISI. This, so called, cyclic prefix is a copy of the last G samples and these are transmitted prior to the actual symbol. Alternatively, a prefix of length G_1 and a postfix of length G_2 can be used. The main idea behind OFDM is to use long symbols, such that the part affected by ISI (the prefix) can be disregarded with only a minor energy loss. The signal is often also windowed in order to improve the performance. This is especially important for reducing out-of-band emissions and to decrease the interchannel interference, ICI, due to frequency offsets [7]. Different approaches for the windowing functions are proposed in the literature. In, for example, the IEEE 802.11a system [3] a smoothly decaying extension of the symbols can be used, while in [7] a multiplication with a raised cosine function in time is proposed. The complex envelope including both pulse shaping/windowing and the pre/postfix can be written as

$$x(t) = \sqrt{\frac{E_s}{N+G}} \sum_{l=-\infty}^{\infty} \sum_{k=-G_1}^{N+G_2} \sum_{n=n}^{N-1} c_{k,l} e^{j2\pi n \frac{k}{N}} \cdot w(t - \frac{k}{f_s} - lT_{tot}), \tag{1.6}$$

where $c_{k,l}$ denotes the data symbol on the kth sub-channel during the lth OFDM symbol, $w(t)$ denotes the windowing function, and $T_{tot} = (N+G)/f_s$ denotes the total OFDM symbol time including the cyclic pre/postfix.

An often used model for wireless systems is that the transmitted signal is convolved with the channel impulse response. Then, noise and interference are added to this signal. In an OFDM receiver, a DFT is performed in order to detect the data symbols on each of the sub-channels. If we assume ideal synchronization, a cyclic extension exceeding the maximum excess delay (delay between the first and last tap of the channel) and no frequency offset then the received symbol after the DFT is given by [7]

$$r_k = H_k c_k + n_k \tag{1.7}$$

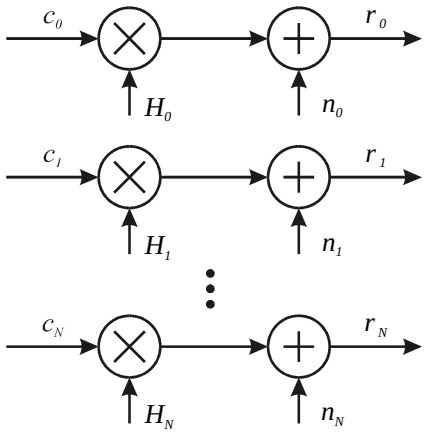

Figure 1.2: The data symbols are multiplied by the values of the transfer function at the sub-channel frequencies and disturbed by noise.

for a slowly varying time dispersive channel with additive Gaussian noise. Here n_k denotes white Gaussian noise and H_k denotes the complex value of the transfer function at sub-channel k. In the ideal case, the symbol is only disturbed by noise and phase shifted and attenuated by the frequency response, according to Figure 1.2, after demodulation and transmission over a multipath channel. This is in contrast to single-carrier systems, where the received signal generally is corrupted by ISI.

A block description of a conventional OFDM modulator and demodulator is presented in Figure 1.3. First the coded bits are fed to the modulator where groups of $\log_2 M$ bits are mapped to M-ary symbols, where M is the size of the symbol alphabet on each carrier. Then N of these symbols are grouped into one OFDM symbol (note, however, that some of the outer channels may be zero-valued). Each data symbol is assigned to a specific sub-channel by the IDFT, the inverse discrete Fourier transform. Normally this latter operation is performed as an inverse fast Fourier transform (IFFT) due to its lower complexity. Then the cyclic extension is added and finally pulse shaping is performed before the signal is up-converted and transmitted over the channel.

At the receiver, the received signal is down-converted and sampled. Then the samples are fed in parallel to the channel estimator, the FFT and the automatic gain control/coarse synchronizer, respectively. The automatic gain control (AGC) tracks the incoming power and adjusts the level so that the full dynamic range of the analog-to-digital converter is used. The task of the synchronizer, or more precise the coarse synchronizer, is to find the start of each data packet and each OFDM symbol, and to find an estimate of the frequency offset so that compensation can be made before the samples are fed to the FFT

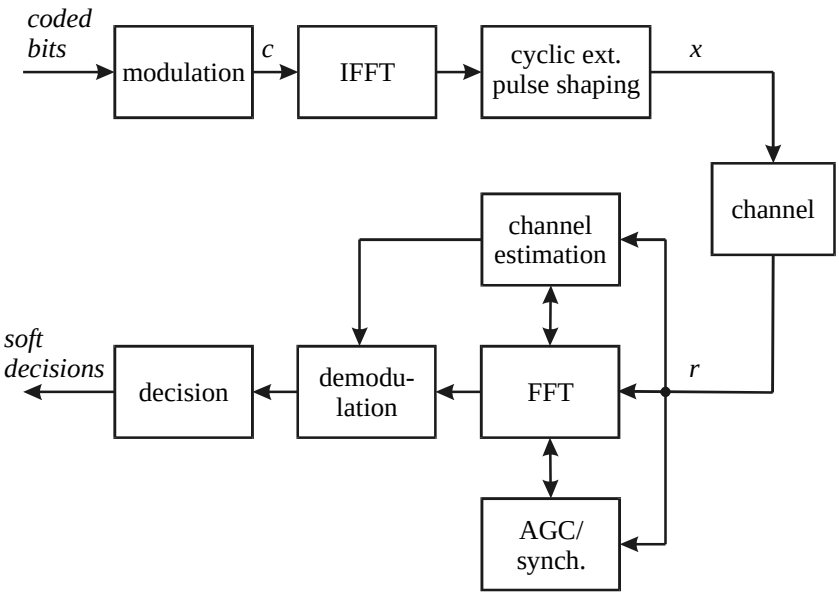

Figure 1.3: Block description of a conventional OFDM modulator and demodulation chain

block. The OFDM demodulator is very sensitive to frequency offsets [64] and in order to avoid ICI frequency offset compensation has to be made before the DFT.

The channel estimator is used to enable coherent modulation schemes. The time dispersive channel results in a phase shift and an amplitude attenuation according to (1.7) and therefore the values of the transfer function at the different sub-channel frequencies have to be estimated in order to use coherent detection. In addition, if there is a small timing offset, this will result in linearly increasing phase between the sub-channels. The channel estimator can estimate and compensate for these phase shifts since they can be regarded as a part of the channel transfer function [26].

The FFT block performs a DFT on the received OFDM symbol in order to retrieve the data symbols. These symbols are then passed to the demodulation stage where the influence of the channel is removed so that hard or soft decisions can be made. These decisions are then passed to the channel decoder as in Figure 1.1.

1.3.1 Single-Carrier vs. Multi-Carrier Systems

The main difference between a single-carrier and a multi-carrier system is how the influence of the channel is mitigated. In an OFDM system the receiver has

to perform an FFT in order to demodulate the data, but the equalizer can on the other hand be really simple since a time dispersive channel only results in a multiplication by the value of the channel transfer function after the FFT. In a single-carrier receiver there is no FFT block, but the equalizer must be designed to mitigate the ISI instead. The complexity of the equalizer can be quite high if the ISI spans over several symbols, but if it spans over only a few symbols the solution is straightforward. An estimate of the complexity of the equalizer is found by comparing the maximum excess delay of the channel to the symbol time; if the ISI spans over several symbols, then OFDM is to prefer to the use of long (and thereby complex) equalizers [7].

Another advantage of multi-carrier modulation is that it is possible to use adaptive modulation on a sub-channel basis. This means that different modulation levels can be used on different sub-channels. On a "bad" sub-channel where the channel attenuation is high, a low-level modulation scheme, such as binary PSK, can be used and on a good sub-channel a higher order modulation, such as 16-QAM, can be used. This means that the channel is more efficiently used and that the bit error rate (BER), can be decreased significantly compared to fixed modulation schemes [26]. Note however, that adaptive schemes require both extra hardware and control signaling.

The main disadvantages of using multi-carrier systems, or OFDM systems in particular, are the sensitivity to frequency offsets and the large dynamic range of the signal to be transmitted [7]. Frequency offsets cause both ICI and attenuation of the desired signal. The sensitivity to frequency offsets can be explained by the large frequency side lobes of the signal and the sensitivity can be decreased by using a window function with a more attractive frequency characteristics compared to the one resulting from the rectangular time domain window [7]. The large dynamic range can be understood by looking at the OFDM signal as a sum of N sinusoids. Sometimes the waveforms add constructively, which means that a peak is produced. For most of the time, however, the waveforms have different signs such that the sum achieves low amplitude.

Chapter 2

Research Topics

In this chapter a detailed background and some of the main publications in each of the research areas are presented. Also, the relation between the papers presented in this thesis and other papers is discussed. As mentioned in the introduction the work has been focused on system aspects of the modulator-demodulator chain. In principle I have been active in four different research projects, and these are the areas that I will present in more detail in this chapter. The four projects are: channel estimation in OFDM systems, wireless OFDM system design, preamble-based synchronization for OFDM, and channel estimation and synchronization using superimposed pilot sequences. Papers A and C belong to the first group on channel estimation and paper D belongs to preamble-based synchronization. Paper B belongs to OFDM system design, but of course the papers regarding channel estimation, pre-compensation and preamble-based synchronization can also be assigned to this group. Finally, papers E, F and G belong to the project on superimposed pilot sequences. Next, there is a presentation of channel estimation in general since this is a fundamental topic for the whole thesis. After that, details of the different research areas are presented.

2.1 Channel Estimation

The task of the equalizer is to compensate for the influence of the channel. This compensation requires, however, that an estimate of the channel response is available. There are many alternatives identifying the channel response and depending on the channel statistics and the modulation scheme different solutions are preferable. Often the channel impulse response or the frequency response is derived from training sequences or pilot symbols, but it is also possible to use non-pilot-aided approaches, so called blind algorithms [48]. The use of training symbols means that the complexity of the receiver can be kept rather low and, in case that the training symbols and data symbols are re-

ceived without influence from each other, the processes of channel estimation and detection can be separated. Also, training symbols are suitable for packet-based communication since these result in robust behavior with no significant transient effects [48].

In the general case the optimal receiver, in the sense of minimum sequence error probability when no prior channel estimates are available, is the maximum-likelihood (ML) receiver making joint detection and channel estimation [48]. The task is to find a channel estimate and a data sequence that make the actual received sequence as probable as possible, i.e. to find

$$(\hat{\mathbf{c}}, \hat{\mathbf{h}})_{ML} = \arg\max_{\mathbf{c},\mathbf{h}} f(\mathbf{r}|\mathbf{c}, \mathbf{h}) \tag{2.1}$$

where $\mathbf{r}$ is the received sequence, $\mathbf{c}$ is the data sequence, $\mathbf{h}$ is the sequence of the channel responses, $f(\cdot)$ denotes the probability density function and $\arg\max$ denotes the argument maximizing the function. This approach leads, however, to complex receiver structures and therefore other less complex alternatives are of interest. In case channel estimation and data detection can be performed separately, the structure can be simplified by deriving the channel estimates before detection takes place. Given a certain channel estimate $\hat{\mathbf{h}}$, the optimal detector, in the sense of minimal sequence error probability, is the Viterbi detector [56]. This equalizer gives the maximum-likelihood sequence, which means that it makes a decision in favor of the sequence that maximizes the probability of receiving the actual received sequence, i.e.

$$\hat{\mathbf{c}} = \arg\max_{\mathbf{c}} f(\mathbf{r}|\mathbf{c}, \hat{\mathbf{h}}). \tag{2.2}$$

In OFDM systems there is no ISI after demodulation (under ideal conditions) and, as not always the case for single-carrier systems, channel estimation and detection can be performed separately in a time dispersive channel. Another advantage is that a very simple equalizer structure can be used, since only a complex division by the value of the channel transfer function is required. A necessity for this is that the receiver knows, or at least makes estimates of, the transfer functions at all the sub-channel frequencies.

One technique suitable in OFDM systems for estimating the values of the transfer function is to use pilot symbol assisted modulation (PSAM) operating in the frequency domain. Other techniques for channel estimation and calculation of the transfer functions in OFDM systems include, for example, the use of correlation-based estimators working in the time domain [18] and channel estimation using singular value decomposition [23] [44]. The latter can be based on pilot symbols, but to calculate the channel estimates the statistical properties of the channel are used in a different way in order to decrease the complexity.

2.1.1 Pilot Symbol Assisted Modulation

PSAM was, as mentioned earlier, introduced by Moher and Lodge in [49], for the use in a trellis coded modulated (TCM) scheme. The channel estimates were derived using a low-pass filter, and the estimates were used both for detection of the PSK modulated data symbols and as estimates of the received symbol energy in the decoder. Aghamohammadi *et al.* independently proposed a similar pilot-based scheme [5] intended for uncoded PSK and QAM schemes. They used a Kalman smoother to derive the channel estimates and they also came to the conclusion that pilot-based techniques lead to robust detection with a performance better than that of non-coherent schemes. Both in [5] and [49] it was concluded that the maximum spacing between the pilot symbols is determined by the Doppler frequency since the pilot technique can be seen as a sampling of the low-pass fading process. Both these papers relied on simulations for their analyses, but Cavers made a thorough analytical analysis of the technique in [15]. Here, he used an optimal Wiener filter for the channel estimates and he derived the resulting bit error rate both under ideal and mismatched conditions. One of the main conclusions was that PSAM outperforms differential detection for all values of the signal-to-noise ratio (SNR) and Doppler frequency. Another important conclusion was that if the channel estimator is designed for the worst-case Doppler frequency, then the performance will in general not be worse for other lower Doppler frequencies. This means that a worst-case design is appropriate for the channel estimation filter. Cavers also studied the impact of the pilot spacing and stated that in order to fulfill the sampling theorem the spacing in time between the pilot symbols, N_T, should not exceed [15]

$$N_T \leq \frac{1}{2f_D T_s}, \tag{2.3}$$

where f_D is the maximum Doppler frequency. This means that the maximum pilot distance is proportional to the coherence time, normalized to the number of symbols.

PSAM for OFDM systems was proposed by Hoeher in [30]. The two-dimensional channel estimation scheme used in OFDM systems can be seen as a generalization of the one-dimensional PSAM scheme working only in the time direction. In OFDM systems a two-dimensional estimation algorithm can be employed since the channel transfer functions are correlated between adjacent symbols both in the time and in the frequency direction. This means that the overall pilot-to-data symbol ratio can be decreased compared to single-carrier systems [30].

The optimal linear channel estimator, in the sense of minimizing the mean square error, is the two-dimensional Wiener filter. This filter is based on knowledge of the time/frequency correlation function and weighs the received pilot values with respect to their correlation and noise properties in order to derive the channel estimates. For the pilot position at position P_i a tentative channel

estimate, $\tilde{H}(P_i)$, for the transfer function can be obtained by back-rotating the received samples as

$$\tilde{H}(P_i) = \frac{r_{P_i}}{p_{P_i}}, \tag{2.4}$$

where $\mathbf{P} = [P_0, P_1, \ldots, P_{L-1}]$ is the set of pilot symbol positions used for the estimate, p_{P_i} is the pilot symbol at position P_i and r_{P_i} is the corresponding received value after the DFT. The channel estimate, $\hat{H}_{k,l}$, of sub-channel k at time l can then be computed by linear filtering

$$\hat{H}_{k,l} = \sum_{i=0}^{L-1} w_{k,l}(P_i) \cdot \tilde{H}(P_i), \tag{2.5}$$

where L is the number of pilot symbols used for the estimate and $w_{k,l}(P_i)$ is the channel estimation coefficient using the pilot symbol at position P_i for the data symbol at position (k, l). If the channel is wide-sense stationary, the optimal $L \times 1$ vector of filter coefficients is the solution to the Wiener-Hopf equation [31]

$$\mathbf{w}_{k,l}(\mathbf{P}) = \Phi^{-1}\boldsymbol{\theta}_{k,l}, \tag{2.6}$$

where $\boldsymbol{\theta}_{k,l}$ denotes the $L \times 1$ cross-correlation vector with elements $\theta_{k,i}(i) = E[H_{k,l} \cdot \tilde{H}^*(P_i)]$ and Φ denotes the $L \times L$ autocorrelation matrix with elements $\Phi_{i,j} = E[\tilde{H}(P_i) \cdot \tilde{H}^*(P_j)]$. The mean square error (MSE) can be used as a performance measure of the estimation filter, but can also be used to calculate the resulting BER. The MSE under ideal conditions (known correlation matrices) is given by [31]

$$J_{k,l} = \sigma_H^2 - \boldsymbol{\theta}_{k,l}^T \Phi^{-1} \boldsymbol{\theta}_{k,l}^*, \tag{2.7}$$

where σ_H^2 is the variance of the complex components of the channel. Finally, after tedious derivations, the BER for Q-PSK can be calculated as [56]

$$P_{k,l}(error) = \frac{1}{2} - \frac{\mu_{k,l}}{2\sqrt{2 - \mu_{k,l}^2}}, \tag{2.8}$$

where

$$\mu_{k,l} = E_s \frac{\sigma_H^2}{\sqrt{(\sigma_H^2 + \sigma_n^2)(\sigma_H^2 + J_{k,l})}} \tag{2.9}$$

and σ_n^2 denotes the variance of the noise.

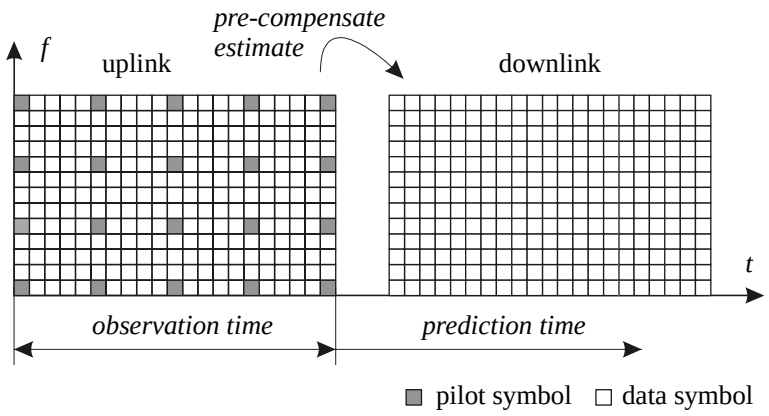

Figure 2.1: The principle of pre-compensation showing two OFDM blocks.

2.1.2 Pre-compensation

Pre-compensation, sometimes called precoding, is a technique where the signal is compensated before it is transmitted in order to improve the performance in some sense. The technique can also be used to move complex functions from the receiver to the transmitter, thereby enabling less complex terminal structures with performance similar to more complex ones. Compensation can, for example, be used to mitigate ISI in single-carrier systems to avoid the need for equalization at the receiver [45] [89] [90] or for antenna diversity using only one receive antenna [11] [28] [51] [68]. The technique can also be used to compensate the fading channel in order to improve the performance [C] [47] or to maintain orthogonality between the different user signals in multi-carrier spread-spectrum (MC-SS) systems [38] [85] [88]. Due to the simplicity and robustness a variant for antenna diversity is used in the DECT system [2] where the antenna receiving the most power during uplink is used for downlink transmission. Also in the UMTS/W-CDMA system a simple phase compensation system is used in time division duplex (TDD) mode to align the phases of different paths [35].

The technique relies on that the channel characteristics are reciprocal and constant during one receive-transmit-cycle. Therefore pre-compensation is suitable when TDD is used and the time between receiving mode and transmission mode is short compared to the coherence time of the channel. The channel characteristics are identified in one direction and then compensation is made prior to transmission in the other direction, as described in Figure 2.1. Since compensation is made in advance, the maximum block lengths are determined by the time-correlation function of the channel. The shape of this correlation function is, among other things, determined by the Doppler frequency. High

Doppler frequencies naturally call for short blocks and therefore the technique is most suitable for systems operating at pedestrian speeds.

When fully compensating the phase and amplitude of a flat-fading channel the complex envelope of the transmitted signal is given by

$$s(t) = \frac{x(t)}{\left|\hat{H}(t)\right|} e^{-\arg\{\hat{H}(t)\}}, \tag{2.10}$$

where $x(t)$ is the complex envelope of the uncompensated signal and $\hat{H}(t)$ is the channel estimate. If the channel is Rayleigh fading the signal fluctuations become severe and therefore the amplitude has to be limited, either by limiting the output power or by inhibiting the transmitter when the channel attenuation is large [C] [47]. In a Rayleigh fading channel the fading dips cause most of the bit errors. If it is possible to compensate for these there will, for uncoded systems, be a large improvement in performance. In [47] it was shown that it is possible to decrease the mean transmitter power by 7 dB (for a fixed BER of 10^{-3}) when introducing power compensation in a non-coherent FSK system. In [C] it is shown that the BER can be decreased 20 times compared to an uncompensated system for coherent QPSK at $E_b/N_0 = 10$ dB. This, however, results in an increase of the peak-to-average power ratio by 10 dB so this strategy might be of limited interest for wireless applications.

Power compensation has also been proposed to maintain orthogonality between users in MC-SS systems [85]. In an MC-SS system, or more specifically in an MC-CDMA system, the data symbol from one user is spread to many sub-channels and then data symbols from several users are transmitted on the same sub-channels simultaneously [39]. Since each data symbol is transmitted over several sub-channels, a frequency diversity gain can be achieved when adding the components from the different sub-channels. If coherent detection is used then equalization is required in the same way as in conventional OFDM systems. However, the orthogonality between users can not be preserved in a multiuser system when equalizing the uplink signal in the receiver. Therefore multiuser detection or pre-compensation is required in order to achieve good performance [85]. The question for pre-compensation in this application is whether one can afford the resulting increase in peak-to-average power ratio reported in [C].

Pre-compensation can also be used for antenna diversity. The technique has been used to choose transmit antenna in the DECT system [2], to achieve constructive addition from two antennas [11] [28] or from an array of antennas [51]. In [11] a diversity gain of 6-10 dB was reported (at BER 10^{-3}) for a system using the antenna diversity arrangement and an equalizer (to mitigate ISI), compared to a system using only the equalizer. This gain is of the same order as the 10 dB gain that can be achieved (for QPSK at BER 10^{-3}) using two receive antennas in a Rayleigh fading channel [56]. Antenna pre-compensation is therefore an efficient technique to get close to optimal antenna diversity using

only one receive antenna. In [68] the switched antenna diversity system was generalized to OFDM systems. In this system, some of the sub-channels were transmitted from one antenna and the remaining sub-channels were transmitted from the other antenna. This system also resulted in a close to optimal diversity gain, assuming independent sub-channels and ideal synchronization and channel estimation. The question is, however, how to derive the necessary channel estimates in an efficient way. It is possible to use some kind of pilot-based channel estimation on a per sub-channel basis, but it is hard to use the correlation between the sub-channels for estimation in the receiver since different antennas are used. Alternatively, phase-only pre-compensation [C] can be used in order to align the phases and avoid channel estimation at the receiver.

A general problem when using pre-compensation in any form is that the channel needs to be constant, or almost constant, during each receive-transmit-cycle. In [C] the maximum block length is calculated where the instantaneous BER gets worse than that of differential detection. It is shown that for phase-only pre-compensation a block length of $f_D T_{block} \approx 0.26$ can be used before the performance deteriorates. To compare this number to an existing TDD system this corresponds to a block length 40 times longer than the one in the DECT system at 1900 MHz and a velocity of 10 km/h. The conclusion also holds when a worst-case design is made for the channel estimators [C]. Similar block lengths can probably be expected for the other pre-compensated systems, but this is of course dependent on the sensitivity of the receiver to different kinds of errors. The BER is dependent on the ability of making channel predictions. If only the last estimates in the uplink block are used for compensation of the downlink block, the performance is degraded compared to the use of optimal prediction. For the phase-only compensation system in [C] the maximum block length is approximately $f_D T_{block} \approx 0.05$ when the last estimate of the uplink block is used for phase compensation. In [38], where the power is pre-compensated in an MC-CDMA system, the maximum block length for a two-fold degradation in the BER is $f_D T_{block} \approx 0.1$. Both these numbers can, for example, be compared to the average duration of a 10 dB fading dip which is $f_D \bar{T}_{dip} \approx 0.13$ for a Rayleigh fading channel with isotropic scattering [87].

2.2 Multi-Carrier System Design

In this section the background of the research areas related to OFDM system design is presented. The topics presented cover optimization of the sub-channel bandwidth, pilot pattern design and preamble-based synchronization. System design of the physical layer includes more topics than presented here, but these are out of scope for the thesis. Other topics include, for example, choice of the modulation format, design of the coding scheme and multiple access scheme.

2.2.1 Sub-Channel Bandwidth

When choosing a suitable sub-channel bandwidth one has to make a compromise between the overhead caused by the cyclic prefix and impairments introduced by the hardware and the transmission channel. The cyclic prefix is used to mitigate the effect of time dispersion and its duration has to exceed the maximum excess delay of the channel in order to avoid ISI. Since the cyclic prefix is a repetition of the samples, the sample rate has to be increased if the data rate is to be constant after introducing the cyclic prefix. Therefore, this leads to an increase in bandwidth. In addition, the receiver uses only samples corresponding to the actual symbol for demodulation, the samples used for the cyclic prefix are disregarded, and therefore there is an energy loss. This energy loss is more evident when the symbol time is short since the relative duration of the cyclic prefix then is longer. Therefore, the cyclic prefix calls for a long symbol time, with its corresponding narrowband sub-channels. The symbol length can not, however, be arbitrarily long because in order to preserve orthogonal sub-channels the impact of the channel on the transmitted signal has to be constant during each OFDM symbol. This means that the channel is not allowed to change, there must be no oscillator offsets and no oscillator phase noise, the sampling time has to be correct and all hardware components must be ideal. Of course, this is not possible in practice and therefore one has to allow a certain level of ICI.

The ICI arising from channel or hardware impairments can often be described as extra additive Gaussian noise and the contributions from different interference sources can be added to get the overall impact on the system [64]. The level of the interference is, among other things, dependent on the Doppler frequency and frequency offset in relation to the sub-channel bandwidth. Since the sub-channel bandwidth is inversely proportional to the number of sub-channels used, the interference level is therefore also dependent on this number. In most cases, the ICI increases as the sub-channel bandwidth is chosen to be smaller and therefore it is important to identify the dominant sources of interference and choose a sub-channel bandwidth such that the ICI does not become the limiting factor in the system performance. Good overviews of the effects contributing to the ICI, including expressions of the resulting ICI levels, can be found in [64] and [66]. There are also many other papers in the literature specialized in one or a few sources of interference. In [66] the power at the kth output of the DFT is decomposed into four groups

$$P[k] = \frac{E_s}{T_s}\frac{N}{N+G_1+G_2}(P_U[k] + P_{ICI}[k] + P_{ISI}[k]) + \frac{N_0}{T_s}, \tag{2.11}$$

where $P_U[k]$ denotes the (normalized) useful power from the symbol to be detected, $P_{ICI}[k]$ denotes the interfering power from other sub-channels, $P_{ISI}[k]$ denotes contributions from other OFDM symbols and N_0 denotes the power spectral density of the noise. It can be shown [66] that the sum of the useful power and the interfering power is constant, $P_U[k]+P_{ICI}[k]+P_{ISI}[k] = 1$, so if

the ICI increases the useful power automatically decreases by the same amount. In the following, we treat ICI in more detail, since this is a very important issue in OFDM system design.

An important effect to consider in the system design is interference due to channel changes during an OFDM symbol. The area has, for example, been studied in [43] [61]. This interference is caused by movements of either the transmitter, the receiver or reflection points and it is hard, if not impossible, to compensate for [61]. Transmitter or receiver movements result in the same kind of interference as oscillator offsets [64], but in a wireless environment they cause a Doppler spread rather than a frequency shift and therefore they are much harder to compensate. The power of the ICI due to the time-selective fading is proportional to the squared normalized Doppler frequency $(f_D T_s)^2$, and therefore it is inversely proportional to the square of the sub-channel bandwidth f_{sc}^2. The variance of the equivalent noise caused by the time-selective channel can be calculated as [64]

$$\sigma_{ts}^2 = \frac{\pi^2}{6}(f_D T_s)^2, \tag{2.12}$$

when the received signal power is normalized to 1.

In [B], the sub-channel bandwidth is optimized with respect to the physical phenomena described above, the energy loss and the ICI due to the time-selective channel. The optimal sub-channel bandwidth is dependent on the signal-to-noise ratio since there is a balance between the ICI and the noise; therefore it is hard to give an exact expression for a suitable sub-channel bandwidth. However, for mobile applications where $E_b/N_0 = 10 - 25$ dB a normalized Doppler frequency of $f_D T_s \approx 0.01 - 0.02$ seems suitable [B] [66]. This corresponds to a degradation in SNR of less than 1 dB [64] and results in a signal-to-interference ratio exceeding 30 dB [43].

ICI can also be caused by frequency offsets and for consumer products, where low-cost implementations are important, this might be a significant source of interference. The interfering noise level can be approximated by [64]

$$\sigma_{fo}^2 \approx \frac{\pi^2}{3}(\Delta f T_s)^2, \tag{2.13}$$

where Δf denotes the frequency offset. As seen in (2.13), the variance is again proportional to the square of the normalized frequency offset, $\varepsilon_c^2 = (\Delta f T_s)^2$, and therefore inversely proportional to the square of the sub-channel bandwidth f_{sc}^2. Normally, frequency offset compensation is made prior to the DFT and the interesting parameter is therefore the remaining frequency offset after compensation. The expression for the variance of the equivalent noise caused by frequency offset is similar to the noise caused by the time-selective channel. Therefore, for the same level of interference the maximum frequency offset is

of the same order as the Doppler frequency, typically $\varepsilon_c \approx 0.01 - 0.02$, but this should be seen only as a rough guideline.

Other effects that need consideration are the influence of phase noise of the oscillator and the effect of sampling errors. Phase noise has been studied in, for example, [6] [55] [60] and it leads to two kinds of distortion. First, there is an ICI term, whose variance increases with the number of sub-carriers. It is hard to get numerical expressions for the variance since this is dependent on the model of the phase noise including possible feedback in the circuit. For Wiener phase noise, the corresponding SNR loss due to the ICI is proportional to the oscillator linewidth and inversely proportional to the sub-channel bandwidth [55]. The phase noise also leads to a phase error common to all sub-channels within an OFDM symbol. The variance of this distortion decreases as the sub-channel bandwidth decreases [60]. The correlation of the common phase error is in general low between successive OFDM symbols, so compensation on a symbol-by-symbol basis is required if the variations are significant.

Sampling offsets lead to a phase rotation of the symbols. Further, if some paths of the channel fall outside the guard interval, ICI and ISI occur. The resulting interference can be approximated as Gaussian noise with variance [64]

$$\sigma_{to}^2 \approx \sum_i |a_i(t)|^2 \left(2\frac{\varepsilon_i}{N} - \left\{ \frac{\varepsilon_i}{N} \right\}^2 \right), \tag{2.14}$$

where ε_i is the normalized "effective" time offset defined as

$$\varepsilon_i = \begin{cases} \frac{\Delta t - \tau_i}{T_{samp}} & \Delta t > \tau_i \\ \frac{\tau_i - GT_{samp} - \Delta t}{T_{samp}} & 0 < \Delta t < \tau_i - GT_{samp} \\ 0 & \text{else} \end{cases} \tag{2.15}$$

through the time offset Δt and the sample time $T_{samp} = 1/f_{samp}$. As before, N denotes the number of sub-channels, G denotes the length of the cyclic prefix, τ_i denotes the delay of path i, and $a_i(t)$ denotes the channel tap value of path i. The variance of the ICI and ISI due to sampling offsets is highly dependent on the power delay profile of the channel [64] and the disturbance is in principle proportional to the sub-channel bandwidth [66]. If many sub-channels are used, the relative importance of samples corresponding to paths falling outside the cyclic prefix becomes smaller and therefore the overall distortion decreases.

Finally some remarks about the relation between the sub-channel bandwidth and the length of packets in wireless systems. The number of sub-channels used determines the minimum number of symbols in a packet and therefore also the granularity when choosing the packet lengths. For large data packets this might not be a problem, but if there are many short packets, or there are low-rate connections with constraints on the delay, this effect has to be considered when determining the sub-channel bandwidth.

2.2.2 Preamble-Based Synchronization

There are many synchronization parameters to be estimated before symbol decisions can be made. The demodulator has to know where the start of a packet is, where the OFDM symbols start, how large the frequency offset is etc. In packet-based systems it is desirable to minimize the acquisition time and therefore pilot-aided approaches for synchronization are suitable, though these lead to a throughput degradation. Often a preamble is used, which is transmitted before each new packet. Such an approach is used in the HiperLAN/2 system [4].

The influences of time and frequency offsets have been described in Section 2.2.1. The variance of the interference caused by frequency offsets is given in (2.13) and it is evident that the OFDM demodulator is very sensitive to frequency offsets. In order to achieve a tolerable degradation in system performance the remaining frequency offset must not exceed a few percent of the sub-channel bandwidth. The requirements on the timing estimator are somewhat lower due to the robustness introduced by the cyclic prefix. If all delayed paths of the channel fall into the cyclic prefix, orthogonality is preserved and neither ISI nor ICI arise due to the time offset; otherwise the level of interference is given by (2.14).

Synchronization in OFDM systems is often based on the transmission of a repeated symbol. The synchronization parameters are then derived by correlating currently received samples with earlier samples. Time synchronization is achieved when the result of the correlation exceeds a certain threshold, which means that the current samples correspond to the earlier received ones. When the correct start of the frame has been found the frequency offset is estimated by the phase shift between the first and second part of the repeated symbol. Normally correlation is performed using the time domain samples [63], but it can also be performed using the signal after the FFT [50]. In the latter case, there is a performance loss for large frequency offsets, but the method is still possible to use.

The specific structure of the preamble varies, but the principle is the same for all these methods. In [50] and [62] a repeated full-length OFDM symbol is used. In [46] and [63] the repetition is accomplished within one symbol by only modulating every nth sub-channel. The shorter time before repeating the signal means an increased frequency offset capture range of the synchronizer at the cost of a corresponding increase in variance of the estimate. Other approaches use the repetition due to the cyclic prefix [37] [8]. If the existing cyclic prefix is used, then there is no overhead introduced and this is of course desirable. In some cases, the cyclic prefix is too short to give sufficient estimation accuracy, and then averaging over several symbols is required, with a corresponding delay. In [41] a special synchronization sequence, a CAZAC-sequence (constant-amplitude zero autocorrelation sequence), is inserted on each side of an OFDM symbol for synchronization purposes. The sequence is

repeated both between the symbols and within the block, which means that coarse synchronization can be performed within the block and fine frequency synchronization can be performed by using adjacent blocks. Finally, in [D] a repeated m-sequence (maximal-length sequence) is used. This approach is motivated by the attractive correlation properties for these sequences.

There are also methods based on principles other than repeated symbols. In [42] the pilot symbols intended for channel estimation are used together with the cyclic prefix to derive an improved time synchronization signal. In [13] the preamble is generated using only the highest or lowest half of the subchannels. In the system, time synchronization is achieved by comparing the spectral content in the two halves of the normal transmission bandwidth. In [65] a special low-rate signaling channel using frequency shift keying (FSK) is used for synchronization purposes. This approach was motivated by complexity arguments since then a simple FSK receiver can be used during sleep mode. In [59] continuous pilot carriers were used to derive a frequency offset estimate. Another approach is based on short power variations in the transmitted signal [62]. By using so called null symbols the receiver only has to look for certain variations in the received power to find a coarse time estimate, which can be performed using a low-complexity receiver without performing an FFT.

The main difference between the approach in [D] and the other correlation-based methods is that in [D] correlation with the known sequence is made prior to multiplication with the second part of the repeated sequence when calculating the timing synchronization. This synchronizer is called a pre-correlating synchronizer. The synchronization signal can for both synchronizers be calculated as

$$\gamma[k,a] = \sum_{l=0}^{L-1} \left[\left(\sum_{n=0}^{K-1} p^*[k-n-lR-a] r[k-n-lR] \right) \cdot \left(\sum_{n=0}^{K-1} p^*[k-n-(l+Q)R-a] r[k-n-(l+Q)R] \right)^* \right], \tag{2.16}$$

where a is the lag between the known synchronization sequence and the received sequence, L is the number of products used for the synchronization signal, $p[k]$ is the known sequence, Q is an extra delay which can be used to improve the performance of the frequency offset estimate, K is the correlator length and R is the repetition length of the sequence. The only difference between the conventional correlation-based synchronizer and the synchronizer in [D] is how these parameters are chosen. The conventional synchronizer is shown in Figure 2.2 and this is obtained by choosing the correlation length to $K = 1$, the number of products used for the estimate to $L = R$, the known sequence to $p[k] = 1$ and $Q = 1$. The pre-correlating synchronizer is shown in Figure 2.3. The synchronization signal is obtained by choosing the number of products to $L = 1$, the correlation length to the repetition length $K = R$ and $Q = 1$. The pre-correlation with the known sequence means that the noise is

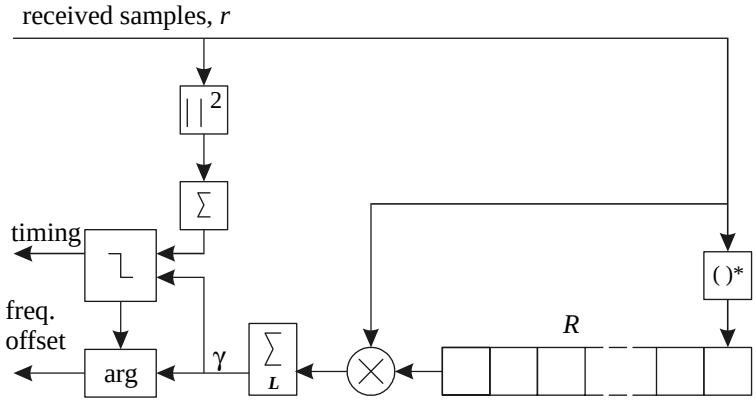

Figure 2.2: Structure of the conventional synchronizer. This corresponds to $K = 1$, $Q = 1$ and $L = R$ in (2.16).

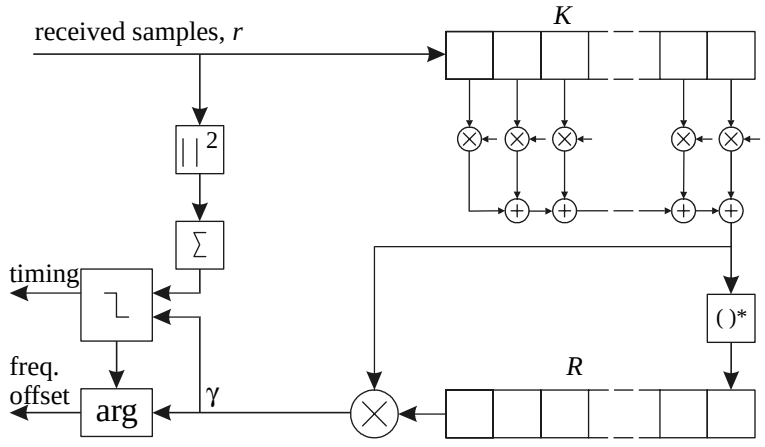

Figure 2.3: Structure of the pre-correlating synchronizer. This corresponds to $K = R$, $Q = 1$ and $L = 1$ in (2.16).

averaged before multiplication is performed. Therefore, the mean amplitude of the synchronization signal when the timing is wrong is decreased when using this synchronizer instead of the conventional synchronizer, which in turn results in a performance enhancement. In [41] a similar approach was mentioned for the special training sequence used, but this was not treated further.

The complexity of the synchronizer is an important issue, especially in sleep mode where low power consumption is important. The synchronization methods based on the power comparison and the separate signaling channel are well suited for low-power implementations, but there is a large performance degradation compared to the correlation-based methods. Low-complexity variants of the correlation-based methods using only one-bit resolution are proposed in [9] and [37] for synchronization using the cyclic prefix and in [D] for the case of a PN-sequence preamble. In [37] the synchronization signal is used both for time and frequency synchronization and it is shown that it is possible to achieve satisfactory results for the frequency offset estimate using only one-bit resolution of the received samples. However, the cyclic prefix was quite long and averaging over several symbols was required in order to decrease the variance. The PN-based preamble is especially suited for the one-bit resolution synchronizer since the transmitted sequence has constant amplitude and correlation is performed before multiplication. In [D] it is shown that the low-resolution PN-based synchronizer in fact can achieve lower false detection probability and lower probability of missing the synchronization signal compared to a full-resolution conventional synchronizer based on a repeated OFDM symbol. The short preamble used is, however, not long enough to give satisfactory results for low-resolution frequency offset estimation.

The variance of the frequency offset estimate is similar for all the correlation-based methods. The performance is in principle determined by the number of observed samples and the delay for the correlation. The OFDM-based methods where sample-by-sample multiplication is used show a very robust behavior. The phase of the synchronization signal has a flat region around the correct timing and the sensitivity of timing errors is therefore low for frequency estimation using these methods. The PN-based method has a very sharp synchronization peak. This makes it very easy to identify the peak, but if there is a timing error the performance of the frequency offset estimate will deteriorate significantly.

2.2.3 Pilot Pattern Design

When designing the pilot pattern the task is to use sufficiently many pilots in order to suppress the noise and to be able to follow the changes in the transfer function without increasing the energy loss and bandwidth expansion too much. The receiver has to estimate the values of the transfer function in a two-dimensional grid, and therefore it is also important where the pilots are located. The pilot pattern describes where in the time and frequency plane to transmit the pilot symbols. There are only a few papers in the literature

regarding the design of this pattern. Kaiser gives a good overview of channel estimation and related topics for multi-carrier systems in [39], but it is hard to find analytical treatments. In [A] it was shown that the pilot pattern has a large influence on the resulting performance. By designing the pattern in an efficient way it is possible to increase the Doppler range over which the system can operate by up to ten times [A] compared to a system using a pilot pattern similar to a single-carrier case. In the paper, a Kalman filter was used to design the pilot pattern by searching for the best place to put the next pilot symbol. Much work can be put into the design of the pilot patterns, and optimal channel estimation filters can be used for the analyses. In real-time operation less complex filters can naturally be used. In [53] a Wiener filter was used for channel estimation and the theoretical symbol error rate was derived based on the MSE of the estimation filter. The influence of the pilot spacing in time and frequency was then investigated using a rectangular pattern. Under optimal conditions, i.e. when the correlation functions are known, the authors in [53] concluded that a pilot spacing close to the sample theorem was sufficient. However, for worst-case designs, or when less complex non-optimal channel estimation filters are used, the distance between the pilot symbols has to be decreased by at least a factor two. The maximum pilot distance in frequency (measured in number of sub-channels) due to the sampling theorem is [19]

$$N_F \leq \frac{1}{\tau_{\max} f_{sc}}, \tag{2.17}$$

where $\tau_{\max}$ denotes the maximum excess delay of the channel. The maximum pilot distance in time is similarly given by (2.3). The maximum pilot distance in time is therefore proportional to the coherence bandwidth, normalized by the sub-channel bandwidth. In [31] the term "balanced design" was introduced, meaning that the estimation uncertainty in the time direction and in the frequency direction should be the same. As a rule of thumb the authors proposed two times oversampling, resulting in the following guideline[1]:

$$2 f_D T_s \cdot N_T \approx \tau_{\max} f_{sc} \cdot N_F \approx 1/2, \tag{2.18}$$

where N_F is the pilot spacing in the frequency direction. This guideline coincides with the results in [53].

The publications mentioned above deal mainly with rectangular pilot patterns according to Figure 2.4. Another alternative is to use a hexagonal pattern and this approach was investigated in [24]. The authors compared the performance of different patterns and concluded that a hexagonal grid gives better performance since it leads to a more efficient sampling of two-dimensional signals. However, to make a more general conclusion, investigations for more channels than the HF channel addressed in [24] are required.

[1] In the original publication [31] "balanced design" was defined as $f_D T_s \cdot N_T \approx \tau_{\max} f_{sc} \cdot N_F \approx 1/4$, but this was changed to (2.18) in [39].

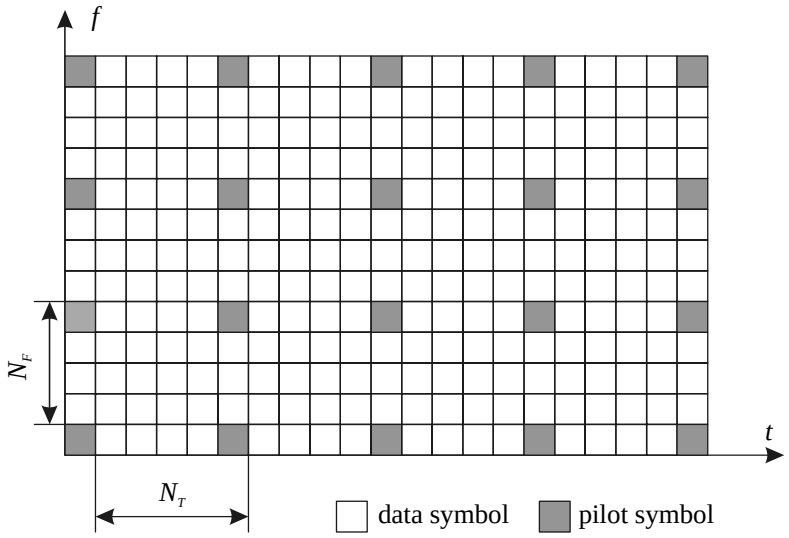

Figure 2.4: Example of a rectangular pilot pattern. N_F is the pilot distance in frequency whereas N_T is the distance in time.

The important conclusion in all the papers is to spread the pilot symbols in both time and frequency, otherwise there will be a performance degradation due to overhead or inaccurate channel estimates. In the HiperLAN/2 system [4] the pilot symbols are continuously transmitted at the same sub-channels during the whole block and this property is therefore not fulfilled. Each data block is preceded by a preamble, but for a long block it is hard to make full use of the time correlation of the channel when deriving the channel estimates. In [80] it was shown that it was possible to gain 1-1.5 dB just by rearranging the pilot symbols so they are located at different sub-channels. This study was performed for an uncoded system without using the preamble for channel estimation, but simulations showed [20] that the same gain could be achieved in a coded system including the preamble for channel estimation.

The channel model for the time correlation used in [A] is an auto-regressive model, which also has been used by other researchers [36] [84]. It is hard to compare the resulting pilot patterns in [A] directly to the guidelines above since the shape of the correlation functions is different. In addition, there is a difference in that only one pilot symbol at a time is used in [A], while many pilot symbols are transmitted simultaneously in the other investigated systems. However, the general result still holds that the pilot symbols shall be spread both in the time and in the frequency direction so that the time correlation and frequency correlation of the channels can be used to improve the channel estimates. In [19] the influence of the edge effects was mentioned

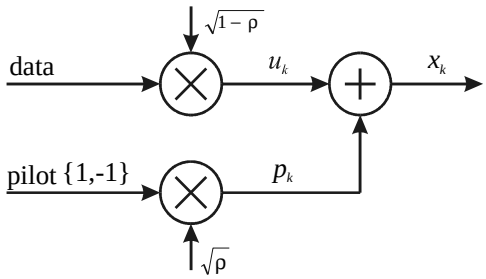

Figure 2.5: Generation of the signal when using superimposed pilot symbols.

and it was suggested to have more pilot symbols at the outer symbols in a packet to mitigate this effect. This corresponds well to the result of the pilot pattern search in [A], where the outer channels are used for pilot symbols more frequently.

2.3 Superimposed Pilot Sequences

Superimposed pilot sequences means that the known pilot sequence is overlaid on the data sequence and transmitted in the same frequency band and at the same time as the data sequence. The transmitted signal contains both a known pilot part and an unknown data part. It is generated according to Figure 2.5 and can be described as

$$x_k = u_k + p_k, \tag{2.19}$$

where p_k is the known pilot sequence with energy $\rho \cdot E_s = E[|p_k|^2]T_s$ and u_k is the unknown data part with energy $(1-\rho) \cdot E_s = E[|u_k|^2]T_s$. The power of the pilot sequence is only a small fraction of the transmitted power, but the pilot sequence can anyhow be used for channel estimation or time and frequency synchronization. The pilot sequence can be treated either as extra noise or as a known displacement of the constellation. In the latter case the disturbance on the data sequence can be made negligible, but otherwise the distortion on the data part is anyhow small.

The idea of superimposing a pilot sequence has been proposed by various authors for different applications. Holden and Feher proposed in [34] a system for carrier recovery based on superimposed pilot sequences. They called the technique spread-spectrum pilot technique (SSPT) since the pilot sequence used was a PN-sequence, which was linearly added to the inphase data channel. By multiplying the incoming signal with the known PN-sequence it is possible to recover the carrier signal in an easy way. After de-spreading, the carrier signal is narrowband and therefore much of the distortion from the data

signal can be filtered out. Compare this to the processing gain achieved in CDMA systems. The method showed to be robust for low SNRs, the system was implemented in hardware and the carrier recovery unit worked well down to $E_b/N_0 = 0$ dB. Bejjani and Belfiore proposed a system using a variant of superimposed pilot sequences [10] for underwater communication. The goal there was to enable tracking of a rapidly changing channel by introducing a continuous pilot sequence. In the system, two different orthogonal pulse shapes were used for the data and pilot sequence. Therefore, there was a loss in spectral efficiency, but the system showed to be robust for high Doppler frequencies. This technique is similar to the UMTS/W-CDMA system in uplink, where the data and pilot sequences are transmitted simultaneously [35], but using the orthogonal inphase and quadrature channels, respectively. Both these methods mean that the spectral efficiency is decreased since no data is transmitted on the pilot channel. Superimposed pilot sequences have also been used for frame synchronization [67]. The authors used a non-orthogonal pilot sequence to determine the start of frames in a wireless system. The pilot sequence was not used for bit synchronization and therefore the influence of the data could be removed before correlating the received pilot sequence with the known pilot sequence. Then, when frame synchronization was achieved, the influence on the data symbols was removed by subtracting the pilot sequence. This technique also showed good synchronization performance compared to conventional systems, especially in fast Rayleigh fading channels where the long continuous pilot sequence means that time diversity is achieved.

In paper [E] superimposed pilot sequences are used for time and frequency synchronization in OFDM systems and in paper [F] the technique is used for joint channel estimation and detection in flat Rayleigh fading channels. The latter paper is not restricted to multi-carrier systems, but can be used for any system where the time dispersion is negligible. The synchronizer structure in [E] has low complexity and is suitable for implementation but the complexity of the channel estimator in [F] is quite high and therefore low-complexity approximations are of interest. The scope of [F] has been to show the potential of using superimposed pilot sequences for channel estimation and to derive an optimal receiver structure. The paper however also addresses complexity issues and possible simplified receiver structures.

2.3.1 Time and Frequency Synchronization for OFDM

Superimposed pilot sequences are used in [E] for time and frequency synchronization in OFDM systems. The advantage of using superimposed pilots for this purpose is that a continuous sequence is present that can be used for synchronization purposes. This means that the frequency-offset range can be increased compared to methods where the cyclic prefix is used for synchronization. Further, the time synchronization signal shows a more attractive behavior since it has a sharper peak compared to the methods incorporating a cyclic pre-

fix. The technique is suitable for both tracking and acquisition and it is very flexible in that the power of the pilot sequence can be adjusted to the specific needs. In acquisition mode all the transmitted power can be devoted to the pilot sequence and then the transmitted signal is equivalent to the preamble described in [D]. By using all the power for the pilot sequence during acquisition, the acquisition time can be made small. During tracking, the pilot power can be low without loosing performance since the parameters do not change that fast and averaging over several symbols can be used.

The synchronization signal is calculated in the same way as for the pre-correlating synchronizer according to (2.16) and Figure 2.3. The received signal is first correlated with the known sequence of length R. Then the correlated values are multiplied in pairs and finally L of these products are summed in order to get the synchronization signal. The only difference compared to the preamble synchronization scheme is that the data sequence is transmitted at the same time as the synchronization sequence and therefore acts like an additional disturbance. Therefore, the correlator length or the number of products used for the synchronization signal has to be increased compared to the case with the synchronization signal only, i.e. the preamble. The amplitude of the synchronization signal is used for time synchronization and when the correct timing is found the frequency-offset estimate is given by the phase angle of the synchronization signal.

In the literature no similar work can be found for OFDM systems, at lest as far as I know. The method has some similarities to synchronization in CDMA systems. Often in CDMA systems, for example in the UMTS/W-CDMA system, there is a special synchronization channel transmitted on top of the data channels. Since the synchronization channel can be shared between the users, the overhead is small anyway. However, as opposed to the superimposed pilot sequences, the synchronization channel in the CDMA system is orthogonal to the data channels by means of different codes. This means somewhat better detection properties, since the synchronization scheme based on superimposed pilot sequences relies on statistical averaging for detection of the synchronization signal.

In the spread-spectrum pilot system in [34] superposition of the pilot sequence and data sequence was made in the analog domain. This means that for this system bit synchronous transmission could not be guaranteed. This is in contrast to the superimposed pilot sequences in [E], where this property is essential for time synchronization and later for channel estimation in [F].

2.3.2 Channel Estimation using Superimposed Pilot Sequences

The superimposed pilot sequences can also be used for channel estimation and detection. Due to the continuous transmission, the technique is also very useful in fast fading conditions. Pilot-based systems fail when the sampling theorem is

not fulfilled, and for conventional systems this becomes an important limitation at high Doppler frequencies. Different approaches can be used for detection based on superimposed pilot sequences. The pilot sequence can be seen as a pilot tone that is spread over the whole transmission bandwidth or as a small shift of the signal constellation. In [F], a system based on joint maximum likelihood detection of the data and the channel is described. This means that the pilot sequence and the data sequence both are used for channel estimation and detection. Other possible approaches would be to first estimate the channel by help of the pilot sequence and then detect the data based on this estimate. In [34] the synchronization signal was extracted and after that the disturbance of the data sequence was decreased by changing the detection threshold with regards to the pilot sequence. In [67] the data signal was first detected and its influence on the synchronization signal was removed before extraction of the synchronization signal. However, the latter is not possible for data detection and channel estimation since the influence of the data can not be removed before extraction of the synchronization parameters and vice versa.

Joint channel estimation and detection is well described in [48]. The aim is to find the data sequence and channel estimate that maximizes the probability density function of the received sequence according to (2.1). Assume that the received signal sample at time k can be described by

$$r_k = h_k c_k + n_k, \tag{2.20}$$

where the noise n_k and multiplicative channel c_k are jointly Gaussian variables and that the data symbols are independent random variables with equal probabilities. It can be shown that [48] the optimal joint estimation-detection rule can be formulated as

$$(\hat{\mathbf{c}}_{\mathbf{ML}}, \hat{\mathbf{h}}) = \arg\max_{\mathbf{c},\mathbf{h}} f(\mathbf{r}, \mathbf{h}|\mathbf{c}) = \arg\max_{\mathbf{c},\mathbf{h}} f(\mathbf{h}|\mathbf{r}, \mathbf{c}) f(\mathbf{r}|\mathbf{c}). \tag{2.21}$$

A suitable strategy is therefore to maximize $f(\mathbf{h}|\mathbf{r}, \mathbf{c})$ to find the MAP (maximum a posteriori) channel estimate for each possible data sequence, $\mathbf{c}$, and use this as a true parameter when selecting the sequence with the largest likelihood. Since the channel and noise variables are assumed to be Gaussian, the MAP estimate is equal to the minimum mean square error estimate, which in turn can be calculated using a Wiener filter [40]. This is the approach in [F]. For each hypothesis of the data sequence the channel estimate is calculated using a Wiener filter. This channel estimate is then used to calculate the metric such that the most probable sequence can be found. In [F] both an optimal and an approximate metric is derived. The use of the optimal metric requires that all possible data sequences are taken into account when looking for the sequence resulting in the lowest metric, whereas the approximate metric is can be used for a less complex detector by using per-survivor processing [57].

In [G] there is a comparison between channel estimation and detection using superimposed pilot sequences and other detection and estimation methods. The

other analyzed methods include PSAM, multiple-symbol differential detection, a novel hybrid detection method introduced in [G] and differential detection. The analysis is based on the calculation of the pairwise error probability that the decision is made in favor of a specific erroneous data sequence instead of the correct one. The pairwise probability has previously been used to analyze TCM [16] and multiple-symbol differential detection [29]. The method is very useful when the decisions are dependent between successive symbols. Such situations occur, for example, when using TCM or superimposed pilot sequences for channel estimation as described above. In these situations, it is impossible or tricky to calculate exact expressions for the BER, but by expressing the decision variable in a quadratic form of Gaussian variables it is possible to derive expressions for the pairwise error probability and approximate expressions for the BER.

Chapter 3

Summary and Contributions

In this chapter, some general results are presented and the contents and main results in each paper are described. There is also a presentation of the contributions in each paper to the specific research field. Thereafter, there is a discussion of possible directions for future wireless communication systems. Finally, some general conclusions are presented and possible further work in the research fields is discussed.

3.1 General Results

Future communication systems have to provide high data rates in highly disturbed environments. The demand for the services is expected to be high. Therefore new frequency bands have to be used and, due to the lack of bandwidth, these will be located at higher frequencies. These things lead to increased importance of synchronization and channel estimation issues in the future. The use of higher frequency bands means that the Doppler frequency is increased for a certain velocity of the mobile terminal. The many users mean that the overall interference level will increase. This makes it necessary to be able to operate in highly disturbed environments. In addition, the requirements of spectral efficiency call for the use of higher order modulation schemes, and these are more sensitive to synchronization and channel estimation errors. In most textbooks on digital communications, the synchronization issues are, however, neglected and synchronization parameters are often provided by a genie. In real systems no genie exists and the synchronizers and channel estimators have a large influence on the performance. The goal of the research has been to study these key components and suggest improvements such that the requirements on future systems can be met.

The focus of the thesis is problems related to the wireless mobile channel. This channel has some important properties that call for special attention; the channel changes rapidly and it has large amplitude fluctuations. Therefore, the system has to be optimized according to these specific properties. The task of the synchronizer and channel estimator is to enable communication also under severe conditions. In the thesis, we have therefore studied robust synchronization and channel estimation techniques, suitable also for low SNRs and fast fading channels. The work has been focused on future systems and systems under standardization. However, the principles are the same for existing systems and the analyses can be applied to these as well. In the thesis the influence of the pilot pattern in OFDM systems is studied and guidelines for the design are given. Synchronization in OFDM systems based on preambles is also analyzed and a new synchronizer structure is proposed. Further, the subchannel bandwidth in wireless OFDM systems is optimized with regards to the characteristics of the channel. The use of pre-compensation is also investigated by deriving expressions for the BER and determining feasible block lengths. Finally, the use of superimposed pilot sequences is analyzed and proposed for synchronization and channel estimation issues.

The requirements on the synchronizer and channel estimator are not static over a longer time. They change when new technologies and techniques are introduced. Improved semiconductor technology has meant that more complex algorithms than before can be used. The new algorithms, used as an example for detection or error correction, are often more accurate than the previously used ones, but also more sensitive to the accuracy of the synchronization parameters. Errors that previously were "hidden" due to a limited resolution and did not affect the result can play an important role in new systems with better resolution. The goal is that the "supporting functions" should not be a limitation of the system performance. Complexity is however still an important issue for the algorithms, since the complexity should be spent on functions providing the largest gain.

3.2 Paper Contributions

3.2.1 Pilot Assisted Channel Estimation for OFDM in Mobile Cellular Systems

This paper contains an analysis of the influence of the pilot pattern on the performance of a mobile OFDM system. The bit error performance is calculated for five different pilot patterns based on the mean square error of the channel estimate. In the paper an autoregressive process is assumed for the changes of the Rayleigh fading channel and a Kalman filter is used for channel estimation. The estimation filter is also used to calculate the optimum pilot pattern in the sense of minimum uncoded BER for each of the OFDM symbols. The performance from this pilot pattern is compared to the performance when using other

pilot patterns found in the literature. It is shown that the pilot pattern has a large influence on the resulting bit error rate. A well-designed pilot pattern means that the channel changes can be tracked better, with a corresponding performance improvement. In a numerical example it is shown that the BER can be decreased by a factor of 5 or that the system can cope with 10 times higher Doppler frequency when using a suitable pilot pattern instead of one similar to a single-carrier case.

The main contribution of the paper is to show the importance of a well-designed pilot pattern and the large improvements that can be achieved. The paper is one of the first papers using an analytical approach to optimize the pilot pattern. The analysis is general and can be used to derive the amount of pilot symbols required and to find a suitable pilot pattern for any OFDM-based system.

3.2.2 Optimization of Sub-Channel Bandwidth for Mobile OFDM Systems

In this paper the bandwidth of the sub-channels is optimized with respect to interchannel interference (ICI) caused by the time-selective channel and the energy loss caused by the cyclic prefix. For a given duration of the cyclic prefix, the energy loss increases as the sub-channel bandwidth increases. On the other hand, the ICI caused by channel changes during an OFDM symbol increases as the sub-channel bandwidth decreases. An expression for the resulting uncoded BER as a function of the sub-channel bandwidth is therefore derived, including these effects. The influences of the Doppler frequency, SNR, and total data rate are investigated. In addition, the effects of introducing channel estimation are analyzed with respect to the choice of sub-channel bandwidth. It is concluded that the total data rate has virtually no effect on the optimum and that the optimization can be performed independent of channel estimation. It is shown that the optimal sub-channel bandwidth depends on the Doppler frequency and SNR. Higher Doppler, as well as an increased SNR, call for an increased sub-channel bandwidth to avoid that the ICI becomes the dominant source of errors. Finally, it is argued that a worst-case design with respect to the Doppler frequency should be performed.

The main contribution of the paper is the analytical optimization of the sub-channel bandwidth. There are not many papers addressing the choice of the number of sub-channels in OFDM systems, and at the time of publication the choice had previously been based only on simulations. In addition, the influence of channel estimation is investigated and it is concluded that the optimization can be performed independent of channel estimation and on a worst-case assumption of the Doppler frequency.

3.2.3 Pre-Compensation for Rayleigh Fading Channels in Time Division Duplex OFDM Systems

This paper investigates the possibility of using pre-compensation for compensating the influence of the Rayleigh fading channel in a wireless OFDM time division duplex system. Three compensation methods are analyzed:

1. phase-only compensation
2. phase and amplitude compensation up to a certain power level, above which the amplifier is saturated
3. phase and amplitude compensation up to a certain power level, above which the amplifier is switched off

The BERs and increase in peak-to-average power ratios are calculated for the different methods assuming perfect channel knowledge. The maximum block lengths are calculated for the phase compensated system, since this is assumed most appropriate for wireless applications. It is concluded that block lengths much longer than the one in the DECT system can be used before the instantaneous performance gets worse than that of a differential system.

The main contributions are the derivation of the resulting uncoded BER for the phase-only compensation method and the analysis of the influence of both the uplink block length and the downlink block length, respectively. The calculations are performed both for perfect knowledge of the channel correlation functions and under correlation mismatch. The maximum block length is a very important design parameter in the design of pre-compensated time division duplex systems and to my knowledge no analytical analysis of its influence on the performance exists in the literature. In other papers the channel is assumed to be constant or simulations are used to investigate the influence of block length and Doppler frequency. Other contributions in this paper include the derivation of the uncoded BER for the three methods when assuming perfect channel knowledge and the corresponding increase in peak-to-average power ratio.

3.2.4 Preamble-Based Time and Frequency Synchronization for OFDM Systems

This paper compares the synchronization performances achieved by two different preambles and two different synchronizer structures. The preambles under investigation are based on a repeated OFDM symbol and a repeated PN-sequence, respectively. The investigated synchronizers are a conventional synchronizer, where multiplication is performed on a per sample basis before summation, and a pre-correlating synchronizer, where correlation is performed with the known sequence prior to multiplication. In the paper a low-complexity

version of the pre-correlating synchronizer is also introduced, suitable for low-power operation during, for example, sleep mode. The performance when using each of the two preambles together with each of the two synchronizers is analyzed analytically. The probability density function of the time synchronization signal is derived for the four different combinations. Then the probability of false detection and the probability of missing the synchronization preamble are calculated for AWGN channels. The performance in Rayleigh fading channels is investigated by simulations. In addition, the variance of the frequency-offset estimate is given. It is shown that the pre-correlating synchronizer using the PN-based preamble results in significantly better detection performance compared to the conventional synchronizer using an OFDM-based preamble. The low-complexity synchronizer is also shown to have an attractive detection performance. The probability of false detection and the probability of missing the synchronization preamble is even lower than the corresponding probabilities using the full-resolution conventional synchronizer.

The main contributions are the introduction and analysis of the pre-correlating synchronizer for preamble-based synchronization in OFDM systems. The analysis is purely analytical and the statistics of the time synchronization signals are derived for the case of correct and wrong timing in AWGN channels. An important conclusion of the analysis is that it is possible to enhance the detection performance by introducing the pre-correlator. Another contribution is the introduction and analysis of the low-complexity coarse time synchronizer. The performance of this synchronizer is also analyzed analytically by calculating the probability distribution of the synchronization signal.

3.2.5 OFDM Time and Frequency Synchronization by Spread Spectrum Pilot Technique

In this paper, superimposed pilot sequences are introduced for time and frequency synchronization in OFDM systems. The paper starts with a description of the synchronization system based on superimposed pilot sequences. The main idea behind superimposed pilot sequences is to linearly add a known pilot sequence to the unknown data sequence. The two sequences are transmitted simultaneously and in the same frequency band. In the paper, a suitable synchronizer structure is also presented. The synchronization signal is calculated in the same way as for the pre-correlating preamble-based synchronizer described above. First, the received signal is correlated to the known sequence, and then the result is multiplied by an earlier correlation value to get a differential signal. The delay between the terms determines the frequency-offset range; decreasing the delay increases the frequency-offset range. The maximum frequency offset range is half the OFDM bandwidth, which should be compared to half the sub-channel bandwidth for conventional synchronization based on the cyclic prefix. The amplitude of the differential signal, or a sum of them, is the time synchronization signal and, when the correct timing is found, its

phase is used for the frequency-offset estimate. In the paper, the distribution of the synchronization signal is derived for AWGN channels, and verified by simulations. The presented synchronizer structure is of low complexity and it is shown that the presented method works well also at low SNRs. The power of the pilot sequence can be adjusted to the specific requirements, but it can in general be low due to suppression of noise and data when performing the correlation.

The main contributions include the introduction of superimposed pilot sequences for synchronization purposes in OFDM systems. The proposed method is independent of the cyclic prefix and is very flexible. To my knowledge, the derivation of the distribution functions of the time synchronization signal is also new, even though similar synchronization structures can be used in CDMA systems. Another contribution in the paper is the generalization of other synchronization methods such that the analysis can be applied to these as well.

3.2.6 Channel Estimation using Superimposed Pilot Sequences

In this paper, superimposed pilot sequences are introduced for the purpose of channel estimation. The system and receiver structure are thoroughly described. The main idea is to linearly add a known pilot sequence to the transmitted data sequence and perform joint channel estimation and detection in the receiver. The detection method is based on hypotheses of the transmitted data sequence and for each hypothesis, a channel estimate is derived by means of per-survivor processing. The final decision is made in favor of the sequence with the lowest distance to the received sequence. The proposed system is analyzed analytically for flat Rayleigh fading channels. The pairwise error probability is derived as well as approximations of the BER. The performance of the system is verified by simulations and the influences of different parameters such as pilot power, filter coefficients and filter lengths are investigated. It is shown that the technique is robust; it works well for a wide range of SNR values and also for high Doppler frequencies. The superimposed pilot sequences are transmitted in the same band as the data symbols and, as opposed to conventional pilot-based channel estimation techniques, there is no bandwidth expansion. The power of the pilot sequence can be quite low. Typically 5% of the total power is enough. The technique can be used both in existing and new systems and it results in better power efficiency than conventional pilot symbol assisted modulation, especially in fast fading environments.

The main contributions are the introduction of channel estimation based on superimposed pilot sequences and the corresponding derivation of the optimal receiver. The receiver relies on joint detection and channel estimation and both an optimal metric and a metric suitable for a recursive receiver structure are derived. The analysis of the system is also new, even though similar analyses can be found for trellis-coded modulation. Finally, the system is verified by

extensive simulations and the influence of filter lengths, non-optimal filters, pilot-to-data power ratio, and Doppler frequency have been investigated.

3.2.7 A Comparative Analysis between Different Detection Schemes in Flat Rayleigh-fading Channels

This paper contains a comparison between different coherent and non-coherent detection methods. The analysis is based on calculation of the pairwise sequence error probability and a common framework is used for the analysis of all the investigated methods. The analyzed methods include: pilot symbol assisted modulation (PSAM), multiple-symbol differential detection, superimposed pilot sequences, differential detection, and a novel hybrid method, similar to PSAM, which uses both pilot symbols and hypotheses of the data symbols for channel estimation. In the paper, it is shown that the decision variable for all methods can be written in a quadratic form of Gaussian variables, for which the probability density function is derived. The pairwise error probability is compared for the error sequences dominating the error rate. The performance is analyzed for different SNR and Doppler values and it is concluded that superimposed pilot sequences, multiple-symbol differential detection and the hybrid method lead to robust detection both for a wide range of SNRs and for high Doppler frequencies. Conventional differential detection shows a performance loss due to the lack of noise averaging when using the previously received symbol as a reference for detection. Together with PSAM this method also suffer from an error floor for high Doppler values due to inaccurate "channel estimates".

The main contribution in the paper is the introduction of the common framework used for the analysis of the different coherent and non-coherent detection methods. This common framework is used to find the sequences dominating the error rate and also to investigate the influence of different noise levels and different Doppler frequencies. The presented analysis is general and can be applied to other detection methods as well.

3.3 Conclusions and Discussion

While being at the department, I have had the opportunity to follow the development of wireless techniques and standards for some years. In the thesis I have presented research results that I have produced during the last four years. The question is now what will happen next? Here I will give some short personal views, but as always when regarding the future, they are only guesses.

If we start at a high level, there are some certain events that will change the availability of access to fast wireless or mobile communication. During 2001, the first third-generation systems based on wideband-CDMA will be on the market. At the same time there will be large investments to upgrade the existing second-generation systems so that these also can offer higher data

rates and packet based communication. This means that data rates of up to 384 kbit/s can be offered over wide areas. In a few years the new W-LAN systems will also be on the market. These systems will work in unlicensed frequency bands and offer data rates of around 20 Mbit/s, but with a limited coverage. Finally, the first commercial Bluetooth devices will be launched soon. This means that short distance wireless communication with data rates of up to around 700 kbit/s is soon available for a wide range of electronic devices.

I think that the importance of the latter two systems can not be overestimated. These systems will offer fast wireless communication "for free" and can be used as cheap links between any electronic device and the access point. This means that it is possible for low-cost devices to have access to fast and advanced communication links where the mobile terminal acts more like a relay station. Probably there will be a wide range of low-cost third party products offering new services using indirect access to the existing network.

I also think that the W-LAN systems will have a very important role in the near future. Except for coverage and maybe mobility, these systems are actually fulfilling all the requirements stated by some of the leading actors in the area for a fourth generation system. This means that you can achieve data rates high enough both for video and high-quality audio applications. It is also possible to be your own operator and offer broadband services in an office or a residential area "free of charge". Further, you can set up your own personal area network where electronic devices in your home can communicate with each other. Hopefully these systems will be user-friendly enough not to prevent their usage. The use of the unlicensed bands is very important since this means that in your "own" system there will be no reluctance using it due to high transmission fees.

In this scenario there is no cell planning and the interference level will sometimes be high. The high transmission rates means that, in order to limit the transmitted power to appropriate levels, the energy per bit will be decreased compared to today's systems. Therefore, the systems need to be able to adapt the transmission to the present situation and operate in highly disturbed environments. Fast and reliable synchronization and channel estimation are important in order to prevent the systems from being limited by these functions, rather than by noise or interference.

Maybe the most important thing that I have learned during these years is that wireless communication systems are complex systems and the true challenge is to optimize the *system* and get everything to work together. It is no use to only improve certain parts as long as the rest of the system can not make use of these improvements. To me it seems that in the highly optimized systems the different functional blocks become more dependent on each other, the information exchange between the blocks has become more important. The channel estimator, for example, has not only to deliver channel estimates to the equalizer, but also to the decoder such that the decoding can be optimized; and maybe to the transmitter such that the transmitted signal can be adjusted

to the specific channel. In order to challenge today's limits the receiver has to incorporate more and more of the available information, often to the price of increased complexity. There is also a question whether this increase in complexity can be afforded and how large the resulting gain is. As always, it is a question of cost vs. performance.

In order to fulfill the requirements on synchronization and channel estimation I think that it is well worth spending a few percent of the transmitted power and bandwidth on known pilot symbols. The pilot symbols mean that faster and less complex algorithms can be used, often resulting in more robust detection and better performance. The loss caused by the pilot symbols is often small compared to the gain that can be achieved, so pilot symbols are generally a good "investment"

In OFDM systems it is really important to limit the influence of the frequency offset. This effect often results in undesirable constraints in the system design, especially when using low cost terminals where, in general, the accuracy of the oscillators is low. Therefore, I think that it is important to look for fast and robust estimation techniques as well as cheap and accurate oscillators. In order to get efficient systems it is important that the performance is not limited by hardware impairments, but rather by the characteristics of the physical channel or by interference from other users.

Regarding future work on the material presented in this thesis, I think that the paper on pre-compensation rather easily can be extended to account for antenna diversity as well. This makes the paper more complete by also covering systems similar to the antenna pre-compensation system in the UMTS/W-CDMA system. The preamble-synchronization paper can be extended to include preamble-based channel estimation. In its present form, the focus is on acquisition of time and frequency synchronization parameters only. In packet based systems there are, however, some initial transients when a packet arrives and it is not meaningful to make a final channel estimate before reaching some kind of steady state. Anyway, the required information is there, after correlation the channel impulse response is available as a coarse channel estimate. The use of superimposed pilot sequences is also worth further investigations. First of all, it would be interesting to look at less complex receiver structures. The technique seems promising, especially for fast fading channels, but the complexity is still too high. It would be very interesting to combine synchronization and channel estimation to actually get a working system suitable for implementation.

Bibliography

[1] 802.11b-1999 suppplement to information technology–telecommunications and information exchange between systems–local and metropolitan area networks–specific requirements–part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications: Higher speed physical layer (PHY) extension in the 2.4 GHz band. Technical Report P802.11b, IEEE - Institute of Electrical and Electronics Engineers, Inc., New York, USA, 1999.

[2] Digital Enhanced Cordless Telecommunications (DECT); common interface (CI); part 2: Physical layer (PHL). EN 300 175-2 V1.4.2, ETSI – European Telecommunications Standards Institute, Sophia Antipolis, France, June 1999.

[3] Draft supplement to information technology-telecommunications and information exchange between systems - local and metropolitan area networks - specific requirements - part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications: High speed physical layer in the 5 GHz band. Technical Report P802.11a, IEEE - Institute of Electrical and Electronics Engineers, Inc., New York, USA, 1999.

[4] Broadband Radio Access Networks (BRAN); HIPERLAN Type 2; Physical (PHY) layer. TS 101 475 V1.1.1, ETSI – European Telecommunications Standards Institute, Sophia Antipolis, France, Apr. 2000.

[5] A. Aghamohammadi, H. Meyr, and G. Ascheid. A new method for phase synchronization and automatic gain control of linearly-modulated signals on frequency-flat fading channels. *IEEE Trans. Commun.*, 39(1):25–29, Jan. 1991.

[6] A. G. Armada and M. Calvi. Phase noise and sub-carrier spacing effects on the performance of an OFDM communication system. *IEEE Communications Letters*, 2(1):11–13, Jan. 1998.

[7] A. R. S. Bahai and B. R. Saltzberg. *Multi-Carrier Digital Communications: Theory and Applications of OFDM.* Kluwer Academic Publishers, Dordrecht, The Netherlands, 2000.

[8] J.-J. van de Beek, M. Sandell, and P. O. Börejesson. ML estimation of time and frequency offset in OFDM systems. *IEEE Trans. Signal Proc.*, 45(7):1800–1805, July 1997.

[9] J.-J. van de Beek, M. Sandell, M. Isaksson, and P. O. Börejesson. Low-complex frame synchronization in OFDM systems. In *Int. Conf. on Universal Personal Communication*, pages 982–986, Tokyo, Japan, Nov. 1995.

[10] E. Bejjani and J.-C. Belfiore. Multicarrier coherent communications for the underwater acoustic channel. In *Proc. Oceans'96*, volume 3, pages 1125–1130, Fort Lauderdale, FL, USA, 1996.

[11] R. Benjamin, I. Kaya, and A. Nix. Smart base stations for dumb time-division duplex terminals. *IEEE Commun. Mag.*, pages 124–131, Feb. 1999.

[12] J. A. C. Bingham. Multicarrier modulation for data transmission: An idea whose time has come. *IEEE Commun. Mag.*, 28(5):5–14, May 1990.

[13] K. Brüninghaus and M. Radimirsch. Coarse frame synchronization for OFDM based wireless communication systems. In *Proc. Int. Symp. on Personal, Indoor and Mobile Radio Commun.*, pages 806–810, Boston, MA, USA, Sept. 1998.

[14] S. Carlsson. Generationsväxling för mobiltelefon (in Swedish). *TELE*, 89(2):37, 1983.

[15] J. K. Cavers. An analysis of pilot-symbol assisted modulation for Rayleigh-fading channels. *IEEE Trans. Vehic. Technol.*, 40(4):686–693, Nov. 1991.

[16] J. K. Cavers and P. Ho. Analysis of the error performance of trellis-coded modulation in Rayleigh-fading channels. *IEEE Trans. Commun.*, 40(1):74–83, Jan. 1992.

[17] R. W. Chang. Synthesis of band-limited orthogonal signals for multichannel data transmission. *Bell System Tech. J.*, 45:1775–1796, Dec. 1966.

[18] L. J. Cimini. Analysis and simulation of a digital mobile channel using orthogonal frequency-division multiplexing. *IEEE Trans. Commun.*, COM-33(7):665–675, July 1985.

[19] F. Classen, M. Speth, and H. Meyr. Channel estimation units for an OFDM system suitable for mobile communication. In *ITG Conference on Mobile Radio*, Neu-Ulm, Germany, Sept. 1995.

[20] M. de Courville, B. Muquet, and S. Simoens. New elements underlying the need for a midified pilot pattern. ETSI EP BRAN doc. no. HL13.5Mot1A, European Telecommunications Standards Institute, Sophia Antipolis, France, Apr. 1999.

[21] D. Divsalar and M. K. Simon. Multiple-symbol differential detection of MPSK. *IEEE Trans. Commun.*, 38(3):300–308, Mar. 1990.

[22] O. Edfors. *Low-complexity algorithms in digital receivers.* Phd. thesis, Luleå University of Technology, Luleå, Sweden, Sept. 1996.

[23] O. Edfors, M. Sandell, J.-J. van de Beek, S. K. Wilson, and P. O. Börjesson. OFDM channel estimation by singular value decomposition. *IEEE Trans. Commun.*, 46(7):931–939, July 1998.

[24] M. J. Fernandez-Getino Garcia, J. M. Paez-Borrallo, and S. Zazo. Efficient pilot patterns for channel estimation in OFDM systems over HF channels. In *Proc. IEEE Vehic. Technol. Conf.*, volume 4, pages 2193–2197, Amsterdam, The Netherlands, Sept. 1999.

[25] J. C. Haartsen. The Bluetooth radio system. *IEEE Pers. Commun.*, 7(1):28–36, Feb. 2000.

[26] L. Hanzo, W. Webb, and T. Keller. *Single- and Multi-carrier Quadrature Amplitude Modulation: Principles and Applications for Personal Communications, WLANs and broadcasting.* John Wiley & sons, Chichester, England, 2nd edition, 2000.

[27] S. Haykin. *Adaptive Filter Theory.* Prentice Hall, Upper Saddle River, NJ, USA, 3rd edition, 1996.

[28] P. S. Henry and G. B. S. A new approach to high-capacity digital mobile radio. *Bell System Tech. J.*, 60(8):1891–1904, Oct. 1981.

[29] P. Ho and D. Fung. Error performance of multiple-symbol differential detection of PSK signals transmitted over correlated Rayleigh fading channels. *IEEE Trans. Commun.*, 40(10):1566–1569, Oct. 1992.

[30] P. Hoeher. TCM on frequency-selective land-mobile fading channels. In E. Biglieri and M. Luise, editors, *Coded Modulation and Bandwidth-Efficient Transmission*, pages 317–328. Elsevier, Amsterdam, The Netherlands, Sept. 1991.

[31] P. Hoeher, S. Kaiser, and P. Robertson. Pilot-symbol-aided channel estimation in time and frequency. In K. Fazel and G. P. Fettweis, editors, *Multi-Carrier Spread Spectrum*, pages 169–178. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1997.

[32] P. Hoeher and F. Tufvesson. Channel estimation with superimposed pilot sequence. In *Proc. Globecom*, pages 2162–2166, Rio de Janeiro, Brazil, Dec. 1999.

[33] P. Hoeher and F. Tufvesson. Channel estimation with superimposed pilot sequence applied to multi-carrier systems. In K. Fazel and S. Kaiser, editors, *Multi-Carrier Spread-Spectrum and Related Topics*, pages 295–302. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1999.

[34] T. Holden and K. Feher. A spread spectrum based system technique for synchronization of digital mobile communication systems. *IEEE Trans. Broadc.*, pages 185–194, Sept. 1990.

[35] H. Holma and A. Toskala, editors. *WCDMA for UMTS, radio access for third generation mobile communications*. John Wiley & sons, Chichester, England, 2000.

[36] S. J. Howard and K. Pahlavan. Autoregressive modelling of wide-band indoor radio propagation. *IEEE Trans. Commun.*, 9(40):1540–1552, Sept. 1992.

[37] M.-H. Hsieh and C.-H. Wei. A low-complexity frame synchronization and frequency offset compensation scheme for OFDM systems over fading channels. *IEEE Trans. Commun.*, 48(3):1596–1609, Sept. 1999.

[38] I. Jeong and M. Nakagawa. A time division duplex CDMA system using assymetric modulation in duplex channel. *IEICE Trans. Commun*, E82-B(12):1956–1963, Dec. 1999.

[39] S. Kaiser. *Multi-carrier CDMA mobile radio systems - Analysis and Optimization of Detection, Decoding, and Channel Estimation*. Ph.D. thesis, Technische Universität München, Munich, Germany, Jan. 1998.

[40] S. M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, Upper Saddle River, NJ, USA, 1993.

[41] U. Lambrette, M. Speth, and H. Meyr. OFDM burst frequency synchronization by single carrier training data. *IEEE Communications letters*, 1(2):46–48, Mar. 1997.

[42] D. Landström, S. K. Wilson, J.-J. van de Beek, P. Ödling, and P. O. Börjesson. Symbol time offset estimation in coherent OFDM systems. In *Proc. Intern. Conf. Commun.*, volume 1, pages 500–505, Vancover, Canada, June 1999.

[43] J. Li and M. Kavehrad. Effects of the time selective multipath fading on OFDM systems for broadband mobile applications. *IEEE Communication Letters*, 3(12):332–334, Dec. 1999.

[44] Y. G. Li, L. J. Cimini, and N. R. Sollenberger. Robust channel estimation for OFDM systems with rapid dispersive fading channels. *IEEE Trans. Commun.*, 46(7):902–915, July 1998.

[45] A. Masoomzadeh-Fard and S. Pasupathy. Combinde equalization and differential decoding using precoding. *IEEE Trans. Commun.*, 45(3):274–278, Mar. 1997.

[46] D. Matic, N. Petrochilos, T. A. Coenen, F. Schoute, and R. Prasad. Acquisition of synchronisation parameters for OFDM using a single training symbol. In K. Fazel and S. Kaiser, editors, *Multi–Carrier Spread Spectrum and Related Topics*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1999.

[47] H. Matsuki and H. Takanashi. Theoretical analysis of BER improvement by adaptive transmitter power control for Rayleigh fading compensation. *Electron. Lett.*, 29(17):1520–1521, Aug. 1993.

[48] H. Meyr, M. Moeneclaey, and S. A. Fechtel. *Digital Communication Receivers*. John Wiley & sons, New York, USA, 1998.

[49] M. L. Moher and J. H. Lodge. TCMP – A modulation and coding strategy for Rician-fading channels. *IEEE J. Select. Areas Commun.*, 7(9):1347–1355, Dec. 1989.

[50] P. H. Moose. A technique for orthogonal frequency-division multiplexing frequency offset correction. *IEEE Trans. Commun.*, 42(10):2908–2914, Oct. 1994.

[51] M. Münster, T. Keller, and L. Hanzo. Co-channel interference suppression assited adaptive OFDM in interference limited environments. In *Proc. IEEE Vehic. Technol. Conf.*, pages 284–288, Amsterdam, The Netherlands, Sept. 1999.

[52] R. van Nee and R. Prasad. *OFDM for wireless multimedia communications*. Artech House, Boston, USA, 2000.

[53] R. Nilsson, O. Edfors, M. Sandell, and P. O. Börjesson. An analysis of two-dimensional pilot-symbol assisted modulation for OFDM. In *Proc. Intern. Conf. Pers. Wireless Commun.*, pages 71–74, Bombay, India, Dec. 1997.

[54] A. Peled and A. Ruiz. Frequency domain data transmission using reduced computational complexity algorithms. In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, pages 964–967, Denver, CO, USA, 1980.

[55] T. Pollet, M. van Bladel, and M. Moeneclaey. BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise. *IEEE Trans. Commun.*, 43(2/3/4):191–193, Feb/Mar/Apr 1995.

[56] J. G. Proakis. *Digital communications*. Prentice-Hall, New York, USA, 3rd edition, 1995.

[57] R. Raheli, A. Polydoros, and C.-K. Tzou. Per-survivor processing: A general approach to MLSE in uncertain environments. *IEEE Trans. Commun.*, 43(2/3/4):354–364, Feb/Mar/Apr 1995.

[58] T. S. Rappaport. *Wireless Communications, Principles and Practice*. Prentice Hall, Upper Saddle River, NJ, USA, 1996.

[59] P. Robertson. Close–to–optimal one–shot frequency synchronization for OFDM using pilot carrieres. In *Communication Theory Mini Conference in conjunction with IEEE Globecom '97*, pages 97–102, Phoenix, AZ, USA, Nov. 1997.

[60] P. Robertson and S. Kaiser. Analysis of the effects of phase–noise in orthogonal frequency division multiplex (OFDM) systems. In *Proc. Intern. Conf. Commun.*, pages 1652–1657, Seattle, USA, 1995.

[61] M. Russel and G. Stüber. Terrestrial digital video broadcasting for mobile reception using OFDM. *Wireless Pers. Commun.*, 2(1-2):45–66, 1995.

[62] G. Santella. A frequency and symbol synchronization system for OFDM signals: Architecture and simulation results. *IEEE Trans. Vehic. Technol.*, 49(1):254–275, Jan. 2000.

[63] T. M. Schmidl and D. C. Cox. Robust frequency and timing synchronization for OFDM. *IEEE Trans. Commun.*, 45(12):1613–1621, Dec. 1997.

[64] M. Speth, S. A. Fechtel, G. Fock, and H. Meyr. Optimum receiver design for wireless broad-band systems using OFDM - Part I. *IEEE Trans. Commun.*, 47(11):1668–1677, Nov. 1999.

[65] B. Stantchev and G. Fettweis. Burst synchronization for OFDM–based cellular systems with separate signalling channel. In *Proc. IEEE Vehic. Technol. Conf.*, pages 758–762, Ottawa, Canada, May 1998.

[66] H. Steendam and M. Moeneclaey. Analysis and optimization of the performance of OFDM on frequency-selective time selective fading channels. *IEEE Trans. Commun.*, 47(12):1811–1819, Dec. 1999.

[67] A. Steingass, A. Wijngaarden, and W. Teich. Frame synchronization using superimposed sequences. In *Proc. IEEE Intern. Symp. Inf. Theory*, page 489, Ulm, Germany, June 1997.

[68] H. Takanashi and M. Nakagawa. Antenna and multi-carrier combined diversity system. *IEICE Trans. Commun.*, E79–B(9):1221–1226, Sept. 1996.

[69] F. Tufvesson. Some channel estimation strategies for a cellular OFDM system. In *Proc. RVK'96 RadioScience and Commun. Conf.*, pages 161–164, Luleå, Sweden, June 1996.

[70] F. Tufvesson. Channel related optimization of wireless communication systems. Licentiate thesis no. 3, ISSN 1402-8662, Department of Applied Electronics, Lund University, Lund, Sweden, Feb. 1998.

[71] F. Tufvesson and O. Edfors. A comparative analysis between different detection schemes in flat Rayleigh-fading channels. In *Proc. IEEE Vehic. Technol. Conf.*, Tokyo, Japan, May 2000.

[72] F. Tufvesson and O. Edfors. Preamble-based time and frequency synchronization for OFDM systems. *Submitted to IEEE J. Select. Areas Commun.*, 2000.

[73] F. Tufvesson, M. Faulkner, and O. Edfors. Time and frequency synchronization for OFDM using PN-sequence preambles. In *Proc. IEEE Vehic. Technol. Conf.*, pages 2203–2207, Amsterdam, The Netherlands, Sept. 1999.

[74] F. Tufvesson, M. Faulkner, O. Edfors, and T. Maseng. Time and frequency synchronization for BRAN using PN-sequence preambles. In *Proc. RVK'99 - RadioScience and Commun. Conf.*, pages 654–658, Karlskrona, Sweden, June 1999.

[75] F. Tufvesson, M. Faulkner, P. Hoeher, and O. Edfors. OFDM time and frequency synchronization by spread spectrum pilot technique. In *Eighth Communication Theory Mini Conference in conjunction with IEEE ICC '99*, pages 115–119, Vancouver, Canada, June 1999.

[76] F. Tufvesson, M. Faulkner, and T. Maseng. Pre-compensation for Rayleigh fading channels in time division duplex OFDM systems. In *Proc. 6th IEEE Int. Workshop on Intelligent Signal Proc. and Commun. Systems*, pages 53–57, Melbourne, Australia, Nov. 1998.

[77] F. Tufvesson, M. Faulkner, and T. Maseng. Pre-compensation for Rayleigh fading channels in time division duplex OFDM systems. *Accepted for publication in Wireless Pers. Commun.*, 2000.

[78] F. Tufvesson and P. Hoeher. Channel estimation using superimposed pilot sequences. *Submitted to IEEE Trans. Commun.*, 2000.

[79] F. Tufvesson and P. Karlsson. Some deisgn issues for OFDM HIPERLAN 2 systems based on radio channel measurements. ETSI EP BRAN doc. no. 3TRS081A, European Telecommunications Standards Institute, Sophia Antipolis, France, Apr. 1998.

[80] F. Tufvesson and P. Karlsson. Frequency stability and design of pilot patterns for channel estimation and coherent detection. ETSI EP BRAN doc. no. HL12.5TRS2a, European Telecommunications Standards Institute, Sophia Antipolis, France, Feb. 1999.

[81] F. Tufvesson and T. Maseng. On channel estimation strategies for a cellular OFDM system. In *Proc. Nordic Radio Seminar*, pages 28–31, Lund, Sweden, Aug. 1996.

[82] F. Tufvesson and T. Maseng. Pilot assisted channel estimation for OFDM in mobile cellular systems. In *Proc. IEEE Vehic. Technol. Conf.*, pages 1639–1643, Phoenix, USA, May 1997.

[83] F. Tufvesson and T. Maseng. Optimization of sub-channel bandwidth for mobile OFDM systems. In D. Everitt and M. Rumsewicz, editors, *Multiaccess, mobility and teletraffic - advances in wireless networks*, pages 103–114. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.

[84] H. S. Wang and P.-C. Chang. On verifying the first order Markovian assumption for a Rayleigh fading channel model. *IEEE Trans. Vehic. Technol.*, 45(2):353–357, May 1996.

[85] S. Waranabe, T. Sato, and T. Abe. Forward subchannel control scheme for TDD multi-carrier mobile communication system. *IEICE Trans. Fundamentals*, E82-A(7):1172–1178, July 1999.

[86] S. B. Weinstein and P. M. Ebert. Data transmission by frequency-division multiplexing using the discrete Fourier transform. *IEEE Trans. Commun.*, COM-19(5):628–634, Oct. 1971.

[87] M. D. Yacoub. *Foundations of Mobile Radio Engineering.* CRC Press, Boca Raton, FL, USA, 1993.

[88] P. Zhiyong, Y. Xiaohu, and C. Shixin. Transmission and reception of TDD multicarrier CDMA signals in mobile communications system. In *Proc. IEEE Vehic. Technol. Conf.*, pages 2134–2138, Huston, TX, USA, May 1999.

[89] W. Zhuang and W. V. Huang. Phase precoding for frequency-selective Rayleigh and Rician slowly fading channels. *IEEE Trans. Vehic. Technol.*, 46(1):129–142, Feb. 1997.

[90] W. Zhuang, W. A. Krzymein, and P. A. Goud. Adaptive channel precoding for personal communications. *Electron. Lett.*, 30(19):1570–1571, Sept. 1994.

Part II

Included Papers

Paper A

This part has been published as:

Fredrik Tufvesson and Torleiv Maseng, "Pilot Assisted Channel Estimation for OFDM in Mobile Cellular Systems", Proceedings of IEEE Vehicular Technology Conference, Phoenix, USA, pp. 1639-1643, May 1997.

Pilot Assisted Channel Estimation for OFDM in Mobile Cellular Systems

Fredrik Tufvesson and Torleiv Maseng
Department of Applied Electronics, Lund University,
Box 118, S-221 00 Lund, Sweden
E-mail:Fredrik.Tufvesson@tde.lth.se

Abstract

The use of pilot symbols for channel estimation introduces overhead and it is thus desirable to keep the number of pilot symbols to a minimum. The number of needed pilot symbols for a desired bit error rate and Doppler frequency is highly dependent on the pilot pattern used in orthogonal frequency division multiplexed, OFDM, systems. Five different pilot patterns are analysed by means of resulting bit error rate, which is derived from channel statistics. Rearrangement of the pilot pattern enables a reduction in the number of needed pilot symbols up to a factor 10, still retaining the same performance. The analysis is general and can be used for performance analysis and design of pilot patterns for any OFDM system.

1 Introduction

The mobile channel introduces multipath distortion of the signalling waveforms. Both the amplitude and phase are corrupted and the channel characteristics changes because of movements of the mobile. In order to perform coherent detection, reliable channel estimates are required. These can be obtained by occasionally transmitting known data or so called “pilot symbols”. The receiver interpolates the channel information derived from the pilots to obtain the channel estimate for the data signal, see figure 1.

OFDM, orthogonal frequency division multiplexing, is used and proposed for several broadcast systems and there is a growing interest in using the technique for the next generation land mobile communication system. In OFDM systems the information signal can be seen as divided and transmitted by several narrowband sub-carriers. Typically, for practical OFDM systems, the frequency spacing is less than the coherence bandwidth and the symbol time is less than the coherence time. This means that a receiver and pilot estimation pattern that take advantage of the relatively large coherence bandwidth and coherence time can manage with less pilot symbols,

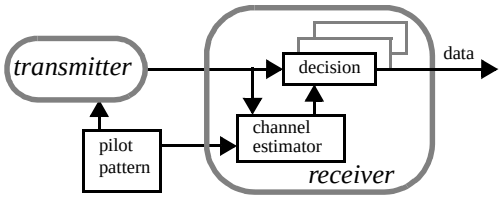

Figure 1: The pilot pattern is known both by the receiver and the transmitter. The purpose is to minimize the number of transmitted pilot symbols without increasing the bit error rate.

thereby minimizing the overhead introduced by the pilot symbols. The problem is to decide where and how often to insert pilot symbols. The spacing between the pilot symbols shall be chosen small enough to enable reliable channel estimates but large enough not to increase the overhead too much. This paper includes among all an algorithm of how to design a suitable pilot estimation pattern.

In a multicarrier system there exist a unique opportunity to determine various parts of the channel impulse response, as opposed to a single-carrier system. It is no use to make efforts to determine already known parts. Until now it seems like no one has looked into the impact of the placement of the pilot symbols for multicarrier systems. Cavers [5] made an exhaustive theoretical analysis for single-carrier systems. He pointed out that it was appropriate that 14% of the sent symbols were pilot symbols to be able to handle large Doppler values ($f_d T_s = 0.05$). Some pilot estimation patterns for OFDM has been presented in the literature, see e.g. [6], [7]. Comparisons between these and the one proposed here are shown later.

2 Estimation Strategies

Five different pilot patterns are analyzed, see figure 2.

1. Measure all channels at the same time, compare to a broadband single-carrier system.
2. Measure the channels in increasing order, one at a time.
3. Measure the channels one at a time in a smart, but predetermined, way. The measurement order is derived upon channel statistics and is optimal in the sense that the total bit error rate is kept at a minimum each symbol time.
4. A pilot pattern presented by T. Mueller *et al*. [6], where the pilots are located with equidistant spacings in time and frequency.

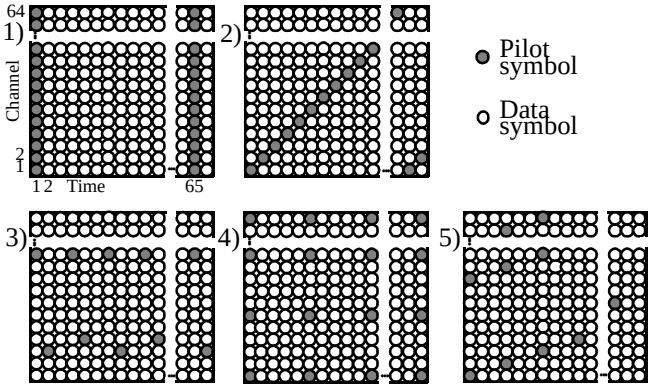

Figure 2: Examples of pilot patterns analyzed.

5. A pilot pattern by P. Hoeher [7], used in [4]. The pilot symbol locations are shifted one step in frequency each pilot interval.

For comparison the same amount of pilots is used, one of 64 sent symbols. This means that only 1.6% extra overhead is introduced by the pilots, but this is not sufficient for large Doppler frequencies in some of the cases.

3 System

At different signal to noise ratios the resulting bit error rate from each pilot pattern is evaluated using the algorithms given in Section 8. A matched filter receiver and coherent BPSK or QPSK are used. Additive white Gaussian noise is assumed for every sub-channel. The channel has delays and Rayleigh distributed amplitudes according to the COST 207: “Typical Urban profile”. The complex parts of the transfer function are assumed to change according to a first order auto-regressive process as described in Section 6. The reason for using this model is to get a linear system which rather easy can be hand-led algebraically. A Kalman filter is used to estimate the frequency response of the channel. From the filter, a time dependent covariance matrix is given as described in Section 7. This is used to calculate the expected bit error rate for each channel. See figure 3 for a description of the system.

4 Results

The resulting bit error rate curves of the pilot patterns are presented for different Doppler frequencies in figure 4.

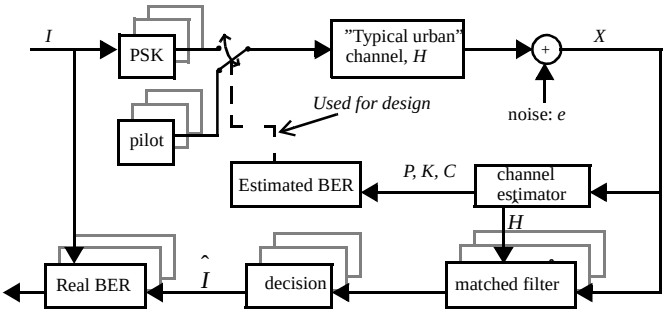

Figure 3: Overview of the system used to analyze the estimation patterns.

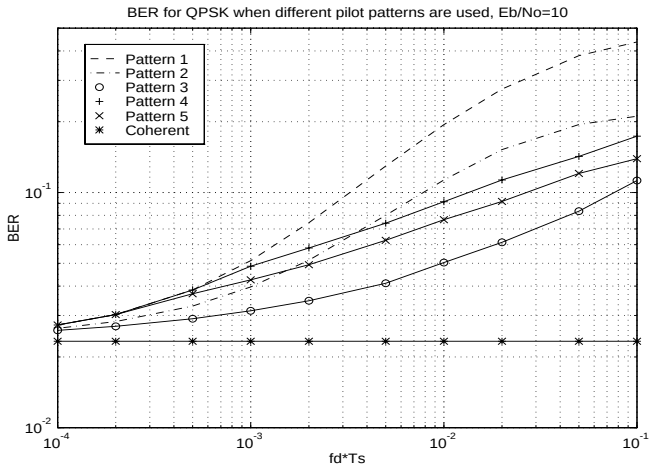

Figure 4: Resulting bit error rate as a function of Doppler value when the different pilot patterns are used. The same number of pilot symbols are used for all curves.

For a given Doppler frequency the pilot pattern used sets the limit for the lowest pilot density to be used, alternatively for a given pilot density it limits the maximum Doppler frequency allowed. The calculations are made at 1800 MHz using 10 kHz channel separation between 64 OFDM channels carrying in total 640 ksymbols/s. No intersymbol interference, perfect synchronization and known Doppler frequency, f_d, is assumed. 1.6% of the sent symbols are pilot symbols and the average bit error rate of the 64 channels is presented.

The bit error rate is degraded both by imperfect channel estimates and noise disturbances. The pilot pattern used determines the conversion between noise limited

and estimate limited region, see figure 5.

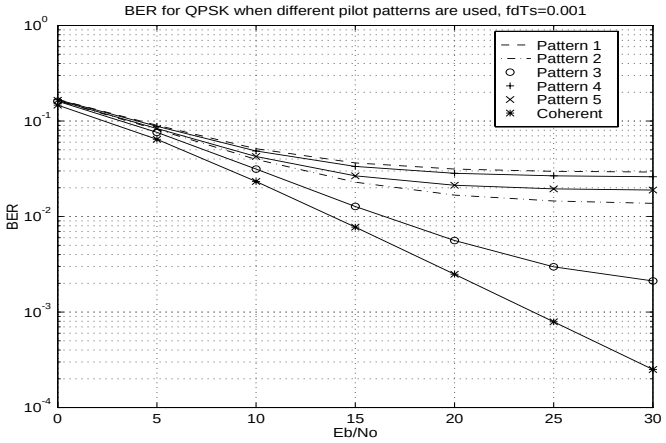

Figure 5: Bit error rate at different signal to noise ratios. Note the difference between the error floors.

It is interesting to study the resulting pattern when estimating the channel that gives the lowest bit error rate (strategy 3). A steady state pilot pattern is often achieved where only few of the sub-channels are measured, see figure 6. Channel estimates of the other sub-channels are achieved by filtering.

To see the effect of mismatch between the pilot pattern design parameters and the actual parameters, the optimal pilot patterns (pattern 3) for three Doppler values were

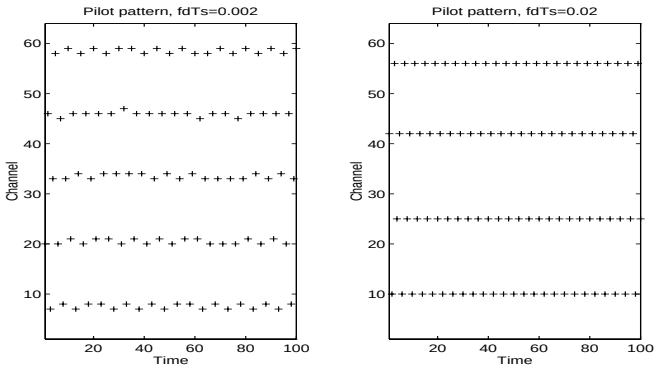

Figure 6: Resulting pilot pattern when sending a pilot at the channel which gives the lowest total bit error rate. Note that only few channels will be used for pilot symbols.

used when moving at another speed. The actual Doppler frequency was as before assumed to be known by the estimator, just to see the effect of the pilot pattern without influences of the estimation algorithm. The designed pilot patterns for the ”typical urban” environment were also used in ”hilly terrain” and ”rural area” specified in [2] to see the influence of the propagation environment on the bit error rate, see figure 7. As seen in the figure, the pilot patterns are robust to mismatches in the design parameters.

5 Discussion

To minimize the bit error rate it is desirable to spread the pilot symbols both in time and in frequency, as seen in figure 4 and figure 5. Normally a worst case design is made for the channel estimation system and then we suggest to tailor the pilot pattern to each base station site. A suitable pilot pattern can be calculated once the propagation environment and maximum expected speed in the particular cell is known. In such a system, the pilot pattern used in the cell is among the information transferred to the mobile when it logs into a new cell. When designing the pilot pattern one also has to take the estimation algorithm into account. Here the estimation algorithm is used only for evaluation and the complexity of the used algorithm is not a problem. In some cases the pilot pattern has to aid the estimation algorithm to enable a less complex one. The estimation algorithm used here, the Kalman filter based on an AR-process, has no delay and the received signal can be detected immediately, i.e. no future pilot symbols is taken into account when making the channel estimate. If the

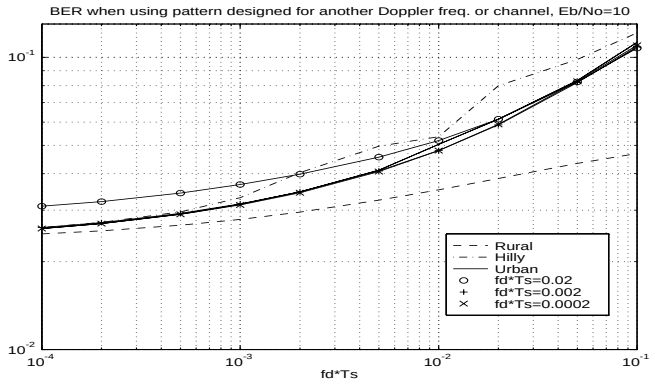

Figure 7: Changes in the bit error rate due to Doppler mismatch and power delay profile mismatch. The pilot patterns designed for f_dT_s=*0.02, 0.002* and *0.0002* were used at different Doppler values and the designed patterns for typical urban environment were used in hilly terrain and rural area.

received signal is stored in a memory, the pilot symbols can be used in both ”time directions” and the time spacings between pilot symbols can be increased, retaining the same performance.

The degradation due to mismatch in design parameters is mainly caused by the estimation algorithm and therefore the curves for different Doppler values do not differ much. When the pilot pattern is designed for higher Doppler values than the actual one, an increased error rate is achieved since the pilot symbols are not located as close to each other in frequency as desired. In rural area the bit error rate becomes lower due to the increased frequency correlation while the opposite happens in hilly terrain. In the first case, an even better result is achieved with less pilots along the frequency axes and more along the time axes.

The bit error rates within the sub-channels differ depending on where the pilots are located. When minimizing the total bit error rate (pattern 3) the channels of the sides get higher error rate, see example in figure 8.

6 Channel Model

The time dependent impulse response, $h(\tau, t)$, is assumed to be a sum of reflections, see (1) where $\delta(t)$ denotes the dirac-function.

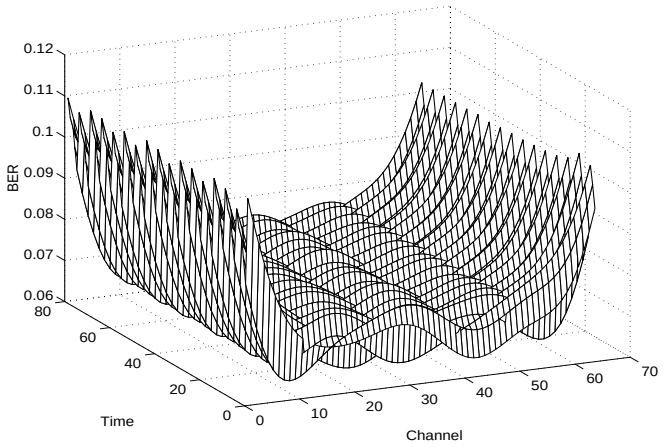

Figure 8: The bit error rate becomes higher for the side channels when the total bit error rate is kept at a minimum

$$h(\tau, t) = \sum_{n=1}^{N} a_n(t) \cdot \delta(t - \tau_n) \quad (1)$$

The tap coefficients, $a_n(t)$, and the tap delays, τ_n, are chosen according to the COST 207 ”Typical Urban” model in the GSM specification [2]. The transfer functions, $H(f,t)$, are obtained by the Fourier transform and these are the functions we want to estimate for the different carriers. These channel transfer functions are regarded as flat fading and constant during a symbol time.
A first order AR-process is used to model how the different taps may change from one time instant to another. If we look at all the transfer functions at the same time, it is possible to set up a state-space model of the form:

$$\begin{aligned} H(k+1) &= \phi H(k) + v(k) \\ y(k) &= C(k)H(k) + e(k) \end{aligned} \quad (2)$$

The matrix ϕ is a diagonal $N*N$-matrix (here treated as a scalar) with elements

$$e^{-k_{AR} 2\pi f_d T_s} \quad (3)$$

that define the AR-process. T_s is the symbol time including any cyclic prefix or guard space. The white noise $v(k)$ has covariance matrix $R_1 = F \cdot R_{GSM} \cdot F^*$, where R_{GSM} corresponds to the multipath intensity profile described in [2]. The vector $y(k)$ is the measured transfer functions. $C(k)$ is an observation vector with ones only at the positions (channels) measured at time KT_s and $e(k)$ is measurement noise with a diagonal covariance matrix R_2.

The parameter k_{AR} in the auto-regressive process for the channel changes is chosen to adjust the ”memory” so that it corresponds to the ”memory” of Jakes’ model. Channels corresponding to the U-shaped spectrum given in [8] were simulated and then estimators based on an AR-model with different k_{AR} were used, see Figure 9: Simulations were performed with one sub-channel ($E_b/N_o = 10$, $f_dT_s = 0.002$) with every tenth symbol as a pilot symbol. In the figure the bit error rates of the nine data symbols are shown. figure 9 shows that the best fit, in this case, is reached for $k_{AR} \approx 0.15$.

The adjustment of the AR-process can also be seen as an adjustment of the bandwidth. If we set the 90% power bandwidth of the AR process equal to the bandwidth of the Doppler spread, see figure 10, a value of $k_{AR} \approx 0.158$ can be calculated. This is the value used for all bit error rate calculations throughout the paper.

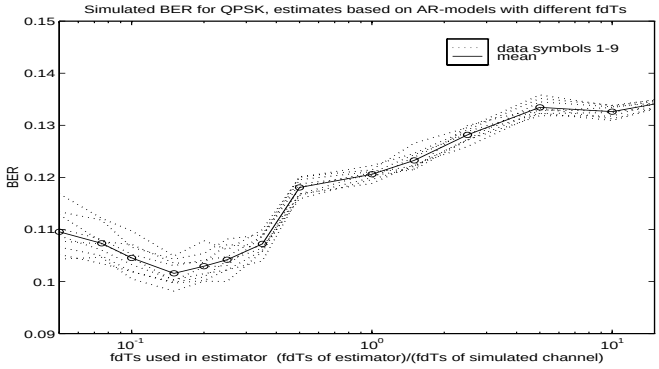

Figure 9: Bit error rate for a simulated channel by Jakes when estimates are based on an AR-process. The minimum value is reached for $k_{AR}=0.15$.

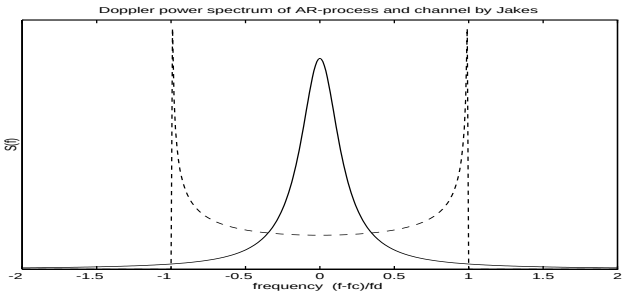

Figure 10: Power density spectrum of the channel by Jakes and a first order AR-process. The latter is adjusted to have the 90% power bandwidth equal to the Doppler spread.

7 Channel Estimator

For the analysis and pilot pattern design a Kalman filter is used to estimate the transfer functions. The Kalman filter is causal and uses measurements up to time k to estimate the transfer functions, $H(k)$. The Kalman filter is given [3] by (4)-(8), where X^* denotes conjugate transpose of X:

$$\hat{H}(k|k) = \hat{H}(k|k-1) + P(k|k-1)C(k)^* \cdot \{C(k)P(k|k-1)C(k)^* + R_2\}^{-1}(y(k) - C(k)\hat{H}(k|k-1)) \quad (4)$$

$$\hat{H}(k+1|k) = \phi\hat{H}(k|k-1) + K(k)[y(k) - C(k)\hat{H}(k|k-1)] = \phi\hat{H}(k|k) \quad (5)$$

$$K(k) = \phi P(k|k-1)C(k)^* \cdot [C(k)P(k|k-1)C(k)^* + R_2]^{-1} \quad (6)$$

$$P(k+1|k) = \phi P(k|k-1)\phi^* + R_1 - K(k)[C(k)P(k|k-1)C(k)^* + R_2]K(k)^* \quad (7)$$

$$P(k|k) = P(k|k-1) - P(k|k-1)C(k)^* \cdot [C(k)P(k|k-1)C(k)^* + R_2]^{-1}C(k)P(k|k-1) \quad (8)$$

The reconstruction error $\tilde{H}(k|k) = H(k) - \hat{H}(k|k)$ is given by:

$$\tilde{H}(k|k) = [\phi - K(k)C(k)]\tilde{H}(k-1|k-1) + v(k-1) - P(k|k-1)C(k)^*[C(k)P(k|k-1)C(k)^* + R_2]^{-1} \cdot \{e(k) + C(k)v(k-1)\} \quad (9)$$

The Kalman filter is optimal in the sense that the variance of the reconstruction error is minimized. The matrix *P(k|k)* is the variance matrix and this is used together with *K* to make an estimate of the bit error rate, which in turn is used to decide the order in which the channels are going to be measured. For pattern 3, the channel is chosen that minimizes the total bit error rate of the channels after the measurement. The matrices *P(k|k)* and *K* are independent of the measured values and therefore it is possible to precompute the order in which the channels are going to be measured.

8 Bit Error Rate Calculations

The bit error rate is calculated for BPSK and QPSK. A matched filter receiver is assumed. The sampled output of this filter is given by

$$X_m = 2\varepsilon H_m e^{j\frac{2\pi}{M}(n-1)} + e_m \quad (10)$$

where ε is the signal energy, H_m is the current transfer function at channel m, n is the signal alternative among M sent and e_m is white gaussian noise. The bit error rate at channel m is calculated as [1]

$$P_{bmBPSK} = \frac{1}{2}(1-\mu_m)$$
$$P_{bmQPSK} = \frac{1}{2}\left(1-\frac{\mu_m}{\sqrt{2-\mu_m^2}}\right) \quad (11)$$

where

$$\mu_m = \frac{E[X_m \hat{H}_m^*]}{\sqrt{E[|X_m|^2]E[|\hat{H}_m|^2]}} \quad (12)$$

$\hat{H}_m$ are the outputs of the Kalman filter. These are not known in advance and therefore (13) -(15) are used. If matrix notation is used and signal alternative 1 is used for the pilot symbol (does affect the analysis here, but in practise different pilot symbol alternatives should be used), the expectations for all the channels can be calculated as:

$$E[X(k)\hat{H}(k)^*] = E[\{2\varepsilon H(k)+N(k)\}\{H(k)-\tilde{H}(k)\}^*] \quad (13)$$
$$= 2\varepsilon E[H(k)H(k)^*] - 2\varepsilon E[H(k)\tilde{H}(k)^*]$$

$$E[|X(k)|^2] = E[\{2\varepsilon H(k)+N(k)\}\{2\varepsilon H(k)+N(k)\}^*] \quad (14)$$
$$= 4\varepsilon^2 E[H(k)H(k)^*] + E[N(k)N(k)^*]$$

$$E[|\hat{H}(k)|^2] = E[\{H(k)-\tilde{H}(k)\}\{H(k)-\tilde{H}(k)\}^*] \quad (15)$$
$$= E[H(k)H(k)^*] - 2E[H(k)\tilde{H}(k)^*] + E[\tilde{H}(k)\tilde{H}(k|k)^*]$$

The expected values $E[X(k)\hat{H}(k)^*]$, $E[|X(k)|^2]$ and $E[|\hat{H}(k)|^2]$ are independent of the measured values but dependent upon which channels that are measured. Therefore they are time dependent. The expectations can be calculated as:

$$E[H(k)H(k)^*] \quad (16)$$
$$= E[\{\phi H(k-1)+v(k-1)\}\{\phi H(k-1)+v(k-1)\}^*]$$
$$= \phi^2 E[H(k-1)H(k-1)^*] + E[v(k-1)v(k-1)^*]$$
$$= R_1/(1-\phi^2)$$

$$E[N(k)N(k)^*] = R_2 \quad (17)$$

$$
\begin{aligned}
&E[H(k)\tilde{H}(k|k)^*] = E[H(k)H(k)^* - H(k)\hat{H}(k|k)^*] = \\
&E[\phi H(k-1)\tilde{H}(k-1)^*\phi^* + R_1 - R_1 C(k)^* a^* \\
&- \phi H(k-1)\tilde{H}(k-1)^*\phi^* C(k)^* a^*] \\
&= (\phi E[H(k-1)\tilde{H}(k-1)^*]\phi^* + R_1)(I - C(k)^* a^*)
\end{aligned}
\tag{18}
$$

where

$$
E[\tilde{H}(k|k)\tilde{H}(k|k)^*] = P(k|k) \tag{19}
$$

$$
a = P(k|k-1)C(k)^*\{C(k)P(k|k-1)C(k)^* + R_2\}^{-1} \tag{20}
$$

The bit error rate calculation is compared and verified by simulations. Proakis [1] gives expressions for the bit error rate when estimating a constant Rayleigh channel using different numbers of pilot symbols, see figure 11.

The bit error rate when using only one pilot symbol corresponds to estimating a rapidly changing channel or the first estimate of an unknown channel. Then, only the latest measurement is useful. In a similar way, the bit error rate when estimating a constant channel with use of (infinitely) many pilot symbols corresponds to that of coherent detection. Normally in a practical system the effect of the estimation is somewhere between these two cases.

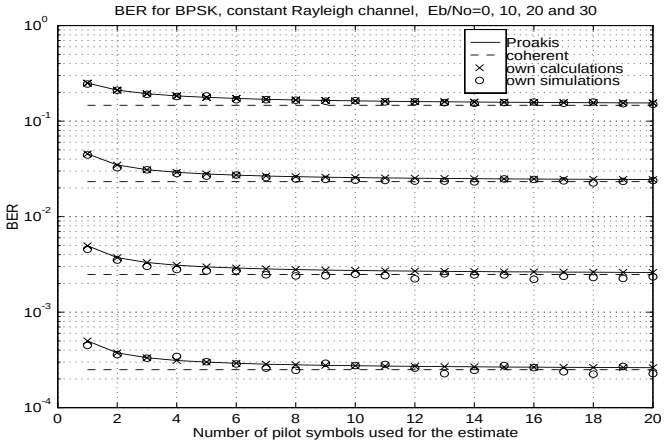

Figure 11: Bit error rate for 2PSK when estimating a constant Rayleigh channel with several pilot symbols at different signal to noise ratios.

9 Conclusions

The bit error rate for pilot assisted QPSK modulation is calculated when using different pilot patterns. The ability to estimate the channel reliably when it changes due to e.g. vehicle movements is highly dependent on the pilot pattern used. By rearranging the pilot pattern it is, in some cases, possible to handle 10 times higher Doppler frequency alternatively possible to reduce the number of needed pilot symbols the same amount, still retaining the same bit error rate. Alternatively the new pilot pattern could be used to reduce the bit error rate up to a factor 5, even more in a low noise environment. The pilot symbols in the proposed pilot pattern are spread out both in time and frequency. For a given propagation environment, e.g. a base station site, it is possible to pre-calculate a suitable pilot pattern.

References

[1] J. G. Proakis, Digital Communications, third edition, McGraw-Hill, USA, 1995.

[2] European Telecommunications Standard Institute, GSM 05.05, version 4.6.0, July 1993.

[3] J. M. Mendel, Lessons in Estimation Theory for Signal Processing, Communications and Control, Prentice-Hall, USA, 1995.

[4] M. Sandell, Design and Analysis of Estimators for Multicarrier Modulation and Ultrasonic Imaging, Ph.D. thesis, Luleå University of Technology, Sweden, 1996.

[5] J. K. Cavers, An Analysis of Pilot Symbol Assisted Modulation for Rayleigh Fading Channels, IEEE Trans. Vehic. Tech. vol 40, no 4, pp 686-693, nov 1991.

[6] T. Mueller, K. Brueninghaus and H. Rohling, Performance of Coherent OFDM-CDMA for Broadband Mobile Communications, Wireless Personal Communications 2, pp 295-305, Kluwer, Netherlands, 1996.

[7] P. Hoeher. TCM on Frequency-Selective Land-Mobile Fading Channels, Proc. of 5th Tirrenia Int. Workshop in Dig. Communications, Tirrenia, Italy, Sept. 1991.

[8] W. Jakes, Microwave Mobile Communications, Wiley-Interscience, USA 1974

Paper B

This part has been published as:

Fredrik Tufvesson and Torleiv Maseng, "Optimization of Sub-Channel Bandwidth for Mobile OFDM Systems", in D. Everitt and M. Rumsewicz, editors, Multiaccess, mobility and teletraffic - advances in wireless networks, pp. 103-114, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.

Optimization of Sub-Channel Bandwidth for Mobile OFDM Systems

Fredrik Tufvesson and Torleiv Maseng
Department of Applied Electronics, Lund University,
Box 118, SE-221 00 Lund, Sweden.
E-mail: Fredrik.Tufvesson@tde.lth.se

Abstract

The bit error rate in orthogonal frequency division multiplex (OFDM) systems is affected by the number of sub-channels used, changes in the channel characteristics and, to some extent, the excess delay of the channel. This paper presents an analytical expression for the bit error rate on Rayleigh fading channels when interchannel interference (ICI) caused by channel changes during a symbol and energy loss due to the cyclic prefix are regarded. This expression is used to optimize the number of sub-channels, and thereby the sub-channel bandwidth in the system. It is argued that the system can be optimized neglecting the effect of imperfect channel estimation and on a worst case assumption for the Doppler frequency and signal to noise ratio. The analysis is general and can be used for performance analysis and optimization of any mobile OFDM system.

1 Introduction

OFDM, orthogonal frequency division multiplexing, is used and proposed for several broadcast systems [1] [2] and there is a growing interest in using the technique for the next generation land mobile communication system. In OFDM systems the information signal can be seen as divided among and transmitted by several narrowband subcarriers. The bandwidths of the sub-carriers depend on the bit rate sent on each of them and consequently, for a given total bitrate the sub-carrier bandwidth is dependent on the number of sub-channels used. The number of sub-channels used is set in the system design and the problem is to find a good trade off between bandwidth, limitations by the hardware and the physical channel. As the sub-carrier bandwidth is reduced, the symbol time on each carrier gets longer, the channel changes during a symbol gets larger and channel compensation gets more difficult. When the changes in the channel characteristics during a symbol time become evident, the orthogonality between the subchannels is lost and interchannel interference (ICI) arises. On the other hand as the carrier bandwidth and the sub-channel symbol rate increases, intersymbol interference (ISI) becomes a problem. This problem can be avoided by

introducing a cyclic prefix [3], but this results in an energy loss. In between these two extremes there exists an optimal bandwidth where the bit error rate is minimized, this bandwidth is found in this paper.

There is some previous work regarding the choice of number of sub-channels in OFDM systems, see e.g. [4] [5], but often the optimization is performed empirically or by simulation. In [4] a noncoherent OFDM DPSK system is analysed with respect to random FM noise and a frequency selective channel. It is concluded that a sub-channel bandwidth of 1.5 kHz is suitable for a bitrate of 200 kbit/s and 100 Hz Doppler frequency, but without consideration to the interchannel interference caused by channel changes during a symbol. The latter can cause severe degradation of the bit error rate if the sub-channel bandwidth is chosen too small.

2 System description

When designing an OFDM system one often start with a requirement on the total bit rate. Assume that a total bitrate of R_{tot} bits per second is required. If these are equally divided between the channels then the bitrate

$$R_1 = \frac{R_{tot}}{N} \tag{1}$$

is transferred on each of the N sub-channels. The symbol time for M-ary modulation on each sub-channel becomes

$$T_{sub} = \frac{\log_2 M}{R_1} = \frac{N \log_2 M}{R_{tot}}. \tag{2}$$

In the following the influence of some of the most important design parameters are discussed.

2.1 Cyclic prefix

To avoid intersymbol interference and to maintain orthogonality between the sub-channels in a time dispersive channel, a cyclic prefix can be used [3]. ISI is avoided if the length of the cyclic prefix, T_{cp}, is chosen large enough to exceed the maximum excess delay of the channel. The total symbol time, T_{sub}, is extended when a cyclic prefix is used but the sub-channel bandwidth, Δf, is equal to the inverse of the symbol time excluding the cyclic prefix [4], T_s, see figure 1.

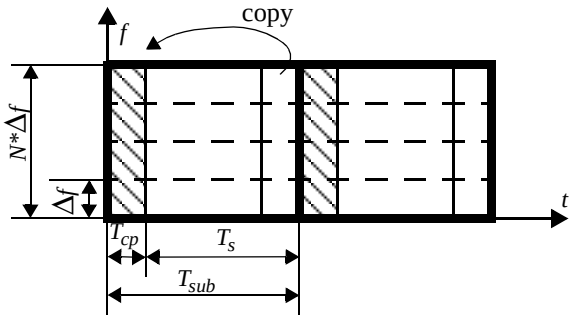

Figure 1: Eight symbols sent on four sub-channels. The striped area indicates insertion of the cyclic prefix.

$$\Delta f = \frac{1}{T_s} = \frac{1}{T_{sub} - T_{cp}} = \frac{R_{tot}}{N\log_2 M - T_{cp}R_{tot}} \tag{3}$$

Insertion of a cyclic prefix means that the sub-channel bandwidth has to be increased in order to keep the bit rate R_1 constant. The bit rate determines the symbol time T_{sub} and if a part of this is used for the cyclic prefix, then the time T_s has to be shortened which in turn leads to increased sub-channel bandwidth.

The cyclic prefix also leads to a power loss, α_{cp}. The receiver uses the energy received during the time T_s and discards the energy corresponding to the duration of the cyclic prefix. The remaining signal energy can be calculated as $\alpha_{cp}E_s$ where

$$\alpha_{cp} = \frac{T_s}{T_{sub}} = 1 - \frac{T_{cp}R_{tot}}{N\log_2 M}. \tag{4}$$

2.2 Interchannel interference

Even though the channel is perfectly estimated there are some variations in the transfer function during each symbol interval. These variations becomes evident when many sub-channels are used due to the long symbol time, they are hard to track and result in interchannel interference, ICI. When the number of sub-channels is sufficiently large the resulting ICI can be modeled as additive white Gaussian noise, which is added to the channel noise [6] with spectral density N_0. The variance of the ICI-noise, $\sigma^2_{ICI} = E[n_{ICI}n_{ICI}^*]$, is calculated as[6]:

$$\sigma_{ICI}^2 = 2E_s\left(1-\frac{1}{N}-\frac{2}{N^2}\sum_{n=1}^{N-1}(N-n)J_0\left[2\pi f_D n\left(\frac{\log_2 M}{R_{tot}}-\frac{T_{cp}}{N}\right)\right]\right) \quad (5)$$

2.3 Bit error rate for coherent detection

If we neglect the effect of channel estimation errors, the equivalent signal to noise ratio after the cyclic prefix is removed and the ICI-noise is added becomes

$$\left(\frac{E_s}{N_0}\right)_{eq} = \frac{E_s\alpha_{cp}}{N_0+\sigma_{ICI}^2}. \quad (6)$$

The BER for QPSK/4QAM in a Rayleigh fading channel with coherent detection can then be calculated as

$$p_b = \int_{y=0}^{\infty} p(y)Q\left(\sqrt{y\left(\frac{E_s}{N_0}\right)_{eq}}\right)dy = \frac{1}{2} \quad (7)$$

$$-\frac{1}{2}\sqrt{\frac{NE_s(2N-T_{cp}R_{tot})}{N^2(4N_0+10E_s)-NE_s(8+T_{cp}R_{tot})-16E_s\sum_{n=1}^{N-1}(N-n)J_0\left[2\pi f_D n\left(\frac{\log_2 M}{R_{tot}}-\frac{T_{cp}}{N}\right)\right]}}$$

where y is the power attenuation of the channel with an exponential distribution of mean 1, i.e. $p(y)=e^{-y}$. In figure 2 the bit error rate for coherent detection is presented for four different E_s/N_0 at 50 Hz Doppler frequency and a total bit rate of 320, 640 respectively 1280 kb/s. The reference curve represents coherent detection in a Rayleigh channel without losses due to cyclic prefix or ICI.

A cyclic prefix of duration 10 µs is chosen to combat time dispersion, a duration long enough to exceed the expected maximum excess delay of the channel. In this paper a COST 207 ”Typical Urban” channel [8] is considered with an maximum excess delay of 5 µs. In figure 2 the bit error rate is shown insensitive to bitrate changes but the effect of the ICI-noise and the energy loss due to the cyclic prefix is apparent. On the right hand side few sub-channels are used, the symbol time is short and the length of the cyclic prefix needed to suppress ISI causes a big loss of symbol energy. On the left hand side the symbol time is long, the changes in the channel characteristics during a symbol becomes significant and ICI arises. The optimal sub-channel bandwidth is sensitive to the Doppler frequency since a large Doppler frequency results in large ICI-noise and the optimum is pushed toward broader channels, see figure 3.

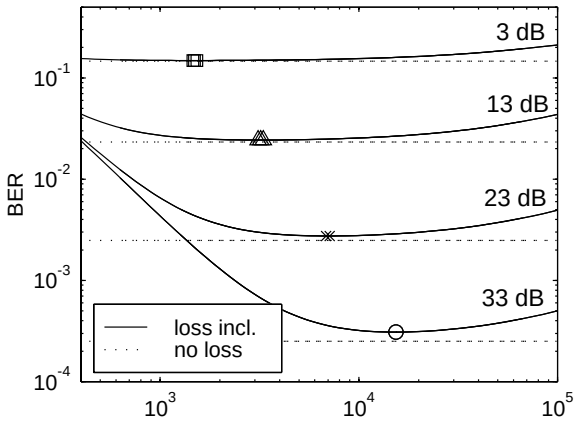

Figure 2: The BER curves for R_{tot}=320, 640 and 1280 kbit/s overlap each other totally in this case. Coherent QPSK is used in a Rayleigh fading channel with T_{cp}=10 µs and f_D=50 Hz when the number of sub-channels and signal to noise ratio are varied. The dotted lines represent coherent detection without losses.

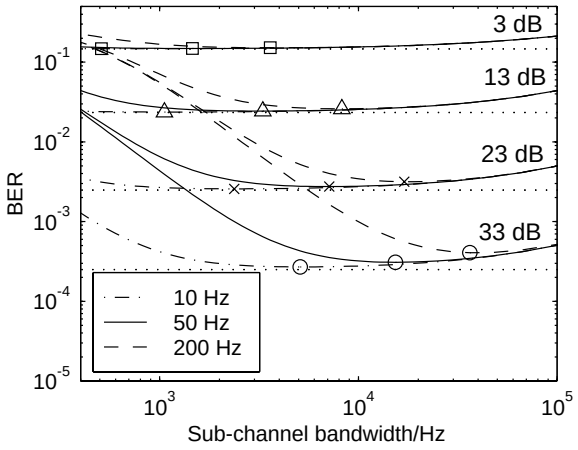

Figure 3: Bit error rate for coherent detection when the Doppler frequency is varied, f_D={10, 50, 200} Hz, R_{tot}=320 kbit/s and E_s/N_0= {3, 13, 23, 33} dB. The dotted lines represent coherent detection without losses.

2.4 Channel estimation

The radio channel corrupts the transmitted signal and in order to make coherent detection possible we have to know the impact of the channel. Both the amplitude and phase are corrupted by the fading channel, whose characteristics vary because of movements of the mobile terminal. In order to keep track of the channel characteristics pilot symbol assisted modulation, PSAM, can be used. This means that known training symbols are multiplexed into the data stream at certain sub-channels and certain times. The receiver interpolates the channel information derived from the pilot symbols and makes channel estimates for the data symbols. Since the pilot symbols carry no data the pilot density is to be kept at a minimum not to increase the overhead too much. In order to make it possible for the receiver to achieve nearly maximal channel information from each of the pilots the pilot pattern can be made ”balanced” [7]. This means that the spacings between the pilots are approximately the same in both frequency and time when normalized by the minimum expected coherence bandwidth and the minimum expected coherence time respectively. Few sub-channels means that we have to insert several data symbols between the pilot symbols in the time domain. Many sub-channels gives us the opportunity to use several of the channels for pilots simultaneously without increasing the pilot density, see figure 4.

Let N_T and N_F denote the pilot time distance respectively pilot frequency distance in number of symbols, f_{Dmax} denote the maximum expected Doppler frequency (Hz) and τ_{max} denote the maximum expected excess delay (s) of the channel, then the condition for ”balanced design” becomes [7]

$$f_{Dmax} \cdot T_{sub} \cdot N_T \approx \tau_{max} \cdot \Delta f \cdot N_F \ . \tag{8}$$

The overall pilot density is $\beta \approx \frac{1}{N_T N_F}$, and as a rule of thumb the pilot spacings can be chosen as [7]

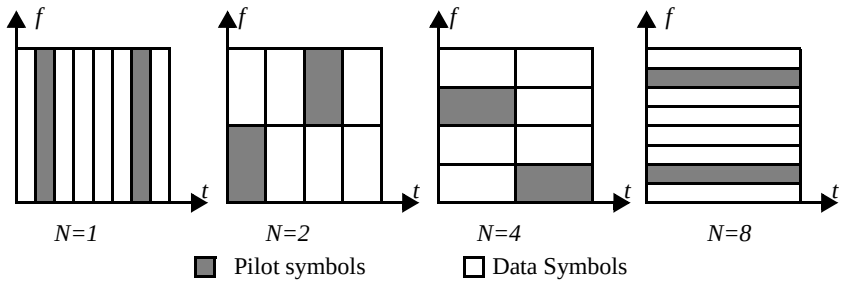

Figure 4: Time - frequency distribution of a pilot symbol differs depending on the number of sub-channels, *N*, used. Note that the same pilot density, 1/4, is used for the three alternatives.

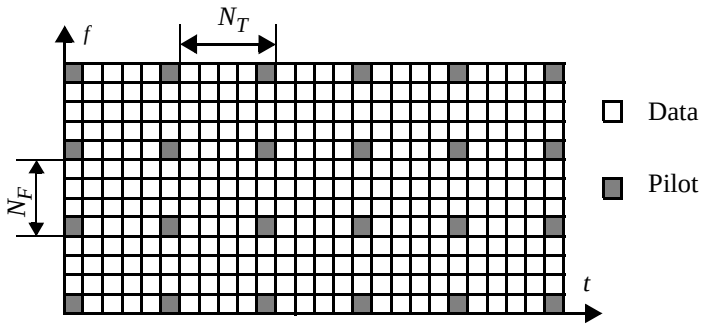

Figure 5: An example of a ”balanced” pilot pattern with pilot time distance N_T=5 and pilot frequency distance N_F=4.

$$f_{Dmax} \cdot T_{sub} \cdot N_T = \tau_{max} \cdot \Delta f \cdot N_F = \frac{1}{4} \tag{9}$$

An example of a pilot pattern with N_F=4 and N_T=5 is given in figure 5.

With a fixed time-frequency block, edge problems can arise when designing the pilot pattern. In order to get rid of this problem in the following and in order to suppress the bit error rate at the edges, pilots are put on the outer sub-channels respectively time instants and then sufficiently many pilot symbols are evenly distributed between these edge pilots in frequency respectively time in order not to exceed the pilot distances given by (9).

To see the effect of the channel estimation only, the bit error rate was calculated without ICI-noise and energy losses due to the cyclic prefix. For the estimation a two-dimensional Wiener filter [7] was used, which is optimal in the sense that the variance of the estimation error is minimized. Furthermore, known Doppler frequency, perfect synchronization, a COST 207 ”Typical Urban” channel [8] and a matched filter receiver was assumed, see Appendix A for more details. The resulting bit error rate under these assumptions with a bit rate of 320 kbit/s, 200 Hz Doppler frequency is presented infigure 6.

The reason why the different pilot patterns virtually do not affect the bit error rate is explained by the fact that the pilot symbols are placed at approximately the same frequencies and time instances independent of the sub-channel bandwidth and symbol time used. There are some changes in the bit error rates due to the fact that the pilot distances can not be below one, due to numerical reasons in the pilot pattern design and due to the slight increase in the total bandwidth to estimate when few channels are used, but these changes are rather small.

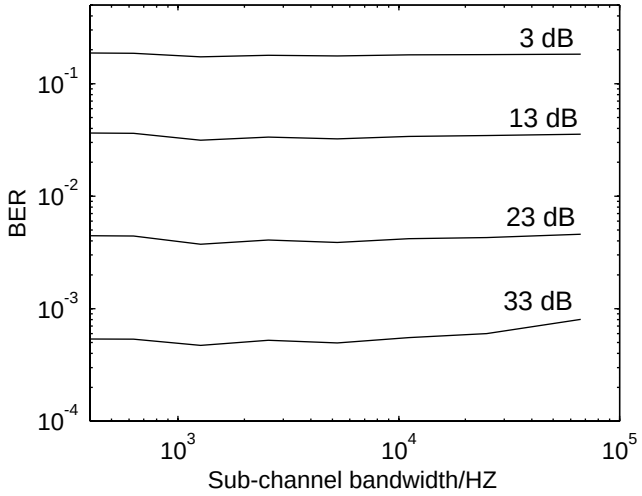

Figure 6: Changes in the bit error rate caused by the pilot pattern, E_s/N_0={3, 13, 23, 33} dB. As seen, the pilot pattern has virtually no effect on the bit error rate when ”balanced design” is used.

3 Complete System

In a real system the channel estimation is of course not perfect. The bit error rate when both channel estimation and the losses due to the cyclic prefix and ICI-noise are included is lower bounded by the bit error rate for coherent detection in (7) and the deviation gets smaller as the pilot density increases. The bit error rate for the same parameters as in last section (R_{tot}=320 kbit/s f_D=200 Hz, ”Typical Urban” channel, T_{cp}=10 µs) but with the losses included is presented in figure 7.

4 Discussion

As indicated in figure 7 it is possible to optimize the number of sub-channels and thereby the sub-channel bandwidth by assuming perfect channel estimates. The resulting optimum when channel estimation effects are included is principally unaffected by the number of sub-channels used. The reason for this is that the variance of the estimation error, see Appendix A, is approximately the same, independent of the sub-channel values chosen. This, in turn, means that it is possible to optimize the number of channels neglecting the effect of channel estimation.

More narrowband sub-channels can be used when terminal movements are slow

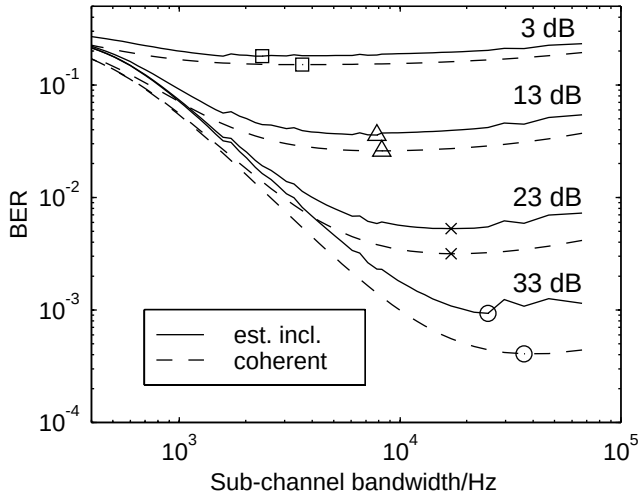

Figure 7: Bit error rate for coherent detection and for the total system when channel estimation errors are included, E_s/N_0={3, 13, 23, 33} dB.

since the channel is not expected to change during the longer symbol time. This means that the relative duration of the cyclic prefix becomes shorter and that the bandwidth expansion is kept small. However, the sub-channel bandwidth has to be optimized on a worst case assumption for the Doppler frequency, otherwise the increase in bit error rate can become severe for high Doppler values.

The optimal sub-channel bandwidth is unfortunately dependent on the signal to noise ratio. Higher noise levels call for more narrowband sub-channel since the noise makes the energy loss due to the cyclic prefix more evident. In the same way as for the Doppler frequency a worst case design has to be made with respect to the noise level. It is important to keep the maximal bit error rate as low as possible, even though this results in a larger BER increase for lower noise levels. Actually, often it is the performance of a coded system that is interesting and since the coding gives more distinct optima it is important not to increase the worst case BER.

The effect of frequency offsets is not included in the analysis and this effect becomes more significant when many narrowband sub-channels are used. The carrier to interference ratio due to frequency offsets is proportional to Δf^2 [9], thus moving the optimal sub-channel bandwidth towards broader channels when frequency offsets are present. Finally, the calculations are based on the assumption that the sub-channels are flat fading, which may not be the case when broadband sub-channels are used in some channels, e.g. hilly terrain.

5 Conclusions

The impact of the number of sub-channels, and thereby the sub-channel bandwidth, used in a mobile OFDM-system was analysed. An analytical expression for the bit error rate in Rayleigh fading channels was presented where the impact of the cyclic prefix and the channel changes during a symbol time are encountered. This BER showed an optimum for the sub-channel bandwidth to be used depending on Doppler frequency, length of cyclic prefix and signal to noise ratio. The effect of channel estimation on the optimum was also examined and this showed that the optimization can be performed on the coherent system, without taking the channel estimation into account. The excess delay of the channel, high noise levels and efficient usage of the spectrum call for many narrowband sub-channels while a rapidly moving terminal call for the use of fewer more broadband sub-channels. Finally it was argued that the sub-channel bandwidth has to be designed on a worst case assumption for the Doppler frequency and the signal to noise ratio.

Appendix A. System Models

In the following section the channel model, estimation algorithm and bit error rate calculations are presented in detail.

The time dependent channel impulse response, $h(\tau, t)$, is assumed to be a sum of reflections, see (A.1) where $\delta(t)$ denotes the dirac-function.

$$h(\tau, t) = \sum_{n=1}^{N} a_n(t) \cdot \delta(t - \tau_n) \tag{A.1}$$

The tap coefficients, $a_n(t)$, and the tap delays, τ_n, are chosen according to the COST 207 ”Typical Urban” model in the GSM specification [8]. The transfer functions, $H(f,t)$, are obtained by the Fourier transform of (A.1) and these are the functions we want to estimate for the different carriers. These channel transfer functions are regarded as flat fading.

For the estimation and analysis a two-dimensional Wiener filter [7] is used, which is optimal in the sense that the variance of the reconstruction error is minimized. The sampled signal from the matched filter receiver of a pilot symbol at channel m in time k is given by

$$X_m^k = E_s H_m^k e^{j\frac{2\pi}{M}(p-1)} + n_m^k \tag{A.2}$$

where E_s is the signal energy, p is the signal alternative sent among the M possible ones and n is white Gaussian noise. The noise terms have a variance given by (7), but the covariance matrix $R_e^k = E[n^k(n^k)^*]$ is not diagonal due to the correlation between the channels introduced by the ICI. The filter coefficients are given by

$$\omega_m^k = ((\theta_m^k)^T \phi^{-1})^T \tag{A.3}$$

where $\phi = E[XX^*]$ is the covariance matrix for the received pilot symbols and $\theta_m^k = E[H_m^k(X_m^k)^*]$ is a cross-covariance vector between the transfer function to be estimated and the received pilot symbols. The estimates of the channel transfer functions can then be calculated as

$$\hat{H}_m^k = (\omega_m^k)^T X \tag{A.4}$$

where X is a vector of the sampled values from all of the pilot symbols.

Finally, the mean square error of the estimates is calculated as

$$J_m^k = \sigma_H^2 - (\theta_m^k)^T \phi^{-1} (\theta_m^k)^* \tag{A.5}$$

and this matrix is in turn used to calculate the QPSK bit error rate for each of the symbols. The BER is given by [10] [11]

$$P_b = \frac{1}{2} - \frac{\mu}{2\sqrt{2-\mu^2}} \tag{A.6}$$

where

$$\mu = E_s \frac{\sigma_H^2}{\sqrt{(\sigma_H^2 + \sigma_e^2)(\sigma_H^2 + J)}}. \tag{A.7}$$

References

[1] ”Digital broadcasting systems for television, sound and data services”. European telecommunications standards institute, prETS 300 744, Valbonne, France, 1996

[2] ”Radio broadcasting systems; Digital Audio Broadcasting (DAB) to mobile, portable and fixed receivers. European telecommunications standards institute, ETS 300 401, Valbonne, France, 1995

[3] A. Peled and A. Ruiz, ”Frequency domain data transmission using reduced computational complexity algorithms”, Proc. IEEE Int. Conference Acoustics, speech, signal processing, Denver, USA, pp 964-967, 1980

[4] S. Hara, M. Mouri, M. Okada and N. Morinaga, ”Transmission performance analysis of multi-carrier modulation in frequency selective fast Rayleigh fading channel”, Wireless Personal Communications, nr 2, Kluwer Academic Publishers, Dordrecht, Netherlands, pp 335-356, 1996

[5] M. Wahlqvist, C. Östberg, J.-J. van de Beek, O. Edfors and P. O. Börjesson, ”A conceptual study of OFDM-based multiple access schemes: Part 1 - Air interface requirements. Technical report Tdoc 116/96 ETSI STC SMG2 meeting no 18, Helsinki, Finland, May 1996

[6] M. Russell and G. L. Stüber, ”Terrestrial digital video broadcasting for mobile reception using OFDM”, Wireless Personal Communications, vol 2, nr 1-2, Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 45-66, 1995

[7] P. Hoeher, S. Kaiser and P. Robertson, ”Pilot-symbol-aided channel estimation in time and frequency”, In K. Fazel and G. P. Fettweis, editors, Multi-Carrier Spread-Spectrum, chapter IV, pp 169-178. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1997.

[8] European Telecommunications Standard Institute, ”GSM 05.05”, version 4.6.0, July 1993.

[9] Y. Zhao and S.-G. Häggman, ”Sensitivity to Doppler shift and carrier frequency errors in OFDM systems -- The consequences and solutions”, IEEE Vehicular Technology Conference, Atlanta, USA, April 29 - May 2, pp. 1564-1568, 1996.

[10] S. K. Wilson, ”Digital audio broadcasting in a fading and dispersive channel”, Ph.D. thesis, Stanford University, USA, 1994.

[11] J. G. Proakis, ”Digital communications”, third edition, McGraw-Hill, New York, USA, 1995.

Paper C

This part will be published as:

Fredrik Tufvesson, Mike Faulkner and Torleiv Maseng, "Pre-Compensation for Rayleigh Fading Channels in Time Division Duplex OFDM systems", Wireless Personal Communication, 2000.

Pre-Compensation for Rayleigh Fading Channels in Time Division Duplex OFDM systems

Fredrik Tufvesson[1], Mike Faulkner[2] and Torleiv Maseng[1]

Abstract

Coherent demodulation results in good detection performance but requires channel estimation. Fading pre-compensation (precoding) at the transmitter can lead to low-complex receiver structures with good performance capabilities, without the need for channel estimation. Time division duplex systems based on orthogonal frequency division multiplex (OFDM) are particularly suited to this because intersymbol interference effects can be neglected, simplifying transmitter adjustments. Methods that involve amplitude and/or phase pre-compensation are compared in terms of resulting bit error rate and increase in peak-to-average power ratio. Dynamic channels degrade the performance as the block lengths get longer. For a certain block length the performance degrades below that of traditional differential decoding. A block length of up to 40 times that used in the Digital European Cordless Telephone system, DECT, is possible using channel estimation and ideal Wiener prediction.

1 Introduction

In time division duplex (TDD) systems the same frequency is used both by the base station and the mobile terminal, but at different time instants. In these systems the possibility of making predictions of the channel impact is rather straightforward. The used radio channel is often assumed to be reciprocal, see e.g. [1] [2], since we are interested in compensating the multipath channel and not the interference. If the block length is short compared to the speed of change of the channel, then predictions of the channel influence are likely to be accurate for adjacent blocks and we can adjust the signal to be transmitted to achieve better performance.

Pre-compensation can be used for different purposes, for example to control the

1 Department of Applied Electronics, Lund University, Box 118, SE - 221 00 Lund, Sweden, E-mail: Fredrik.Tufvesson@tde.lth.se

2 School of Communications and Informatics, Victoria University of Technology, PO Box 14428, MCMC, Melbourne, Vic. 8001, Australia

output power [3], to reduce interference from other users [4], or to modify the transmitted signal in order to reduce intersymbol interference (ISI) [1]. In this paper we examine the possibilities and limits when using pre-compensation of the channel impact on the downlink symbols in orthogonal frequency division multiplex (OFDM) systems. Figure 1 shows the principle of the technique. In OFDM systems it is

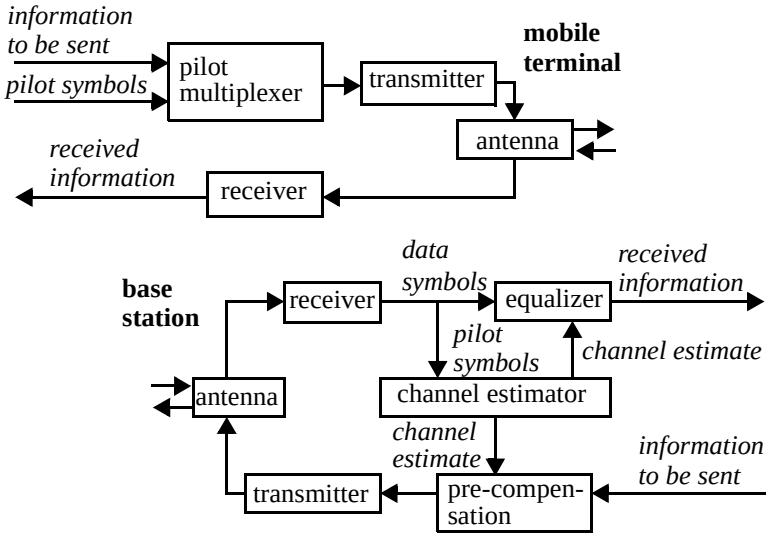

Figure 1: Pre-compensation where the channel influence is compensated in advance on the downlink.

straightforward to compensate for a time dispersive channel since its impact after the FFT only is an amplitude attenuation and a phase shift of the data symbols. Channel estimation is normally a complex operation to perform and every mobile terminal has to make channel estimates in order to use coherent detection. By pre-compensation it is possible to move the task of channel estimation to the base station, but still achieve the performance benefits of coherent detection at the mobile terminal [5]. This means that cheap, less complex, mobile receivers can be used. We address here the problem of channel pre-compensation in a simple uncoded OFDM system, but the results can be applied for other systems where the channel can be modeled as a single complex tap or together with coding and diversity arrangements.

Compensating channel attenuation means that we transmit more power where the channel attenuation is high and less power where the channel attenuation is low. This contradicts the water pouring theorem, which states that more power and higher order modulation should be used where the channel attenuation is low in order to maximize

the transmission capacity. However, here the aim is to use fixed modulation and enable a low-complex receiver.

A drawback when compensating channel attenuation is also that the amplitude variations of the transmitted signal become larger. The resulting peak-to-average power ratio has to be limited. For PSK systems it is not necessary to compensate amplitude variations and in this paper we show that phase compensation only may be adequate in these cases. In the paper we are interested in evaluating the performance limits for the system. We therefore assume perfect time synchronization and ideal oscillators without phase or frequency jitter. In a real system such impairments naturally degrade the performance and their influence needs to be taken into account.

The paper is organized as follows: In section 2, the system and the model for the OFDM system are described. In section 3 we then present the three analyzed compensation methods, and in section 4 we derive their performance in case where the channel is known by the base station. In section 5 the resulting increase in peak-to-average power ratio is derived for the three methods, respectively. We then proceed in section 6 with the method where only the phase shift of the channel is compensated and we derive the maximum block length before the performance gets worse than that of differential detection. Finally, the conclusions are presented in section 7.

2 System

In the paper we use an OFDM TDD-FDMA system with M sub-channels, see Figure 2 for a description of the transmitter and receiver. The OFDM signal is transmitted through a time dispersive Rayleigh fading channel. It is assumed that the fading is flat for each of the sub-channels and constant during a symbol interval. We assume further a cyclic prefix exceeding the maximum excess delay, and that the demodulated signals do not suffer from interchannel interference nor ISI.

Under these assumptions we can use a model for the OFDM system where the data

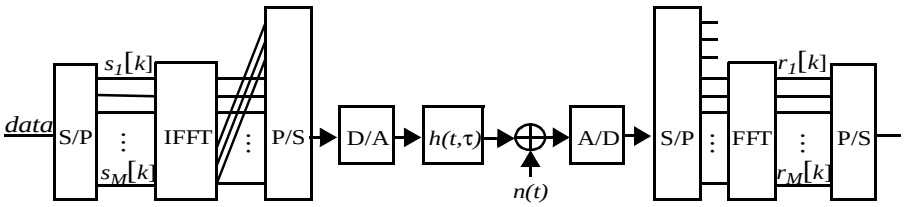

Figure 2: Block description of the transmitter-receiver chain.

signal is transmitted on separate, but correlated, Rayleigh fading channels with additive Gaussian noise. The transmitted OFDM signal is convolved with the channel impulse response, $h(t, \tau)$, but after demodulation the convolution only results in a phase shift, $\theta_m[k]$, and a power gain, $y_m[k]$, of the data symbols. The sampled complex representation of the received signal on channel m at time k can therefore be written as [6]

$$r_m[k] = \sqrt{y_m[k]}e^{j\theta_m[k]}s_m[k] + n_m[k], \quad (1)$$

where $s_m[k]$ is the complex representation of the transmitted data signal and $n_m[k]$ is independent samples of white Gaussian noise.

3 Compensation Methods

We evaluate three compensation methods:

1. Phase compensation only, the transmitted phase is altered by $\theta_m[k]$ radians. This method gives no increase in the transmitted signal dynamics.
2. Phase and amplitude compensation when $y>a_2$, otherwise no transmission. The transmitter output is inhibited when the channel gain does not exceed the threshold a_2. This saves power when the channel is poor.
3. Phase and amplitude compensation when $y>a_3$, otherwise power limiting. The transmitter output saturates when the channel gain does not exceed the threshold a_3. This method limits the output peak power to a reasonable level.

In the following the sub-script m is omitted to simplify notation, but every sub-channel is compensated individually by its own amplitude and/or phase.

The base station makes estimates of the channel amplitude and channel phase e.g. by using known pilot symbols from the uplink. Then the signal to be transmitted is compensated. If we let $u_q'[k]$ denote the complex representation of the uncompensated signal alternative q, the complex representation of the compensated signals, $u_q[k]$, becomes

$$(u_q[k])_1 = e^{-j\theta}u_q'[k] \quad (2)$$

$$(u_q[k])_2 = \begin{cases} 0 & y < a_2 \\ \sqrt{\frac{\eta_2}{y}}e^{-j\theta}u_q'[k] & y \geq a_2 \end{cases} \quad (3)$$

$$(u_q[k])_3 = \begin{cases} \sqrt{\frac{\eta_3}{a_3}} e^{-j\theta} u_q'[k] & y < a_3 \\ \sqrt{\frac{\eta_3}{y}} e^{-j\theta} u_q'[k] & y \geq a_3 \end{cases}, \tag{4}$$

for the three methods respectively. Here η_2 and η_3 are power normalization coefficients. The probability density function of the power gain, *y*, in a Rayleigh channel with unity mean is $p_Y(y) = e^{-y}$. Hence, if the channel is perfectly estimated and E_q' denotes the symbol energy before compensation, then the respective mean transmitted symbol energies are

$$(\overline{E_q})_1 = E_q' \tag{5}$$

$$(\overline{E_q})_2 = \eta_2 E_q' \int_{a_2}^{\infty} \frac{e^{-y}}{y} dy \tag{6}$$

$$\begin{aligned} (\overline{E_q})_3 &= \int_0^{a_3} \frac{\eta_3}{a_3} e^{-y} E_q' dy + \int_{a_3}^{\infty} \frac{\eta_3}{y} e^{-y} E_q' dy \\ &= \eta_3 E_q' \left(\frac{1}{a_3}(1 - e^{-a_3}) + \int_{a_3}^{\infty} \frac{e^{-y}}{y} dy \right). \end{aligned} \tag{7}$$

In order to maintain the same transmitted mean energy for the three methods, the power normalization coefficients are consequently chosen as

$$\eta_2 = 1 \Big/ \left(\int_{a_2}^{\infty} \frac{e^{-y}}{y} dy \right) \tag{8}$$

$$\eta_3 = 1 \Big/ \left(\frac{1}{a_3}(1 - e^{-a_3}) + \int_{a_3}^{\infty} \frac{e^{-y}}{y} dy \right). \tag{9}$$

4 Performance when the Channel is Known by the Base Station

The performance limit for each of the given methods is achieved when all downlink transfer functions are known by the base station and the channel is compensated perfectly. This is of course not the situation in a practical case, but it is anyhow interesting to get a lower limit for the resulting bit error rate (BER).

The BER, P_b, when the Rayleigh fading channel is perfectly known is for the first method the same as for ordinary coherent detection of QPSK signals, i.e. [6]

$$\begin{aligned}(P_b)_1 &= \int_0^\infty Q\left(\sqrt{y\frac{2E_b'}{N_0}}\right)e^{-y}dy \\ &= \frac{1}{2}\left(1-\sqrt{\frac{E_b'/N_0}{1+E_b'/N_0}}\right).\end{aligned} \tag{10}$$

Here N_0 denotes the density of the noise and E_b' denotes the bit energy before compensation, $E_b' = E_q'/2$ for QPSK. The BER, when the base station knows the channel impact, can for the two other methods be calculated as

$$\begin{aligned}(P_b)_2 &= \int_0^{a_2}\frac{1}{2}e^{-y}dy + \int_{a_2}^\infty Q\left(\sqrt{\eta_2\frac{2E_b'}{N_0}}\right)e^{-y}dy \\ &= \frac{1}{2}(1-e^{-a_2}) + e^{-a_2}Q\left(\sqrt{\eta_2\frac{2E_b'}{N_0}}\right)\end{aligned} \tag{11}$$

$$\begin{aligned}(P_b)_3 &= \int_0^{a_3} Q\left(\sqrt{y\frac{\eta_3}{a_3}\frac{2E_b'}{N_0}}\right)e^{-y}dy + \int_{a_3}^\infty Q\left(\sqrt{\eta_3\frac{2E_b'}{N_0}}\right)e^{-y}dy \\ &= \frac{1}{2} + \sqrt{\frac{\eta_3(E_b'/N_0)}{a_3+\eta_3(E_b'/N_0)}}\left[Q\left(\sqrt{2a_3+\eta_3\frac{2E_b'}{N_0}}\right)-\frac{1}{2}\right].\end{aligned} \tag{12}$$

The resulting BER is shown in Figure 3 together with the curves for differential QPSK [6].

At low signal to noise ratios phase-only compensation appears to be best for the majority of power compensating limits. At medium and high signal to noise ratios

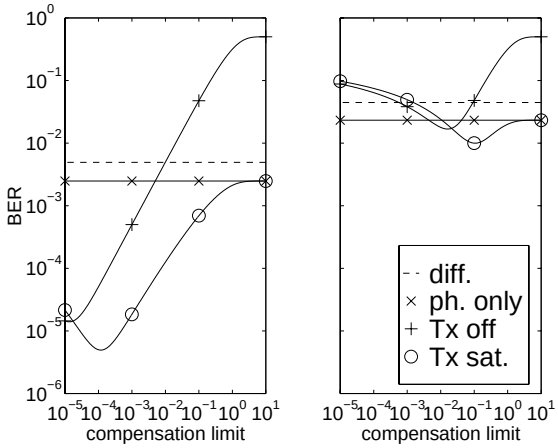

Figure 3: BER for the compensation methods when the channel is known by the base station, E_b'/N_0= 20 dB (left) and 10 dB (right). The compensation limit (i.e. a_1, a_2 or a_3) is the maximum power amplification (attenuation) of the channel to be entirely compensated.

there is some value in adding amplitude compensation with peak power saturation, but the question is whether one can afford the increased peak-to-average power ratio. For the amplitude compensation methods a low compensation limit means that too much energy is spent on compensating very bad channels, while a high limit means that the transmitter is saturated or off even when communication normally is advisable.

5 Peak-to-Average Power Ratio

The compensation limits, a_2 and a_3, can be seen as measures of the dynamic range for the power compensation. In practical measurements the peak-to-average ratio of the transmitted power is often used, which for the pre-compensation methods is the product of the peak-to-average ratio of the modulation method, e.g. OFDM, and the increase due to the compensation. For a single sub-channel the increase in peak-to-average power ratio can be calculated as 1, η_2/a_2 and η_3/a_3 for the three methods respectively, but for an OFDM signal the increase is dependent on the correlation between the sub-channels and which particular channels that are compensated. If we compensate the phase only, then the transmitted signal is a sum of sinusoids with different phases and there is no increase in the peak-to-average power ratio. If we compensate the amplitude as well, then the transmitted signal is a sum of weighted sinusoids and we get a larger peak-to-average power ratio. In Figure 4 we present

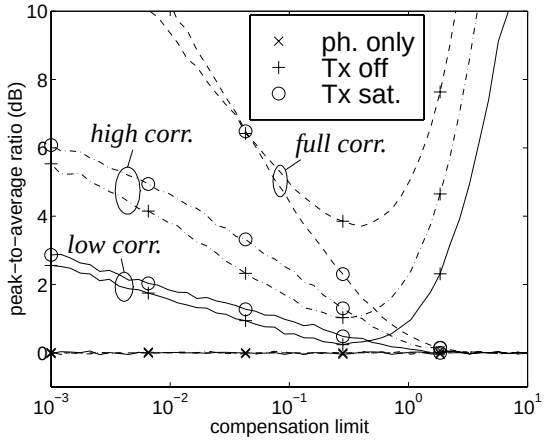

Figure 4: Simulated values of the increase in peak-to-average power ratio when applying the three methods for an OFDM-based wireless-LAN. The dashed lines represent a channel with no delay spread and fully correlated sub-channels. The dash-dotted lines represent an indoor channel with short delay spread and high correlation (τ_{rms} =50ns), and the solid lines represent an indoor channel with long delay spread and low correlation (τ_{rms} =250ns).

simulated values of the increase in peak-to-average power ratio when applying the different methods in channels with various delay spreads. We use here the OFDM based wireless local area network (LAN) specified in ETSI BRAN [7][1] as an example. The system uses 52 sub-channels and has a sub-channel bandwidth of 312.5 kHz. For the simulations we use three different channel models:

- A one tap channel, giving full correlation between the sub-channels.
- An indoor channel with short delay spread (“channel A”, τ_{rms} =50ns), resulting in large sub-channel correlation.
- A channel with long delay spread modeling large open buildings (“channel E”, τ_{rms} =250ns). This gives low correlation between the sub-channels.

The simulated values of the increase in peak-to-average power ratio are achieved as the difference between the 99.5% peak-to-average power ratio levels for the uncompensated and pre-compensated OFDM signal.

1 This system is similar to IEEE 802.11a standardized in the U.S.

In Figure 4 we see that we get a larger increase in peak-to-average power ratio if the sub-channels are correlated, i.e. the delay spread is short. A large peak in the transmitted signal is achieved when the sub-channels contributing to the peak are amplified due to the amplitude compensation. When the correlation is small, the increase becomes smaller. This is because the probability that all sub-channels contributing to the peak are amplified becomes lower. The simulations are performed for a specific system, but similar results can be expected for other OFDM systems if we take into account the relation between the coherence bandwidth of the channel and the OFDM bandwidth.

Phase-only compensation produces no increase in peak-to-average power ratio and for known channels we get the same performance as for coherent detection. If it is possible to increase the peak-to-average power ratio by 2 dB, we see in Figure 4 that we can use a_3=0.15 for the method where the transmitter is saturated. As seen from Figure 3, it is then possible to decrease the BER 2-3 times compared to phase-only compensation. However, it is doubtful whether one can afford to increase the peak-to-average power ratio for wireless OFDM applications, since power variations already are a problem in these systems.

6 Maximum Block Length

In practice the transfer functions are not known, but they have to be estimated e.g. by transmitting known so-called pilot symbols. The impact of the channel on the pilot symbols is observed, and the channel coefficients for the data symbols are achieved by filtering and prediction. Different alternatives can be used to form these channel estimates. Here we use a two-dimensional Wiener filter, which uses both the correlation in time and in frequency to derive the channel estimates. The Wiener filter is optimal in the sense that the mean squared error is minimized. For further details about channel estimation in OFDM systems see e.g. [8].

The channel predictions for the downlink block become obsolete after a certain time. This means that the length of the downlink block is limited before the BER reaches an unacceptable level. This particular level is dependent on the application. Here we compare the BER to that of differential detection and use the latter as a benchmark since the complexity of the mobile terminals is almost the same. The maximum block length is a function of the Doppler frequency, f_D, in the system since it determines the speed of change in the channel. In the following we assume a classical U-shaped Doppler spectra such that the time correlation can be described by a zeroth-order Bessel function. We also concentrate on the performance of the phase-only compensated system, since there is no increase in peak-to-average power ratio. In order to investigate the influence of the observation time we study the performance of a single sub-channel in the OFDM system. We use almost noise free pilot symbols

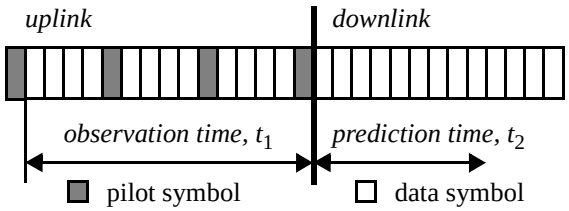

Figure 5: Block structure to investigate the influence of the observation time.

for channel estimation where the observation time t_1, the time between the first and last pilot symbol, is varied, see Figure 5. The low-noise pilot symbols are used in the analysis since we want to separate the impact of the time varying channel. In the analysis we have not compensated the symbol energy of the data symbols due to the boosted pilots. For long and medium block lengths the performance is limited by the time varying channel and not by the noise. Besides, in a well designed OFDM system the noise affecting the pilot symbols are averaged over several symbols when calculating the channel estimates. Therefore, the variance of the channel estimation error is much lower than the variance of the noise affecting data symbols.

In Figure 6 the BER as a function of prediction time t_2, the time since the last pilot symbol, is presented for different observation times. All the pilot symbols, are

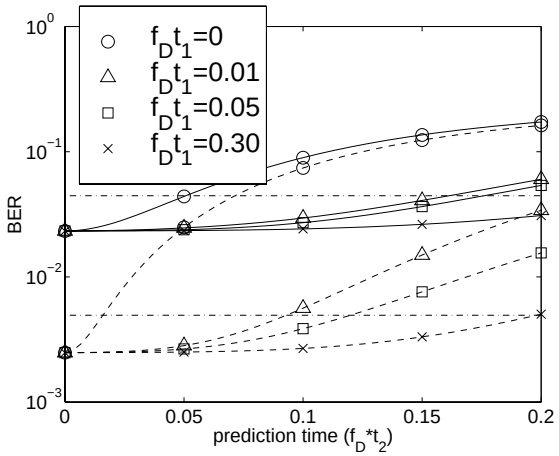

Figure 6: BER as a function of prediction time, f_Dt_2, for different observation times, f_Dt_1. E_b/N_0=20 dB for the dashed lines and E_b/N_0=10 dB for the solid ones. The dash-dotted lines represent differential detection.

equally distributed within the observation time t_1. In this example pilot symbols at P=4 different time instants are used for prediction. The signal to noise ratio, E_b/N_0, is set to 20 dB and 10 dB respectively, except for the pilot symbols where we use a 30 dB higher signal to noise ratio. We assume that the Doppler frequency is known and use a Wiener filter for estimation and prediction of the channel. The BER is derived from the mean squared error, σ_e^2, of the estimate. For QPSK it is given by the correlation coefficient, μ, between the backrotated received samples and the channel estimate as [6, appendix C]

$$P_b = \frac{1}{2} - \frac{\mu}{2\sqrt{2-\mu^2}}. \tag{13}$$

The correlation coefficient when using pilot symbols can be calculated as

$$\mu_{est} = \frac{E_s\bar{y}}{\sqrt{(E_s\bar{y}+N_0)(E_s\bar{y}+\sigma_e^2)}}, \tag{14}$$

where $\bar{y} = E[y]$ is the mean power of the channel. The BER for DQPSK is also given by (13), but with

$$\mu_{diff} = \frac{E_s}{E_s+N_0}. \tag{15}$$

The BER of the downlink block is determined by the ability of making correct channel predictions. This ability is dependent on the location of the pilot symbols in the uplink block, especially their spreading in time, the observation time. Longer observation time means in general better predictions. An observation time of zero corresponds to keeping the last channel estimate in the uplink block as a channel prediction for the whole downlink block. By just adding some information about the trend of the channel values, a considerable decrease in the BER could be achieved. But the enhanced predictions are naturally dependent on the validity of the channel model and the ability to suppress the noise affecting the pilot symbols.

At a certain prediction time, the differential system and the pre-compensated system have the same bit error performance, see Figure 6. This point, the break even point, is evaluated in Figure 7 as a function of the observation time and the number of used pilot symbols using the same assumptions as before. For block lengths smaller than the ones in Figure 7 the (instantaneous) BER is lower for a pre-compensated system than for a differential system. Beyond the break even point, differential detection without pre-compensation gives better results for similar receiver complexity. As an example, look at Figure 6, E_b/N_0=10 dB and the curve for observation time f_Dt_1=0.05. After a certain time f_Dt_2=0.17 the BER is the same as for a differential sys-

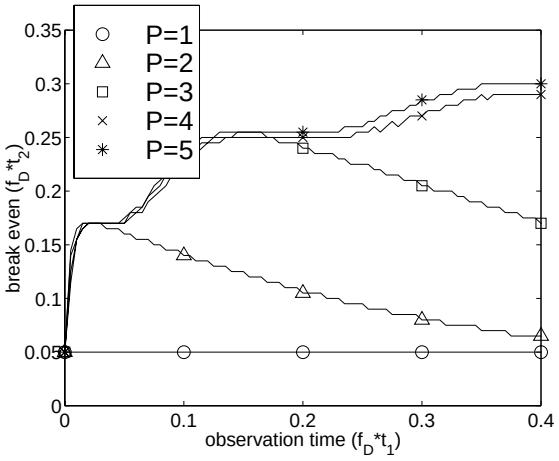

Figure 7: Break even time where the BER for a pre-compensated system starts to exceed the one for a differential system, E_b/N_0= 10 dB. P is the number of time instants in the uplink block used for pilot symbols.

tem. This point can be found in Figure 7. For P=4 and an observation time of 0.05 the break even point is $f_D t_2$=0.17. In Figure 7 we clearly see the effect of the observation time. By using all the pilot symbols for prediction and not only the last estimate there is an increase in the break even time of at least a factor 3. We also see the importance of using sufficiently many pilot symbols to ensure that there always are pilot symbols with high correlation to the point to be predicted. Pilot symbols at 3 or 4 time instants in the uplink block seems enough for this.

To compare the above block lengths to an existing TDD system we can use the Digital European Cordless Telephone (DECT) system, which has a block length of 368 µs [9]. For a terminal at pedestrian speed (10 km/h, 1900 MHz carrier frequency) this results in a normalized block length of $f_D t$=0.0065. With pre-compensation, E_b/N_0=10 dB, P=4 and a known Doppler frequency the break even time becomes $f_D t_2$=0.26 (when the observation time is equal to the break even time). This means that it is possible to use approximately 40 times longer block lengths than the one in the DECT system before the BER becomes worse than that of differential detection.

In a real system the statistics of the time variations may not be known and one often avoids the requirement of measuring Doppler frequency. The estimation and prediction filters are therefore often adjusted to the worst case scenario for the channel changes. This of course decreases the feasible block lengths in the system. In Figure 8 the break even points are given when the Wiener filter is designed for k times

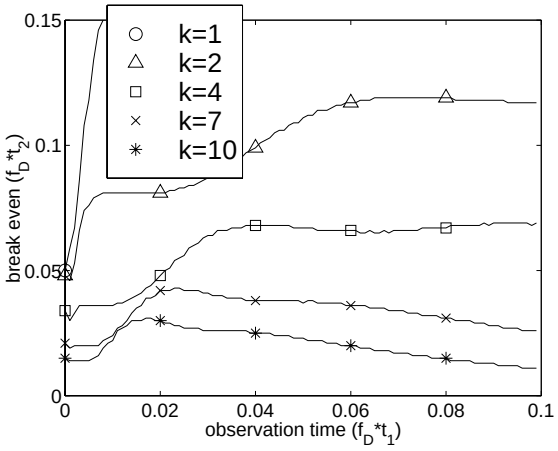

Figure 8: Break even time when the Wiener filter is adjusted to k times the actual Doppler frequency and pilot symbols at P=4 time instants are used for prediction, E_b/N_0= 10.

higher Doppler frequency than the actual one. If a pedestrian for example reduces the speed to 2.5 km/h when the Wiener filter is designed for 10 km/h, the mismatch factor k is 4 and the break even time is reduced from $f_D t_2$=0.18 to 0.07. But this means on the other hand that absolute block length measured in number of symbols or the time t_2 is longer in the slow but mismatched case since the Doppler frequency is smaller there, $(t_2)_{2.5\ \mathrm{km/h}} \approx 0.016$ s instead of $(t_2)_{10\ \mathrm{km/h}} \approx 0.010$ s. The feasible block length in number of symbols is in the general case still determined by the worst case Doppler frequency and therefore mismatch is not regarded as a problem.

7 Conclusions

The BER when applying different methods of pre-compensating QPSK signals in a Rayleigh fading channel is presented. Pre-compensation can be used to achieve a performance better than that of traditional differential detection without using channel estimation and equalization in the mobile receivers. Phase-only compensation is proposed since there is no increase in the peak-to-average power ratio. It is shown that block lengths 40 times longer than the one in the DECT system could be used, and anyhow get a performance better than in differential systems. The block length analysis was made for a single sub-channel assuming low-noise pilot symbols. This assumption is not as restrictive as one may think, since many pilot symbols normally are used in OFDM systems to form the channel estimates. The correlation in time and in frequency results in noise averaging and therefore the channel estimates become

accurate. For unknown Doppler frequencies where the filters are adjusted to the worst case scenario, the feasible block lengths are determined by the maximum Doppler frequency, though the filters are mismatched for lower Doppler frequencies.

References

[1] A. Masoomzadeh-Fard and S. Pasupathy, “Combined equalization and differential detection using precoding”, IEEE Transactions on Communications, vol. 45, no. 3, pp. 274-278, 1997.

[2] W. Zhuang and V. Huang, “Phase pre-coding for frequency selective Rayleigh and Ricean slowly fading channels”, IEEE Transactions on Vehicular Technology, vol. 46, no. 1, pp. 129-142, 1997.

[3] H. Matsuki and H. Takanashi, “Theoretical analysis of BER improvement by adaptive transmitter power control for Rayleigh fading compensation”, Electronics Letters, 29, (17), pp. 1520-1521, 1993.

[4] I. Ghauri and R. Knopp, “Power control for diversity reception in time-division duplex CDMA”, Proc. of IEEE Vehicular Technology Conference, Ottawa, Canada, pp. 2343-2347, May 1998.

[5] P. S. Henry and B. S. Glance, “A new approach to high-capacity digital mobile radio”, The Bell system technical journal, Vol. 60, No. 8, pp. 1891-1904, 1981.

[6] J. G. Proakis, Digital communications, third edition, McGraw-Hill: New York, USA, 1995.

[7] European Telecommunications Standards Institute, “Broadband Radio Access Networks (BRAN); HIPERLAN type 2; Physical (PHY) Layer”, ETSI TS 101 475 V1.1.1, ETSI, Sophia Antipolis, France, 2000.

[8] P. Hoeher, “TCM on frequency-selective land-mobile fading channels”, in Proc. 5th Tirrenia International Workshop in Digital Communication, Tirennia, Italy, pp. 317-328, Sept. 1991; E. Biglieri and M. Luise (eds.), Coded modulation and bandwidth-efficient transmission, Elsevier Science Publishers: Amsterdam, the Netherlands, 1992.

[9] J. D. Gibson, editor, The mobile communications handbook, CRC-press: Boca Raton, Fl, USA, 1996.

Fredrik Tufvesson was born in Lund, Sweden in 1970. He received the M.S. degree in Electrical Engineering from Lund University in 1994, the Licentiate Degree in 1998 and he is currently a Doctoral student at the Department of Applied Electronics, Lund University. His research interests include channel estimation and synchroniza-

tion problems for wireless communication, especially in wireless OFDM systems.

Mike Faulkner received the B.S.E.E. from Queen Mary College, London University, U.K., the M.E. degree from university of New South Wales, Australia and the Ph.D. degree from the University of Technology, Sydney in 1970, 1978 and 1993 respectively. Between 1988 and 1996 he spent three periods at Lund University, Sweden, the last time as a Visiting Professor working in radio hardware design. Since 1977 he has been a Lecturer and then Professor at Victoria University of Technology, Melbourne, Australia. His current interests are signal processing, radio circuit design and spread spectrum.

Torleiv Maseng is Director of Research at the Norwegian Defence Research Establishment where he is responsible for communications and information systems. He has worked as a scientist at SINTEF in Trondheim where he was involved in design and standardization of GSM for ten years. He has also been scientist at NC3A, former STC, which is a NATO research centre in The Hague, for seven years. During 1992-94 he was the technical manager for the new private mobile operator in Norway. Since 1994 he holds a chair in radio communications at Lund University, Sweden. This work is down to 20% since 1996, when he took up his employment at the Norwegian Defence Research Establishment outside Oslo. He is the author of more than 100 papers, holds patents and is a Technical Editor of the IEEE Communications Magazine. He has received an award for outstanding research, has arranged large international conferences and is quoted in “Who is Who in the World”.

Paper D

This part has been submitted as:

Fredrik Tufvesson and Ove Edfors, "Preamble-based Time and Frequency Synchronization for OFDM Systems", submitted to IEEE Journal on Selected Areas in Communications, January 2000.

Preamble-based Time and Frequency Synchronization for OFDM Systems[1]

Fredrik Tufvesson, *student member, IEEE*, and
Ove Edfors, *member, IEEE*
Department of Applied Electronics, Lund University, Sweden

Abstract

Fast and reliable time and frequency synchronization is crucial for packet-based orthogonal frequency division multiplex (OFDM) systems. Various synchronization methods are analyzed to determine their performance when using preambles based on repeated short pseudo noise (PN) sequences or conventional preambles based on repeated OFDM data symbols. We make analytical evaluations for AWGN channels and simulate the performance for one- and two-tap Rayleigh fading channels. Conventionally, the synchronization signal is calculated as a sum of products between received samples. However, it is shown that switching the order of summation and multiplication improves the detection performance in terms of lower probability of false detection and of missing the synchronization signal. The false detection probability is decreased by several orders of magnitude for reasonable parameter settings. The performance in terms of frequency-offset estimation is similar for both preambles. The PN-based preambles have low peak-to-average power ratio and they make it possible to use analog-to-digital converters with only one-bit quantization in stand-by mode. Therefore, the PN-based preambles allow for a great reduction in stand-by mode power consumption.

1 Introduction

There is great interest in using orthogonal frequency division multiplex (OFDM) for high-speed wireless local area network applications. Standardization of systems where OFDM is used is occurring both in the United States [1], Japan [2] and Europe [3]. The systems are based on packet data transmission and therefore fast and reliable time and frequency synchronization with low overhead is crucial to the systems.

Often, a repeated OFDM data symbol is proposed for synchronization issues in OFDM systems, see e.g. [4] [5]. The synchronization signal is achieved by correlat-

1 Parts of this work have been presented at the IEEE Vehicular Technology Conference, Amsterdam, The Netherlands, September 1999.

ing samples one symbol or half a symbol apart. The timing signal is then the amplitude of the correlator output. Time synchronization is achieved when the amplitude exceeds a certain threshold. The frequency-offset estimate is then obtained from the phase of the same signal. When using a cyclic prefix the timing signal has a flat region, which causes an uncertainty about where the actual start of the packet is [6]. On the other hand, if no cyclic prefix is used, the frequency-offset estimate is corrupted by intersymbol interference. This has to be suppressed by weighting the received samples [7].

Synchronization techniques like this, where the synchronization signal is calculated by correlation, require heavy computations while less complex solutions are favorable. Alternative methods resulting in lower complexity include e.g. the use of a preamble, which uses only half of the OFDM spectrum for transmission [8]. The synchronization signal is then achieved as the ratio between the powers in each half of the frequency band. This method works well for frame synchronization at medium and high signal-to-noise (SNR) ratios, but gives only a coarse time synchronization signal and no frequency-offset estimate. There is still a timing uncertainty and, for low SNR values, the probability of missing the synchronization signal is large.

In CDMA systems synchronization is generally achieved by correlation to a known code sequence. The correlation means that the influence of noise and interference is reduced so that synchronization can be achieved at very low SNR values. Further, since the transmitted signal is a single carrier signal, the power variations can be made low. However, the calculation of the frequency-offset estimate is not as straightforward as in OFDM systems [9].

In this paper we propose a synchronization technique for OFDM systems that combines the advantages of time synchronization in CDMA systems with those of the simple frequency estimation scheme in OFDM systems. We propose a preamble that consists of a known short repeated pseudo noise sequence (PN-sequence) to be used together with a synchronizer that correlates received samples before processing them further. The advantages of this technique compared to conventional OFDM synchronization are:

- The synchronization signal has a sharp synchronization peak and no ambiguity due to the cyclic prefix.

- There are low power variations within the preamble, which means relaxed requirements on the automatic gain control (AGC).

- It is possible to use analog-to-digital converters with only one-bit quantization for coarse time synchronization. This means low power consumption in stand-by mode.

- The amplitude of the synchronization signal is small when the timing is wrong compared to the amplitude when the timing is correct. This means low probability of false detection and of missing the correct synchronization signal.

To our knowledge, this is a novel approach for OFDM system synchronization and no analysis of the statistics of the synchronization signal has been made before. The idea of using one-bit quantization has been presented before for conventional OFDM synchronization [10] but there the quantization loss was large. By using the technique presented here this loss becomes low enough to make a one-bit solution attractive.

The paper is organized as follows: In section 2 the synchronization system is described. In section 3 we analytically analyze the performance in AWGN channels. In section 4 the performance for two-tap Rayleigh fading channels is investigated by means of simulations and, finally, in section 5 we present the conclusions.

2 System Description

The preambles are designed for an OFDM system with *M* sub-channels. The aim is to estimate the start of the packet and achieve a reliable frequency-offset estimate using a single-symbol preamble, but other preamble lengths are of course also possible. We compare two preambles, a PN-based and a conventional one. The comparison is performed using two alternative synchronizer structures, the conventional OFDM synchronizer and a pre-correlating synchronizer.

2.1 Preamble Structures

The PN-based preamble was derived from a *R/2* chip long m-sequence, where *R* denotes the repetition length of the preamble. The sequence was oversampled twice and repeated until the total length became *2R+G* samples, as shown in Figure 1, where *G* denotes the length of the guard interval. In order to decrease the power vari-

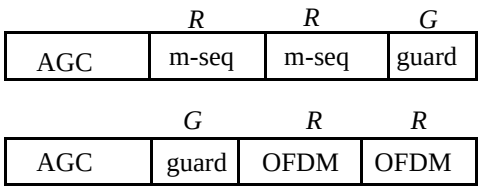

Figure 1: Evaluated preambles and their length in samples.

ations within the preamble we used an offset QAM constellation for transmission where the quadrature sequence was delayed one sample compared to the inphase sequence. The preamble can be generated using the FFT of the PN-sequence as input to the OFDM modulator so no extra device is required for preamble generation.

The conventional preamble consists of a repeated OFDM data symbol as shown in Figure 1. The total preamble length was $2R+G$ samples, where R normally is half the OFDM symbol length, i.e. $R=M/2$. The data symbols were chosen to give good autocorrelation properties and the repetition of the R samples was achieved by using only the even sub-channels for transmission.

2.2 Synchronizer Structures

The main parts of the pre-correlating synchronizer are presented in Figure 2. First, pre-correlation of length K with the known sequence is performed. When the received sequence coincides with the known sequence, a large peak occurs with a phase angle determined by the channel. When there is no synchronization signal, the output is noise only. The correlator output is multiplied by the conjugate of the output delayed PR times. The positive integer P denotes an extra delay, which can be used to decrease the variance of the frequency-offset estimate. Most often no extra delay is used and $P=1$. Correlation peaks spaced PR samples apart produce a large synchronization signal with a (small) phase shift proportional to the frequency-offset. The maximum frequency offset is inversely proportional to the pre-correlator length. If the pre-correlator length is limited, e.g. due to a large frequency offset, and the variations of the synchronization signal is too large it is possible to sum consecutive peaks to decrease the variance. However, normally a long correlator is preferred and no summation is necessary.

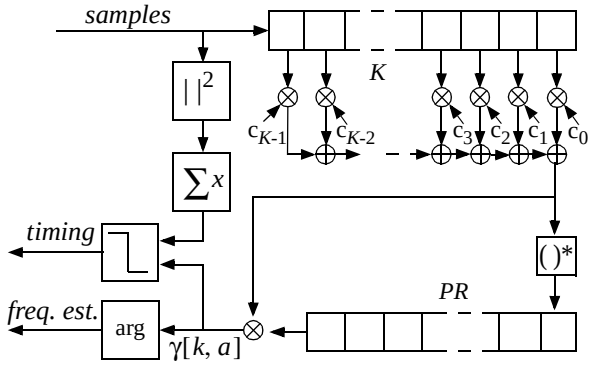

Figure 2: Structure of the pre-correlating synchronizer when using only one pre-correlator product for synchronization ($L=1$).

The conventional synchronizer multiplies samples spaced R samples apart and uses a sliding window to sum R consecutive products. This synchronizer structure is presented in Figure 3. When the repeated signal is present, the phases of the products within the window are nearly the same and they essentially add in phase. Signals that are not repeated generally result in small products with random phases. Then the products do not add in phase and the synchronization signal becomes small.

For both methods the absolute value of the synchronization signal is compared to the short time mean of the received power. If a threshold is exceeded, time synchronization is achieved and the frequency estimate is available from the phase of the synchronization signal.

2.3 Synchronization Signal

For the analytical performance evaluation we consider an AWGN channel. The base band representation of the received signal is

$$r_0(t) = s_0(t)\exp\left(j\left(\frac{2\pi\varepsilon_c t}{T_s} + \theta\right)\right) + n_0(t)\,, \tag{1}$$

where $n_0(t)$ is white Gaussian noise with one-sided spectral density N_0, and $s_0(t)$ is the transmitted preamble. Further, θ is the carrier phase and $\varepsilon_c = \Delta f \cdot T_s$ is the frequency-offset to be estimated when normalized by the sub-channel spacing $1/T_s$.

The synchronization signals from the two synchronizer structures are represented in a common framework. Let $r[k]$ denote the sampled received signal and $c[k]$ the sequence used for pre-correlation. If the pre-correlating synchronizer is used then the pre-correlator length is equal to the repetition length, i.e. K=R. If no pre-correlator is

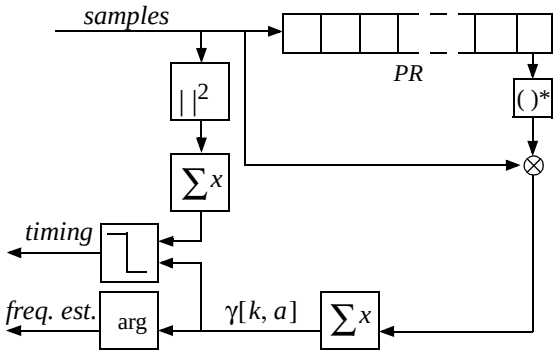

Figure 3: Structure of the conventional OFDM synchronizer.

used, as for conventional OFDM synchronization, then K=1 and all $c[k]$=1. Further, let a denote the slip between the received sequence and the locally generated pre-correlation sequence. In the following analysis this slip is assumed to be an integer. The received sampled signal is multiplied by the conjugate of the known pre-correlation sequence to get a "remodulated" signal, which is summed over length K to get the pre-correlator output. The remodulated signal can be calculated as

$$\begin{aligned} & c^*[k-a]r[k] \\ & = \sigma_s c^*[k-a]s[k]\exp\left(j\left(\frac{2\pi\varepsilon_c k}{M}+\theta_c\right)\right)+n[k], \end{aligned} \tag{2}$$

where * denotes complex conjugation, σ_s^2 is the power of the transmitted signal, M is the number of sub-channels in the OFDM system, $n[k]$ is white Gaussian noise with variance σ_n^2 and $s[k]$ is the transmitted preamble.

The synchronization signal for both the conventional and the pre-correlating synchronizer is given by

$$\begin{aligned} \gamma[k,a] = \sum_{l=0}^{L-1}\Bigg[\Bigg(\sum_{n=0}^{K-1} c^*[k-n-lR-a]r[k-n-lR]\Bigg) \\ \cdot\Bigg(\sum_{n=0}^{K-1} c^*[k-n-(l+P)R-a]r[k-n-(l+P)R]\Bigg)^*\Bigg], \end{aligned} \tag{3}$$

where, again, L is the number of products used for the synchronization signal, P is the extra delay when calculating the products and K is the pre-correlator length. When using the pre-correlator we often have only one product for synchronization and no extra delay so $L=1$ and $P=1$, without the pre-correlator these parameters are set to $L=R$ and $P=R$.

A PN-based preamble permits a very low arithmetic resolution in the synchronizers due to its low peak-to-average power ratio and good autocorrelation properties. There is of course a loss in performance but, if for example the length of the preamble is extended, it is possible to have a sleep mode operation where the synchronizer uses only one-bit resolution for time synchronization. The synchronization unit then only consists of a threshold detector, a K-input adder, a multiplier and a comparator. The synchronization signal is still given by (3), but the received signal now only take the values $r[k] = \pm 1 \pm j$. In this way only the sign of the received signal is used, thus the receiver does not have to compare the timing signal to the received power and can instead use a fixed threshold. This decrease in complexity renders the low-resolution

operation especially suitable in acquisition/sleep mode, where low power consumption is important.

3 Performance Evaluation

In this section we analytically analyze the performance of time and frequency synchronization for the two preambles and the two synchronizer structures in AWGN channels. To compare different parameter choices, we normalize the synchronization signal as $|\lambda[k, a]| = |\gamma[k, a]|/(LK^2\sigma_s^2)$. As we will see below, when the timing is correct the mean of the synchronization signal is almost unity, and when the timing is wrong the mean is close to zero.

In subsection 3.1, we analyze the distribution of the timing signal when the timing is correct, i.e. when *a=0* and there is no slip between the received and known sequence. In subsection 3.2 we derive the distribution of the timing signal when the timing is wrong. By wrong timing we mean that there is no synchronization signal present and the synchronizer is fed by white Gaussian noise only. We concentrate on the cases where the synchronization signal consists of one, two or a large number of products. Normally, only one product is used together with the pre-correlating synchronizer since this results in the best performance. Further, in subsection 3.3, we analyze the performance of time synchronization when using the low-resolution mode. In subsection 3.4 we summarize our results on time synchronization and finally, in subsection 3.5, we analyze the performance of the frequency-offset estimate. For the figures and numerical examples in this section we use an AWGN channel with 0 dB SNR and a preamble length of 60 samples (R=30), except for the evaluation of the frequency-offset estimate where we have other SNR values and the preamble length is 64 for the OFDM-based preamble. The extension is used in that case in order to get a full OFDM symbol as a preamble.

3.1 Time Synchronization, Correct Timing

When the pre-correlator is used, the synchronization signal is calculated as a sum of L products between delayed pre-correlator outputs. Let $\eta[k, a]$ denote the pre-correlator output. If the frequency offset is within the permitted range, $|\varepsilon_c| < M/(2PR)$, then $\eta[k, 0]$ can be calculated for both the conventional and the PN-based preamble as [12]

$$\eta[k, 0] = \frac{1}{\sqrt{L}}\mathrm{sinc}\left(\frac{\varepsilon_c K}{M}\right)\exp\left(j\frac{2\pi\varepsilon_c}{M}\left(k+\frac{1}{2}\right)\right) + n_\eta[k, 0], \tag{4}$$

where $n_\eta[k, 0]$ is white Gaussian noise with variance $\sigma_\eta^2 = \sigma_n^2/(KL\sigma_s^2)$ and the

output has been normalized by $\sqrt{LK^2\sigma_s^2}$. Define, in the same way as in [5], inphase as the phase of the sum,

$$\arg\left(\sum_L \eta[k,0]\eta^*[k-PR,0]\right) \approx \frac{2\pi\varepsilon_c PR}{M}. \tag{5}$$

For reasonable SNR values the noise-times-noise products are negligible and the inphase part of the synchronization signal $\lambda[k,0]$ is much larger than its quadrature part. The amplitude, i.e. the timing signal, can therefore be approximated by the inphase components as

$$\begin{aligned}|\lambda[k,0]| \approx \Lambda[k,0] \;=\; \sum_L \Bigg[& \frac{1}{L}\mathrm{sinc}^2\left(\frac{\varepsilon_c K}{M}\right) + \\ & InPhase\Bigg\{ n_\eta[k,0]\left(\frac{1}{\sqrt{L}}\mathrm{sinc}\left(\frac{\varepsilon_c K}{M}\right)\exp\left(-j\frac{2\pi\varepsilon_c(k-PR+1/2)}{M}\right)\right) \\ & + \left(\frac{1}{\sqrt{L}}\mathrm{sinc}\left(\frac{\varepsilon_c K}{M}\right)\exp\left(j\frac{2\pi\varepsilon_c(k+1/2)}{M}\right)\right) n_\eta{}^*[k-PR,0]\Bigg\}\Bigg].\end{aligned} \tag{6}$$

The approximation, $\Lambda[k,0]$, is Gaussian with mean

$$E[\Lambda[k,0]] \;=\; \mathrm{sinc}^2\left(\frac{\varepsilon_c K}{M}\right), \tag{7}$$

and variance

$$var(\Lambda[k,0]) \;=\; \sigma_n^2\mathrm{sinc}^2\left(\frac{\varepsilon_c K}{M}\right)/(LK\sigma_s^2). \tag{8}$$

The timing signal is always positive and, since the expected value is large compared to the variance, the approximation as a Gaussian gives a negligible probability of negative values. When there is a frequency offset, the mean decreases slightly since the samples do not add totally in phase. However, this loss is small in the interesting frequency-offset range. The cumulative density function (CDF) at the correct timing point can be expressed as

$$F_{|\lambda[k,0]|}(x) \approx F_{\Lambda[k,0]}(x) \;=\; 1 - Q\left(\frac{x - E[\Lambda[k,0]]}{\sqrt{var(\Lambda[k,0])}}\right), \tag{9}$$

where Q is the tail probability of a normalized Gaussian variable.

When the timing is correct, the distribution of the timing signal with the conven-

tional synchronizer is similar to that of the pre-correlating synchronizer. However, since there is no loss due to the phase rotation, the normalized synchronization signal can in this case be regarded as Gaussian with unity mean,

$$E[\Lambda_{conv}[k, 0]] = 1, \tag{10}$$

and variance

$$var(\Lambda_{conv}[k, 0]) = \frac{\sigma_n^2}{LK\sigma_s^2}. \tag{11}$$

Monte Carlo simulations, not presented here, show a good agreement between the true values and the approximations.

3.2 Time Synchronization, Wrong Timing

When using the pre-correlating synchronizer, the normalized complex outputs of the sub-correlators have zero mean and variance $\sigma_\eta^2 = \sigma_n^2/(KL\sigma_s^2)$. If the synchronization signal consists of one product only, i.e. L=1, we can rewrite the timing signal as

$$|\lambda[k, a]| = |\eta[k - lR, a]||\eta[k - (l + P)R, a]|. \tag{12}$$

Since the terms are the amplitude of complex Gaussians, they are Rayleigh distributed with probability density function (PDF)

$$f_{|\eta[k, a]|}(x) = \frac{2x}{\sigma_\eta^2}\exp\left(-\frac{x^2}{\sigma_\eta^2}\right). \tag{13}$$

The two terms in (12) are independent. By using the transformation theorem on their joint distribution, the PDF of the timing signal can be calculated as

$$\begin{aligned} f_{|\lambda[k, a]|}(x) &= \left(\frac{2}{\sigma_\eta^2}\right)^2 \int_0^\infty \frac{x}{b}\exp\left(-\frac{1}{\sigma_\eta^2}\left(b^2 + \frac{x^2}{b^2}\right)\right)db \\ &= \left(\frac{2}{\sigma_\eta^2}\right)^2 x\mathrm{k}_0\left(2\left(\frac{x}{\sigma_\eta^2}\right)\right), \end{aligned} \tag{14}$$

where $\mathrm{k}_0(x)$ denotes the zeroth order modified Bessel function. The CDF of the timing signal is, consequently, given by

$$F_{|\lambda[k,\,a]|}(\alpha) = \int_0^{\alpha} f_{|\lambda[k,\,a]|}(x)dx = 1 - 2\frac{\alpha}{\sigma_\eta^2}\mathrm{k}_1\left(2\frac{\alpha}{\sigma_\eta^2}\right) \tag{15}$$
$$= 1 - 2\frac{\alpha K\sigma_s^2}{\sigma_n^2}\mathrm{k}_1\left(2\frac{\alpha K\sigma_s^2}{\sigma_n^2}\right).$$

For L=2 we are not able to derive exact expressions for the distributions of the timing metric, but we determine bounds for the CDFs instead. By an upper bound we mean that if the CDF of A is upper bounded by the CDF of B then $F_A(x) \le F_B(x)$. The opposite is true for a lower bound. In appendix A we show that the CDF of $|\lambda[k,\,a]|$ is lower bounded by the CDF of an Erlang-2 variable X_λ,

$$F_{X_\lambda}(x) = 1 - \left(1 + \frac{\sqrt{2}}{\sigma_\eta^2}x\right)\exp\left(-\frac{\sqrt{2}}{\sigma_\eta^2}x\right). \tag{16}$$

We also show that the CDF of the timing signal $|\lambda[k,\,a]|$ when L=2 is upper bounded by the CDF of the maximum, Y_λ, of two independent exponential variables,

$$F_{Y_\lambda}(y) = \left(1 - \exp\left(-\frac{\sqrt{2}}{\sigma_\eta^2}y\right)\right)^2. \tag{17}$$

When the number of terms, L, in the sum is large, the distribution of the timing signal can be calculated in the same way for the pre-correlating synchronizer and the conventional synchronizer. In such cases, the distribution of the real and the imaginary part of the timing signal can be approximated as Gaussian due to the central limit theorem. The real and imaginary part of the normalized timing signal, $Re\{\lambda_{conv}\}$ and $Im\{\lambda_{conv}\}$, both have zero mean and a variance

$$\sigma^2_{Re\{\lambda_{conv}\}} = L(\sigma_\eta^2)^2/2. \tag{18}$$

The normalized conventional timing signal can therefore be regarded as Rayleigh distributed with CDF

$$F_{|\lambda_{conv}[k,\,a]|}(x) = 1 - \exp\left(-\frac{x^2}{2\sigma^2_{Re\{\lambda_{conv}\}}}\right) = 1 - \exp\left(-x^2K^2L\left(\frac{\sigma_s^2}{\sigma_n^2}\right)^2\right). \tag{19}$$

The timing signal from the pre-correlating synchronizer is generally smaller compared to the same signal from the conventional synchronizer when the timing is wrong. This in turn means that the false detection probability of the pre-correlating

synchronizer is much lower than that of the conventional synchronizer. In Figure 4 the CDFs of the timing signals from the pre-correlating synchronizer and the and conventional synchronizer are presented for the cases of correct and wrong timing and no frequency-offset. By correlating samples before multiplication the mean and variance of the timing signal are decreased when the timing is wrong. This makes it possible to use a low detection threshold without a resulting large false detection probability. Although in the figure there is no difference between the detection performance of the PN-based preamble and the conventional preamble, there is an advantage for the PN-based preamble when analyzing the case when the timing is almost, but not entirely, correct. The correlation properties are better for the PN-based preamble and it results a sharp synchronization peak. This results in more accurate time synchronization and this property is especially important in multipath channels.

3.3 Low-Resolution Operation

There is a high correlation between the sign of the received signal and the sign of the pre-correlating sequence when the timing is correct. The correlation is zero when the timing is wrong and therefore it is possible to have a low-resolution operation mode where only the sign of the received signal is used to find a coarse timing signal. The probability that the sign of the received sample is equal to the sign of the pre-correlating sequence can, for the I-channel and the Q-channel respectively, be calculated as

$$\begin{aligned} p_{I,\,equal}[k] &= 1 - Q\left(\sqrt{2(\sigma_s^2/\sigma_n^2)}\cos\left(\frac{2\pi\varepsilon_c k}{M} + \theta_c\right)\right) \\ p_{Q,\,equal}[k] &= 1 - Q\left(\sqrt{2(\sigma_s^2/\sigma_n^2)}\sin\left(\frac{2\pi\varepsilon_c k}{M} + \theta_c\right)\right), \end{aligned} \tag{20}$$

when the timing is correct. Note that if the phase of the received signal is between π and 2π then the probability of equal signs in e.g. the quadrature channel is lower than 0.5. However, the same is true for the second part of the preamble so that the product will in general be positive anyway. The correlator adds all the equal and unequal samples and therefore the two complex parts of the correlator output, $\eta_{low-res}$, have a binomial distribution. The probability that the real part of pre-correlator has n_I samples of equal signs is given by

$$P_{Re(\eta_{low-res}),\,equal}(n_I) = \binom{K}{n_I}(p_{I,\,equal})^{n_I}(1 - p_{I,\,equal})^{K-n_I} \tag{21}$$

with a corresponding expression for the imaginary part. For small frequency offsets and correct timing the real part of the synchronization signal is much larger than the imaginary part, so the absolute value can be approximated by the real part only. The

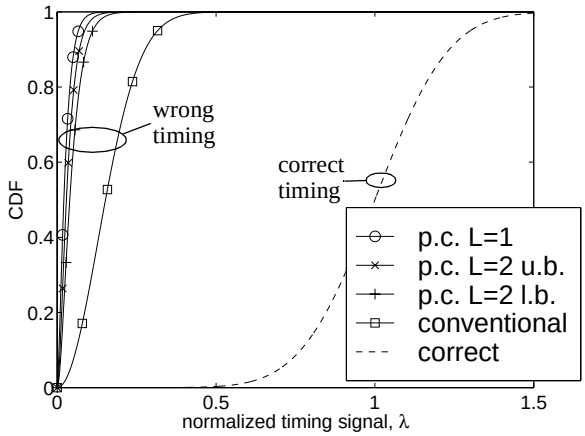

Figure 4: CDFs of the timing signals when using the conventional synchronizer or the pre-correlating synchronizer (p.c.) with L=1 or L=2. In case of L=2 and wrong timing we show an upper (u.b.) and lower bound (l.b.) for the distribution.

real part of the synchronization signal is the sum of the products between the two real parts of the correlator outputs and the two imaginary parts of the correlator outputs. The correlator output is given by $\eta_{low-res} = 2n_I - K + i(2n_Q - K)$. The probability function of the products is calculated by adding all the probabilities for values resulting in a certain product and the probability function, $P_{Re(\gamma_{low-res})}(n)$, of the sum is then achieved by convolution. For larger frequency offsets we have a similar distribution, but the signal is not dominated by the real part.

When the timing is wrong, the input can be seen as random and the probability that the sign of the received sequence is equal to the sign of the pre-correlating sequence is $P_{I,equal}$=$P_{Q,equal}$=50% for the I and Q-channel, respectively. In this case none of the parts dominate and we derive bounds for the distribution. The absolute value of the synchronization signal is larger than the maximum absolute value of the real and imaginary part of the synchronization signal and smaller than the sum of the amplitudes of the real and imaginary part respectively. The CDF of the synchronization signal is therefore upper bounded by the CDF of the maximum absolute value of the real and imaginary part and lower bounded by the CDF of the sum of the absolute values of the real and imaginary parts, i.e.

$$\begin{aligned} &F_{|Re(\gamma_{low-res})| + |Im(\gamma_{low-res})|}(n) \le F_{|\gamma_{low-res}|}(n) \\ &\le F_{max(|Re(\gamma_{low-res})|, |Im(\gamma_{low-res})|)}(n). \end{aligned} \tag{22}$$

Due to the favorable correlation properties, the low-resolution timing signal works well also for low SNRs in AWGN channels. In Figure 5 we plot the CDF of

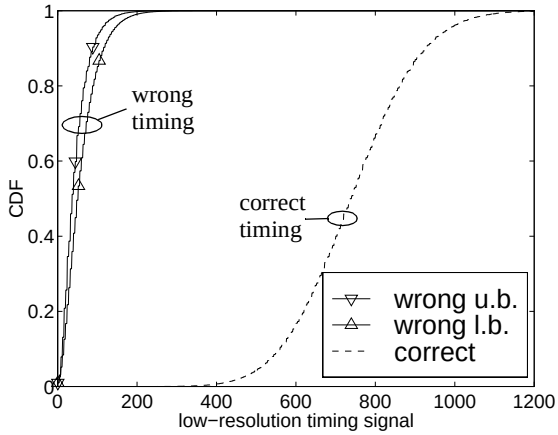

Figure 5: CDFs of the low-resolution timing signal. For wrong timing we have a lower and an upper bound for the distribution.

the low-resolution timing signal using the PN-based preamble and the same parameters as in Figure 4. For wrong timing we show an upper and a lower bound for the distribution. In the figure it can be seen that the probability is very small that the timing signal is larger when the timing is wrong than when it is correct.

3.4 Receiver Operation Characteristics

For the timing signals there is a relationship between the probability of missing the synchronization signal and the probability of false detection, which is dependent on the chosen detection threshold. We plot this relationship, the so-called receiver operation characteristics, in Figure 6 for the conventional and pre-correlated timing signal together with an upper and lower bound for the low-resolution timing signal. Note that we plot the probability of missed synchronization signal and not, as most often, the probability of detection vs. the probability of false detection. We see in the figure that the timing signal from the pre-correlating synchronizer results in much better performance, compared to the conventional synchronizer. For a probability of missed synchronization signal of 10^{-4} we get a false detection probability of $3*10^{-2}$ for the conventional synchronizer while we get $2*10^{-8}$ for the pre-correlating synchronizer. The lower variance for the pre-correlated timing signal when the timing is wrong also means that it is possible to decrease the threshold for detection without having too many false alarms. The low-resolution synchronization signal also results in good detection performance, in AWGN channels it competes well with the conventional full-resolution synchronizer. When the probability of missed synchronization signal is 10^{-4} the false detection probability from the low-resolution synchronization signal

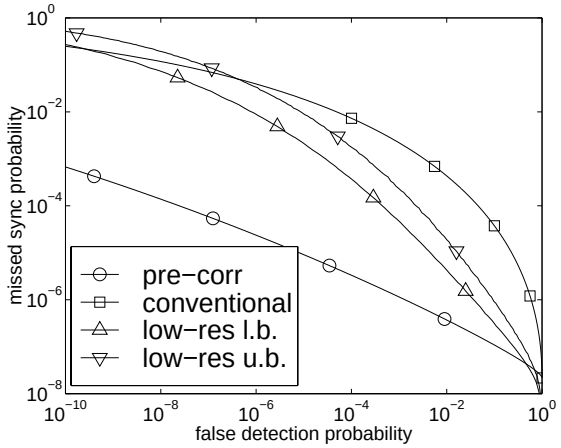

Figure 6: Receiver operation characteristic for the conventional synchronizer, the pre-correlating synchronizer and the low-resolution synchronizer. For the latter one we show an upper and lower bound.

is in the order of 10^{-3}. Compare this to the false detection probability of the conventional, but full resolution, synchronization signal that for these parameters is $3*10^{-2}$. In [10] a synchronizer based on the cyclic prefix using an A/D-converter with one-bit quantization is analyzed. This system is similar to conventional preamble synchronization but the repetition due to the cyclic prefix is used instead of the repetition in the preamble. There the low-resolution synchronizer gave a significant loss compared to conventional full-resolution synchronization. Here the performance of the low-resolution synchronizer is actually better than that of the conventional full resolution synchronizer.

3.5 Frequency-Offset Estimation

The frequency-offset estimate is given by the phase of the synchronization signal when the correct timing is found. The estimate is calculated as

$$\hat{\varepsilon}_c = \frac{M}{(2\pi PR)} \arg(\lambda[k, 0]) . \tag{23}$$

This phase increment method gives a (nearly) unbiased frequency estimate of a complex exponential in white noise [11] and can cope with frequency-offsets which are smaller than $|\varepsilon_c| < M/(2PR)$.

The variance of the normalized frequency estimate can be approximated as [11]

$$var(\hat{\varepsilon}_c) \tag{24}$$
$$= \frac{M^2}{(4\pi^2 KLPR^2)} \cdot \left(\frac{1}{L} \cdot \frac{\sigma_n^2}{\sigma_s^2 \mathrm{sinc}^2(\varepsilon_c K/M)} + \frac{1}{2KP} \cdot \left(\frac{\sigma_n^2}{\sigma_s^2 \mathrm{sinc}^2(\varepsilon_c K/M)} \right)^2 \right).$$

The performance of the frequency-offset estimate is similar for the conventional and the pre-correlating synchronizer. For low SNRs the variance from the pre-correlating synchronizer is slightly lower. In this case the effective SNR after the pre-correlator is often high enough that the second term in (24) can be neglected. In Figure 7 we plot the theoretical standard deviation of the frequency-offset estimation error together with simulated results for the conventional synchronizer using the OFDM-based preamble and the pre-correlating synchronizer using the PN-based preamble. The preamble length was 60 (K=30, L=1) for the PN-based preamble and 64 (K=1, L=32) for the OFDM-based one in an AWGN channel with SNR 0 10 and 20 dB, respectively. The simulated values agree well with the theoretical values. For the pre-correlating synchronizer at large frequency offsets it is possible to observe the loss due to the sinc-term in (24). However, as seen in the figure this increase in the standard deviation is small.

It is possible to use the low-resolution synchronization signal to derive an initial frequency-offset estimate. Simulations show, however, that the loss compared to full-resolution estimation is large, especially for high SNRs. For low SNRs the increase in standard deviation is smaller due to the averaging effect from the noise.

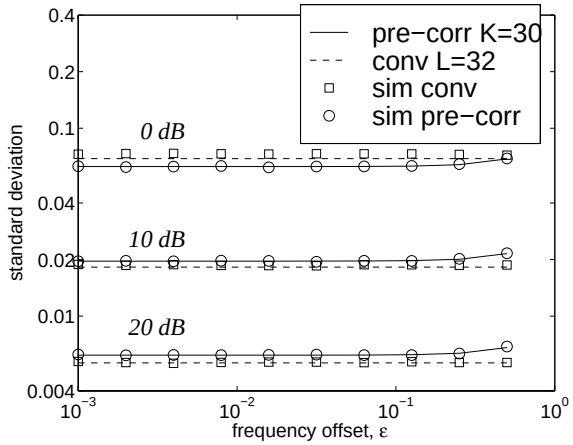

Figure 7: Theoretical and simulated values of the standard deviation of the frequency-offset estimation error. AWGN channel with SNR 0, 10 and 20 dB. Note that the OFDM-based preamble is slightly longer than the PN-based one.

4 Performance in Fading Channels

In multipath channels the timing signal from the pre-correlating synchronizer has a peak for each of the resolvable channel taps. The correlation peaks have the same phase, except for a small distortion due to the noise, and their amplitudes correspond to the tap-power for each delay. For time synchronization the synchronizer can either use the largest peak, a sum of the largest peaks or a sum of all peaks within a time window corresponding to the excess delay of the channel to find the correct timing.

In OFDM systems we want a timing signal that has a peak when the largest part of the channel impulse response is within the window set by the guard interval. This is straightforward to achieve with a rectangular window with length as the guard interval. To get sharp synchronization peaks and to minimize the influence between peaks it is important to use a sequence with good autocorrelation properties. Minimizing the influence is especially important for the frequency-offset estimate, otherwise there will be a small increase in the standard deviation of the estimation error. The PN-based preamble used together with the pre-correlating synchronizer makes it possible to resolve taps spaced one chip (i.e. two samples) apart.

In order to investigate the performance in multipath and fading channels we simulated a one-tap and a two-tap Rayleigh fading channel. In case of the two-tap channel the taps had equal mean power and the delay between them was varied. In Figure 8 we present simulated receiver operation characteristic in a one-tap rayleigh fading channel for the low-resolution synchronizer, the conventional synchronizer using the OFDM-based preamble and the pre-correlating synchronizer using the PN-based and OFDM-based preamble. The preamble length was 60 for the PN-based preamble (K=30, L=1) and 64 for the OFDM-based preamble (K=1, L=32). The OFDM-based preamble was also designed to give good correlation properties by choosing the data symbols properly, the mean SNR was 10 dB and a constant channel during the whole preamble was assumed.

As before, we compare the timing signals when the timing is correct and when the timing is wrong so that the input to the synchronizer in the latter case is noise only. The pre-correlating synchronizer in this and the next figure used the largest peak and not a sum to find the timing. In Figure 9 we present simulated receiver operation characteristic for the two-tap Rayleigh fading channel. The delay between the taps was 4 samples such that we got two fully resolvable Rayleigh fading taps when using the PN-based preamble with the pre-correlating synchronizer. Due to the delay the preambles were extended by 4 samples. For a probability of missed synchronization signal of 10^{-3} the probability of false detection is here $2*10^{-1}$ for the conventional synchronizer, 10^{-2} for the low-resolution synchronizer and $7*10^{-4}$ for the pre-correlating synchronizer, respectively.

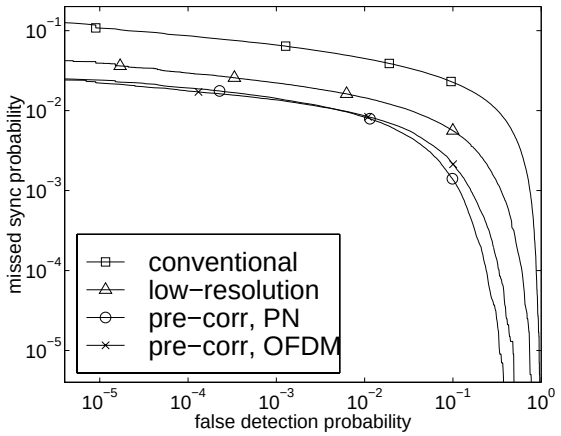

Figure 8: Simulated receiver operation characteristics for the different synchronizers and preambles. One-tap Rayleigh fading channel, mean SNR 10 dB, preamble length 60 (PN-based preamble) or 64 (OFDM-based preamble).

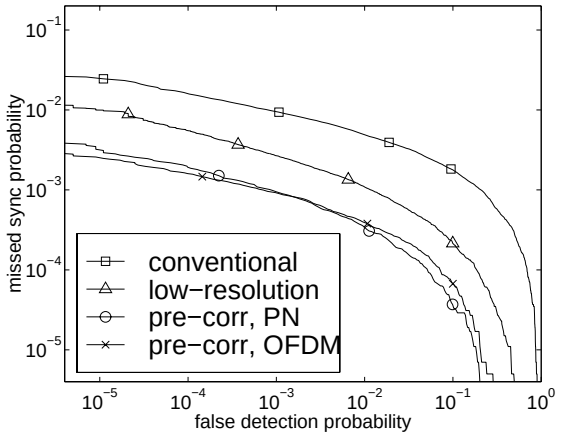

Figure 9: Simulated receiver operation characteristics for the different synchronizers and preambles. Two-tap Rayleigh fading channel with equal power and 4 samples delay, mean SNR 10 dB, preamble length 64 (PN-based preamble) or 68 (OFDM-based preamble) including 4 samples of cyclic extension.

The pre-correlating synchronizer gives better detection performance both for the one-tap and the tow-tap channel. The reason is again that the mean amplitude is lower for that timing signal when the timing is wrong. The low-resolution synchronizer also performs well compared to the conventional one. It also gives lower probability of false detection and lower probability of missing the synchronization signal. When using the pre-correlating synchronizer the difference between the OFDM-based and the PN-based preambles is small. Both preambles have good autocorrelation properties and therefore their performance is similar. The PN-based preamble has a slightly better performance when the threshold is low and the false detection probability thereby is high. This is because the received samples are weighted equally for the PN-based preamble while some samples have higher weight for the OFDM-based preamble. For short preambles this implies that the variance of the timing signal is slightly smaller when the timing is wrong.

In a real application a two-step procedure for synchronization in Rayleigh fading channels can be used. The PN-sequence is repeated four times such that the preamble length becomes $4R+G$ samples. During the first $2R$ samples the low-resolution mode is used to find a coarse timing signal where the received power is unknown. In low-resolution mode the synchronizer has low complexity and therefore the stand by mode power consumption becomes low. At the same time as the coarse synchronization signal is calculated the AGC is set. In the second step, during the following samples, fine time synchronization is performed and a frequency-offset estimate is derived. In this part the synchronization signal is normalized by the received power.

We evaluate the performance of the frequency-offset estimate by comparing the standard deviation of the estimation error. In Figure 10 the standard deviation is presented for different frequency offsets using the conventional synchronizer with the OFDM-based preamble or the pre-correlating synchronizer with the PN-based preamble. It should be noted that also in this example only the maximum peak is used for estimation. A slightly lower standard deviation for the pre-correlating synchronizer is achieved by using more peaks. The simulations showed that the performance is similar for both preambles and that the standard deviation is in principle independent of the actual frequency offset as long as it is within the permitted range. The simulations also showed that outliers caused by fading dips dominate the standard deviation of the frequency-offset estimate.

5 Conclusions

We have analyzed the performance of different preamble-based synchronization schemes in OFDM systems. We analyzed two synchronizer structures, the conventional OFDM synchronizer and the pre-correlating synchronizer, used together with a conventional preamble or a PN-based preamble. It was shown that the pre-correlating

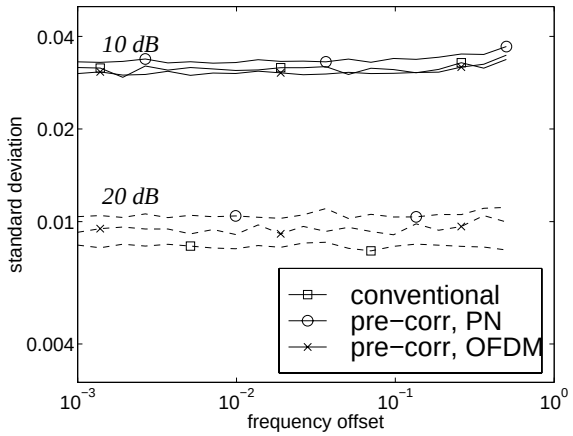

Figure 10: Simulated standard deviation of the frequency-offset estimation error. Rayleigh fading channel with two Rayleigh fading taps. Mean SNR 10 (solid) and 20 dB (dashed), preamble length 64 (PN-based preamble) or 68 (OFDM-based preamble) including 4 samples of cyclic extension.

synchronizer results in better detection performance than the conventional synchronizer. By summing received samples coherently before multiplication, as in the pre-correlating synchronizer, it is possible to decrease the influence from noise or interference when the timing is wrong. Therefore a lower detection threshold can be used, which results in better detection performance in terms of lower false detection probability and lower probability of missing the synchronization signal. By using the PN-based preamble the peak-to-average power ratio of the transmitted signal becomes low and the timing signal from the pre-correlating synchronizer shows a sharp synchronization peak. This is in contrast to conventional synchronization, which is based on repeated OFDM symbols. For frequency-offset estimation the different methods result in similar performance. Synchronization based on PN-sequence preambles and the pre-correlating synchronizer offers great power reductions in stand-by mode. Using this method it is possible to use A/D-converters with only one-bit resolution in stand-by mode. The low-resolution synchronization signal is independent of the received power and therefore a coarse time synchronization signal can be calculated at the same time as the AGC level is set.

Appendix A. Distribution of the Timing Signal, Wrong Timing, *L*=2

If L is even, the synchronization signal can be rewritten as

$$\lambda[k, a] = \sum_{l=0}^{L-1} \eta[k - lR, a]\eta^*[k - (l + P)R, a] \tag{25}$$
$$= \sum_{k=0}^{L/2-1} d_{2k+1}c_{2k+1} + b_{2k+1}e_{2k+1}.$$

All the terms (b, c, d and e) in (25) are independent zero mean Gaussian variables; d and b are real valued and have variance $\sigma_b^2 = \sigma_d^2 = E[bb] = \sigma_\eta^2/2$, whereas c and e are complex valued with variance of the both complex parts $\sigma_{Re\{c\}}^2 = \sigma_{Im\{c\}}^2 = \sigma_\eta^2$. Defining new variables as $g_k = d_{2k+1}Re\{c_{2k+1}\}$ and $h_k = d_{2k+1}$, the joint probability function for g and h becomes

$$f_{G,H}(g, h) = \frac{1}{2\pi\sigma_{Re\{c\}}\sigma_d}\exp\left(-\frac{1}{\sigma_{Re\{c\}}^2}\left(\frac{g}{h}\right)^2 - \frac{1}{\sigma_d^2}h^2\right)\frac{1}{h}. \tag{26}$$

Integrating [13] (3.478.4) over h gives the PDF for the products, i.e.

$$f_G(g) = \int_{-\infty}^{\infty} f_{G,H}(g, h)dh = \frac{1}{\pi\sigma_{Re\{c\}}\sigma_d}\mathrm{k}_0\left(\frac{|g|}{\sigma_{Re\{c\}}\sigma_d}\right) \tag{27}$$
$$= \frac{\sqrt{2}}{\pi\sigma_\eta^2}\mathrm{k}_0\left(\frac{\sqrt{2}|g|}{\sigma_\eta^2}\right).$$

Naturally this PDF also applies for $b_{2k+1}e_{2k+1}$ and for the corresponding imaginary parts. The characteristic function of the product, g, can be calculated as [13] (6.671.14)

$$\phi_g(\omega) = \int_{-\infty}^{\infty} f_G(g)e^{i\omega g}dg = \frac{\sqrt{2}}{\pi\sigma_{\eta_a}^2}\int_{-\infty}^{\infty}\mathrm{k}_0\left(\frac{\sqrt{2}|g|}{\sigma_\eta^2}\right)e^{i\omega g}dg \tag{28}$$
$$= \frac{\sqrt{2}}{\sigma_\eta^2}\frac{1}{\sqrt{(\sqrt{2}/\sigma_\eta^2)^2 + \omega^2}}.$$

The PDF of the sum of products, i.e. the real part of λ, is given by the inverse transform of the product between the L characteristic functions,

$$f_{Re\{\lambda\}}(x) = \frac{1}{2\pi}\left(\frac{\sqrt{2}}{\sigma_\eta^2}\right)^L \int_{-\infty}^{\infty} \left(\left(\frac{\sqrt{2}}{\sigma_\eta^2}\right)^2 + \omega^2\right)^{-L/2} e^{-i\omega x} d\omega. \tag{29}$$

For L=2, we get a Laplacian random variable with PDF [13] (3.389.5)

$$f_{Re\{\lambda\}}(x) = \frac{1}{\sqrt{2}\sigma_\eta^2} \exp\left(-\frac{\sqrt{2}|x|}{\sigma_\eta^2}\right). \tag{30}$$

$|Re\{\lambda[k, a]\}|$ and $|Im\{\lambda[k, a]\}|$ are therefore independent exponential variables and their sum is an Erlang-2 distributed variable. This means that the CDF of $|\lambda[k, a]|$ when L=2 is lower bounded by the CDF of an Erlang-2 variable X_λ,

$$F_{X_\lambda}(x) = 1 - \left(1 + \frac{\sqrt{2}}{\sigma_\eta^2}x\right)\exp\left(-\frac{\sqrt{2}}{\sigma_\eta^2}x\right). \tag{31}$$

The CDF of $|\lambda[k, a]|$ when L=2 is upper bounded by the CDF of the maximum, Y_λ, of two independent exponential variables,

$$F_{Y_\lambda}(y) = \left(1 - \exp\left(-\frac{\sqrt{2}}{\sigma_\eta^2}y\right)\right)^2. \tag{32}$$

6 Acknowledgment

The authors would like to thank Prof. Mike Faulkner, Victoria University of Technology, Australia, for helpful discussions and suggestions.

References

[1] Institute of Electrical and Electronics Engineers, IEEE P802.11a/D7.0, IEEE, New York, USA, July 1999

[2] MMAC system architecture, Ver. 1: MMAC Promotion Council, Tokyo, Japan, March 1998.

[3] European Telecommunications Standards Institute, Broadband Radio Access

Networks (BRAN); HIPERLAN Type 2 Functional Specification Part 1 - Physical (PHY) layer. <DTS/BRAN030003-1> V0.j (1999-09), ETSI, Sophia Antipolis, France, Sept. 1999.

[4] P. Moose, “A technique for orthogonal frequency division multiplexing frequency offset correction”, IEEE Trans. on Communications, Vol. 42, no. 10, Oct. 1994, pp. 2908-14.

[5] T.M. Schmidl and D.C. Cox, “Robust frequency and timing synchronization for OFDM”, IEEE Trans. on Communications, vol. 45, no. 12, Dec. 1997, pp. 1616-1621.

[6] S.H. Müller-Weinfurtner, J.F. Rößler and J.B. Huber, “Analysis of a frame- and frequency synchronizer for (bursty) OFDM”, Proc. of IEEE Communications Theory Mini Conference in conjunction to Globecom’98, Sydney, Australia, Nov. 1998, pp. 201-206.

[7] F. Platbrood, S. Rievers and J. Farserotu, “Analysis of coarse frequency synchronization for Hiperlan Type-2”, Proc. of IEEE Vehicular Technology Conference, Amsterdam, The Netherlands, September 1999, pp. 688-692.

[8] K. Bruninghaus; M. Radimirsch, “Coarse frame synchronisation for OFDM based wireless communication systems”, IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, 1998, pp. 806-810.

[9] N. Nishinaga; M. Nakagami; Y. Iwadare, “A simple synchronization acquisition method for DS/SS system under carrier frequency offset“, IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, vol.E80-A, no.11; Nov. 1997, pp. 2162-2171.

[10] J.J van de Beek, M. Sandell, M. Isaksson and P.O. Börjesson, “Low-complex frame synchronization in OFDM systems”, Proc. of 4th International Conference on Universal Personal Communications, Tokyo, Japan, Nov. 1995, pp. 982-986.

[11] H. Meyr and M. Moeneclaey, “Digital Communication Receivers”, John Wiley & Sons, New York, U.S.A, 1998.

[12] U. Fawer, “A coherent spread-spectrum diversity receiver with AFC for multipath fading channels”, IEEE Trans. on Communications, vol. 42 no. 2/3/4, Feb-April 1994, pp. 1300-1311.

[13] I.S. Gradshteyn And I.M. Ryzhik, “Table Of Integrals, Series, And Products”, Alan Jeffrey, Editor, Academic Press, London, UK, 1994.

Paper E

This part has been published as:

Fredrik Tufvesson, Mike Faulkner, Peter Hoeher and Ove Edfors, "OFDM Time and Frequency Synchronization by Spread Spectrum Pilot Technique", 8th IEEE Communication Theory Mini Conference in conjunction to ICC'99, Vancouver, Canada, June 1999, pp. 115-119.

OFDM Time and Frequency Synchronization by Spread Spectrum Pilot Technique

Fredrik Tufvesson[1], Mike Faulkner[2], Peter Hoeher[3] and Ove Edfors[1]

[1]Department of Applied Electronics, Lund University, Box 118, SE-221 00 Lund, Sweden, e-mail: Fredrik Tufvesson@tde.lth.se

[2]School of Communications and Informatics, Victoria University of Technology, Melbourne, Australia

[3]Information and Coding Theory Lab, University of Kiel, Kiel, Germany

Abstract

Correct time and frequency synchronization is important for the performance of orthogonal frequency division multiplex (OFDM) systems. We evaluate a system where a pseudo noise sequence (PN-sequence) is used for estimation of these synchronization parameters. The PN-sequence is superimposed on the OFDM signal. The system is evaluated by means of the variance of the frequency estimation error and the probability for correct timing synchronization. Both theoretical and simulated results are presented. The proposed technique is very flexible and works also at low signal to noise ratios. The frequency offset range has been significantly increased compared to conventional OFDM synchronizer schemes.

1 Introduction

Time and frequency synchronization are two important issues for the performance of orthogonal frequency division multiplex (OFDM) systems. Correct frequency synchronization is crucial in order to preserve orthogonality of the sub-channels. Timing synchronization is necessary in order to locate the cyclic prefix and identify the start of new packets or frames. Often the synchronization process is divided into two parts, acquisition and tracking. Burst transmission acquisition includes a continuous search for new packets, finding the beginning of the symbols as well as the packets and making a coarse frequency estimate so that demodulation can be performed. For continuous transmission, acquisition means finding the start of symbols and frames and making a coarse frequency estimate. In tracking mode the synchronization parameters are continuously updated to keep the performance of the demodulator as good as possible under changing conditions.

Existing algorithms are often based on correlation of either a repeated [1][2][3] or a known [4][5] symbol. A correlation peak is achieved when the timing is correct and then the frequency offset estimate is achieved by the phase shift between certain samples. The distance between these samples determines the frequency offset range. When, e.g., the cyclic prefix is used as a repeated symbol the frequency offset range is limited to half the sub-channel spacing.

This paper presents a new method for acquisition and tracking of OFDM synchronization parameters, which is based on a continuous transmission of a pseudo noise (PN) sequence. Similar techniques has previously been proposed for synchronization of single carrier systems [6] and for frame synchronization in OFDM systems [7]. In this paper we propose and investigate a system for frequency and timing synchronization, including both symbol and frame synchronization. The system can cope with very large offsets, up to half the OFDM bandwidth, and gives the frequency estimate in one step. The synchronization signal has a sharp peak and is independent of the OFDM parameters used. It can be used for any length of the cyclic prefix and does not increase the transmitted bandwidth.

2 Spread Spectrum Pilot Technique

The estimation of synchronization parameters is based on transmission of a PN-sequence which is *superimposed* on the OFDM information signal. The chip time of the PN-sequence is equal to or double the sample time in the OFDM system, depending on spectrum and sampling requirements. The sequences are synchronous and transmitted in the same band. Figure 1 shows a block description of how the transmitted signal is generated. During acquisition the PN-sequence is sent without transmission of the information signal. After one or more PN-symbols, when the correct timing and carrier frequency have been found, the transmission of the information signal starts and the receiver enters the tracking mode, see Figure 2. In tracking mode the amplitude of the PN-sequence can be chosen so that the distortion of the information symbol is small. The amplitude of the PN-sequence can be adjusted to a specific application and the synchronization scheme works for low signal to noise ratios. It is desirable to use a code length equal to or a multiple of the OFDM symbol length (including the cyclic prefix) so that correlation peaks occur at the symbol boundaries.

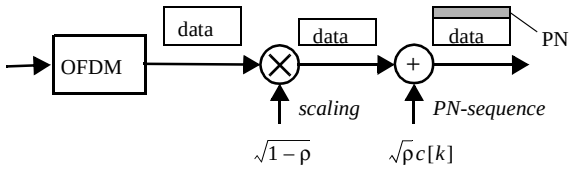

Figure 1: Generation of the transmitted signal

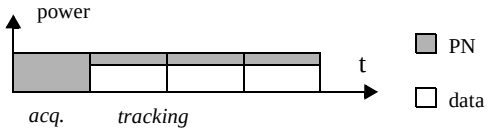

Figure 2: Structure of the transmitted signal.

For frame synchronization it is sometimes advantageous to use a long code.

3 System Description

In the analysis we assume an OFDM system with M sub-channels, having a symbol time T_s in an AWGN channel. The discrete time representation of the received base band signal is

$$\begin{aligned} r[k] &= s[k] + n'[k] \\ &= (\sqrt{\rho}\sigma_s c[k] + \sqrt{1-\rho}\sigma_s d[k]) e^{j\left(\frac{2\pi\varepsilon_c k}{M} + \theta_c\right)} + n'[k]. \end{aligned} \tag{1}$$

where $c[k]$ is the transmitted PN-sequence, $d[k]$ is the OFDM data sequence and $n'[k]$ is white Gaussian noise with spectral density N_0. Further, θ_c is the carrier phase and $\varepsilon_c = \Delta f \cdot T_s$ denotes the frequency offset to be estimated when normalized by the sub-channel spacing. Both the PN-sequence and the data sequence has unity power and the amount of power used for the PN-sequence is controlled by the code power ratio ρ. The received power is given by $\sigma_s^2 \equiv E_s / T_s = E[|s[k]|^2]$, whereas the noise power is given by $\sigma_n^2 \equiv (N_0 / T_s) = E[|n[k]|^2]$. The received signal is despread by the local PN-sequence to get

$$\begin{aligned} c^*[k-a]r[k] &= \sqrt{\rho}\sigma_s c^*[k-a]c[k] e^{j(2\pi\varepsilon_c k/M + \theta_c)} \\ &+ \sqrt{1-\rho}\sigma_s d[k]c^*[k-a] e^{j(2\pi\varepsilon_c k/M + \theta_c)} + n[k]. \end{aligned} \tag{2}$$

For time synchronization we want to find the integer slip, a, between the two codes. For the purpose of time and frequency estimation, the two latter terms in (2) are acting as disturbances. For data demodulation, no code multiplication is performed in the receiver and the first and the last terms cause disturbances. The resulting SNR becomes $\rho\sigma_s^2/((1-\rho)\sigma_s^2 + \sigma_n^2)$ for the code sequence and $(1-\rho)\sigma_s^2/(\rho\sigma_s^2 + \sigma_n^2)$ for the data sequence. In acquisition mode, only the code sequence is sent and the code power ratio ρ equals one. In tracking mode a low code power ratio is then used in order to minimize the impact on data transmission.

The despread signals are summed in groups of K samples to get a so called sub-correlation. The synchronization signal, $\gamma[k, a]$, for a given slip, a, is then achieved as the sum of L products between sub-correlations spaced KP samples apart (P is the delay in number of correlator lengths),

$$\gamma[k, a] = \sum_{l=0}^{L-1} \left[\left(\sum_{n=0}^{K-1} c^*[k-n-lK-a]\, r[k-n-lK] \right) \left(\sum_{n=0}^{K-1} c^*[k-n-(l+P)K-a]\, r[k-n-(l+P)K] \right)^* \right]. \quad (3)$$

Time synchronization is given by the absolute value of this signal whereas the frequency offset estimate is given by the phase. The synchronizer structure is given in Figure 3. This scheme can be seen as a modification of the schemes in [4] and [5]. Here, the length of the sub-correlators is K instead of 1 as in [5], and the differential detectors work on the sub-correlator outputs instead of the correlator outputs as in [4].

4 Timing Estimator Performance

In this section we analyze the distribution of the timing metric in an AWGN channel. Firstly we analyze the case when the timing is correct and later when it is wrong. The calculations are quite tedious and therefore we just summarize the results and verify them by means of simulations. Often it is desirable to have a signal which is independent of received power. By normalizing γ as $\lambda[k, a] = \gamma[k, a]/(LK^2\rho\sigma_s^2)$ we get a timing metric with a mean close to one when the timing is correct and just above zero when it is wrong.

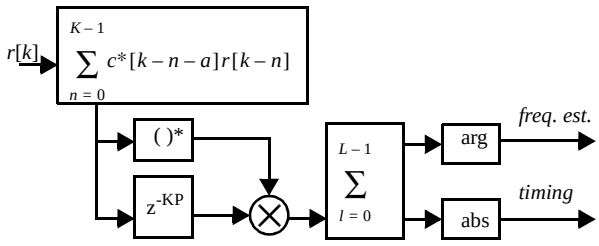

Figure 3: Block diagram of the estimator.

4.1 Correct Timing

When the timing is correct, a=0, the distribution of the timing metric can be analyzed in a similar way as in [3]. The frequency offset entails that the chips are not added totally in phase in the sub-correlator, which leads to an energy reduction. We assume that no ISI is present and $|\varepsilon_c| \ll M$. Normalizing the sub-correlator signal by $\sqrt{LK^2\rho\sigma_s^2}$ gives the normalized output, $\eta[k, 0]$, where [4][1]

$$\eta[k, 0] = \frac{1}{\sqrt{L}}\mathrm{sinc}(\varepsilon_c K/M)\exp\left(j\frac{2\pi\varepsilon_c}{M}\left(k+\frac{1}{2}\right)\right) + n_\eta[k, 0]. \tag{4}$$

All differential products then add coherently since they have nearly the same phase. For medium and high SNR:s, the contribution from the noise vector perpendicular to the signal vector can be neglected. If we consider the data signal as Gaussian, or if K is large, $|\lambda[k, 0]|$ can be approximated as a Gaussian variable with mean

$$E[|\lambda[k, 0]|] = \mathrm{sinc}^2(\varepsilon_c K/M) \tag{5}$$

and variance

$$var[\lambda[k, 0]] = \frac{((1-\rho)\sigma_s^2 + \sigma_n^2)\mathrm{sinc}^2(\varepsilon_c K/M)}{LK\rho\sigma_s^2}. \tag{6}$$

As seen in (6) the variance is inversely proportional to LK. The particular choice of correlator length has small impact on this product if the observation interval is kept constant. The probability of missing the sync signal is therefore mainly determined by the observation interval and the equivalent SNR for the PN-sequence.

4.2 Incorrect Timing

When the timing is wrong, $a \neq 0$, the sub-correlators do not peak. The sub-correlator outputs, $\eta[k, a]$, can be regarded as Gaussian variables with zero mean. The timing metric is therefore the absolute value of a sum of products between zero mean Gaussian variables,

$$|\lambda[k, a]|_{a\neq 0} = \left|\sum_{l=0}^{L-1} \eta_a[k-lK, a]\eta_a{}^*[k-(l+P)K, a]\right|. \tag{7}$$

1 A typo in the original publication has been corrected, the square root was missing.

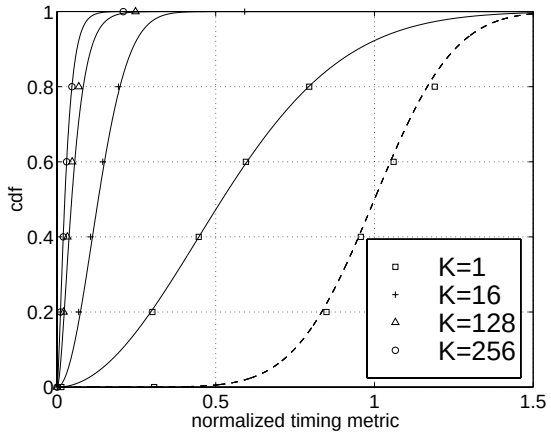

Figure 4: CDF for the normalized timing metric when the timing is wrong (solid) and correct (dashed), SNR=10 dB, ρ =0.1, observation interval 512 samples and $(K, L) = \{(1, 256), (16, 16), (128, 2), (256, 1)\}$. Marker values are obtained by simulations, whereas the lines represent theoretical values.

In appendix A the distribution of the timing metric is analyzed for different choices of the number of differential signals. In Figure 4 the cumulative distribution function (CDF) of the timing metric is presented for various correlator lengths, *K*, and number of differential signals, *L*, when the observation interval is 512 samples, SNR=10dB and ρ =0.1.

4.3 Receiver Operation Characteristics

Depending on how the detection threshold is set we will get a specific probability of missing the sync signal and a probability of false alarm. In Figure 5 we plot the receiver operation characteristics (ROC) for different parameter choices. It is of course desirable to be in the lower left corner where both probabilities are low. Here, again, we see the importance of using long sub-correlator lengths. By correlating samples before differential detection we are not so sensitive to single noise samples. In conventional OFDM synchronization no correlation is performed before differential detection The proposed technique will therefore give better performance, also when only used as a preamble in acquisition mode.

5 Frequency Estimator Performance

The frequency estimator structure is the same both in tracking and acquisition

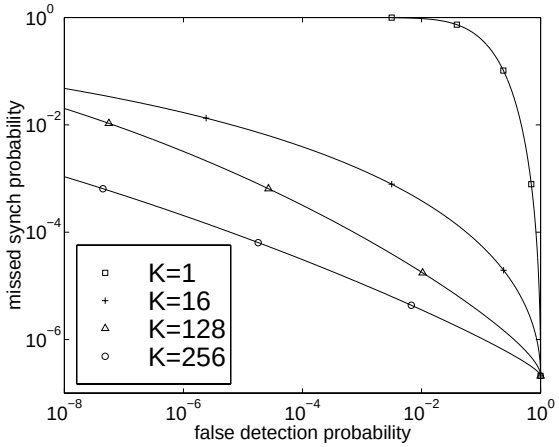

Figure 5: Receiver operation characteristics, probability of missing the synch signal vs. false detection probability. SNR=10 dB, ρ =0.1, $(K, L) = \{(1, 256), (16, 16), (128, 2), (256, 1)\}$ and observation interval 512 samples.

mode, but a two step procedure can be used in order to enhance the performance. In acquisition mode the sub-correlator length, *K*, and the number of delay elements, *P*, are chosen according to the maximum expected frequency offset. In tracking mode, initial compensation has already been made and therefore larger *K* and *P* can be used to decrease the variance of the frequency estimation error.

The phase increment method used gives a (nearly) unbiased frequency estimate of a complex exponential in white noise [8] and can cope with frequency offsets which are upper bounded by[1] $|\varepsilon_c| < M/(2PK)$. The normalized frequency estimate when the timing is known is given by[2]

$$\hat{\varepsilon}_c = \frac{M}{KP} \cdot \frac{\arg(\lambda[k, 0])}{2\pi}. \tag{8}$$

The frequency offset estimators in [1], [2], [4] and [5] can, from a performance point of view, be seen as a special case of the proposed estimator in acquisition mode. See Table 1 for a comparison of parameters, where *G* denotes the length of the cyclic prefix.

1 A typo in the original paper has been corrected, the factor 2 was missing.

2 A typo in the original paper has been corrected, 2π was missing.

Table 1: Comparison of parameters for some frequency offset estimators based on one symbol measurements.

reference	K	P	L
van de Beek [1]	1	M	G
Moose [2]	1	$M/2$	$M/2$
Nishinaga [5]	1	1	M -1
Fawer [4]	$M/2$	1	1
PN-based	$K < \frac{M}{2P\lvert\varepsilon_c\rvert}$	$P < \frac{M}{2K\lvert\varepsilon_c\rvert}$	$\frac{M}{K} - P$

The variance of the normalized frequency estimates can be calculated as [8]

$$var(\varepsilon_c) = \frac{M^2}{4\pi^2 K^3 LP} \left(\frac{1}{L} \cdot \frac{(1-\rho)\sigma_s^2 + \sigma_n^2}{\rho\sigma_s^2 \mathrm{sinc}^2(\varepsilon_c K/M)} + \frac{1}{2KP} \cdot \left(\frac{(1-\rho)\sigma_s^2 + \sigma_n^2}{\rho\sigma_s^2 \mathrm{sinc}^2(\varepsilon_c K/M)} \right)^2 \right). \quad (9)$$

Strictly speaking this is only valid for zero frequency offset, but simulations indicate [8] that it is a good approximation for other frequency offsets as well. An interesting property to note is that the same minimum variance can be achieved for different sub-correlator lengths as long as the number of delay elements are chosen properly.

Figure 6 shows theoretical and simulated values of the variance as a function of the frequency offset when the estimator is designed for a worst case normalized offset of 1, 4 and 10 respectively. Due to the flat region we can conclude that the performance is determined by the frequency offset for which the estimator is designed and that a worst case design has to be made. For large frequency offsets a two mode approach can be used to decrease the variance. First the parameters are adjusted for the worst case scenario and after a first coarse estimate has been made and corrected longer sub-correlators and larger delay can be used. As a rule of thumb the normalized frequency error must not exceed a few percent, see e.g. [2][9]. In order to fulfil this in tracking mode we need to extend the observation interval or increase the code power ratio compared to the parameters used in the figure. A possible solution could be to have correlation length 1 and an observation interval of 3 symbols. Then the frequency error variance becomes $1.3 \cdot 10^{-4}$ for small offsets in tracking mode.

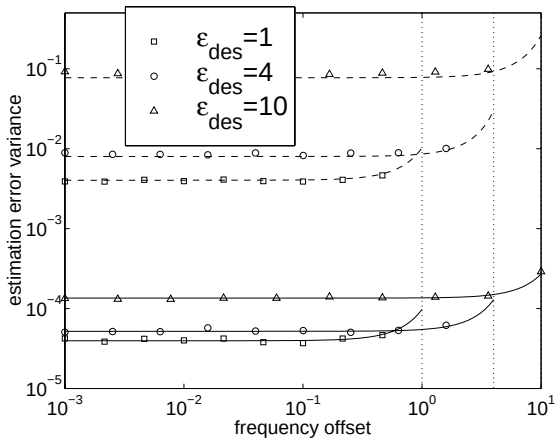

Figure 6: Variance of the normalized frequency estimation error as a function of the offset for SNR=10 dB and code power ratio ρ =0.1 (dashed) and ρ =1 (solid). The sub-correlators are designed for frequency offsets 1, 4 and 10. Marker values are obtained by simulations.

6 Conclusions

We have presented and analyzed a new approach for synchronization in OFDM systems. The technique is very flexible and has a large frequency offset range. Offsets up to half the OFDM bandwidth can be detected in one step, instead of the common offset range equal to one or half the sub-channel spacing. The synchronization signal is achieved by correlating received samples before they are differentially detected. This makes the synchronizer less sensitive to single noise values and results in better detection properties compared to conventional synchronizers. The amount of synchronization signal is controlled by a single parameter and can be tailored to a specific application. Therefore the presented synchronizer scheme is robust and works for low signal to noise ratios. The complexity of the synchronizer is low and it is therefore suitable for implementation.

Appendix A. Distribution of the Timing Metric

In this section we analyze the distribution of the timing metric when the timing is not correct. We analyze the case with one or two differential signals, L, or when the number of differential signals is large. The normalized complex outputs of the sub-correlators have zero mean and variance $\sigma^2_{\eta_a} = ((1-\rho)\sigma_s^2 + \sigma_n^2)/(KL\rho\sigma_s^2)$ if an

ideal code with perfect autocorrelation properties is used, and variance $\sigma^2_{\eta_a} = (\sigma^2_s + \sigma^2_n)/(KL\rho\sigma^2_s)$ for a random code.

For L=1 we can rewrite the timing metric as

$$|\lambda[k, a]| = \left|\eta_a[k - lK, a]\right|\left|\eta_a[k - (l + P)K, a]\right|, \tag{10}$$

where the two terms are Rayleigh distributed. By the transformation theorem we can derive the joint PDF for $|\lambda[k, a]|$ and $\left|\eta_a[k - lK, a]\right|$ and then integrate over $\left|\eta_a[k - lK, a]\right|$ to get the probability density function (PDF) for the timing metric. The PDF for incorrect timing becomes

$$f_{|\lambda[k,\,a]|}(d) = \left(\frac{2}{\sigma^2_{\eta_a}}\right)^2 d k_0\left(2\frac{d}{\sigma^2_{\eta_a}}\right), \tag{11}$$

where $k_0(x)$ denotes the zeroth-order modified Bessel function. The CDF of the timing metric is consequently given by

$$P\{|\lambda[k, a]| \le \alpha\} = F_D(\alpha) = 1 - 2\frac{\alpha}{\sigma^2_{\eta_a}} k_1\left(2\frac{\alpha}{\sigma^2_{\eta_a}}\right), \tag{12}$$

where k_1 is the first order modified Bessel function.

For L=2 we derive bounds for the timing metric. If L is even the metric can be rewritten as

$$\begin{aligned}\lambda[k, a] &= \sum_{l=0}^{L-1} \eta_a[k - lK, a]\eta_a^*[k - (l + P)K, a] \\ &= \sum_{k=0}^{L/2-1} d_{2k+1}c_{2k+1} + b_{2k+1}e_{2k+1}.\end{aligned} \tag{13}$$

All the terms (b, c, d and e) in (13) are independent zero mean Gaussian variables, d and b are real valued and have variance $\sigma^2_b = \sigma^2_d = E[bb] = \sigma^2_{\eta_a}/2$, whereas c and e are complex valued with variance of the both complex parts $\sigma^2_{Re\{c\}} = \sigma^2_{Im\{c\}} = \sigma^2_{\eta_a}$. Defining new variables as $g_k = d_{2k+1}Re\{c_{2k+1}\}$ and $h_k = d_{2k+1}$, the joint probability function for g and h becomes

$$f_{G,H}(g, h) = \frac{1}{2\pi\sigma_{Re\{c\}}\sigma_d}\exp\left(-\frac{1}{\sigma^2_{Re\{c\}}}\left(\frac{g}{h}\right)^2 - \frac{1}{\sigma^2_d}h^2\right)\frac{1}{h}. \tag{14}$$

Integrating [10, integral 3.478.4] over *h* gives the PDF for the products, i.e.

$$f_G(g) = \int_{-\infty}^{\infty} f_{G,H}(g,h)dh = \frac{1}{\pi\sigma_{Re\{c\}}\sigma_d} k_0\left(\frac{|g|}{\sigma_{Re\{c\}}\sigma_d}\right) \tag{15}$$

$$= \frac{\sqrt{2}}{\pi\sigma_{\eta_a}^2} k_0\left(\frac{\sqrt{2}|g|}{\sigma_{\eta_a}^2}\right).$$

Naturally this PDF also applies for $b_{2k+1}e_{2k+1}$ and for the corresponding imaginary parts. The characteristic function of the product, *g*, can be calculated as [10, integral 6.671.14]

$$\phi_g(\omega) = \int_{-\infty}^{\infty} f_G(g)e^{i\omega g}dg = \frac{\sqrt{2}}{\pi\sigma_{\eta_a}^2}\int_{-\infty}^{\infty} k_0\left(\frac{\sqrt{2}|g|}{\sigma_{\eta_a}^2}\right)e^{i\omega g}dg \tag{16}$$

$$= \frac{\sqrt{2}}{\sigma_{\eta_a}^2}\frac{1}{\sqrt{(\sqrt{2}/\sigma_{\eta_a}^2)^2 + \omega^2}}.$$

The PDF of the sum of products, i.e. the real part of λ, is given by the inverse transform of the product between the *L* characteristic functions,

$$f_{Re\{\lambda\}}(x) = \frac{1}{2\pi}\left(\frac{\sqrt{2}}{\sigma_{\eta_a}^2}\right)^L \int_{-\infty}^{\infty}\left(\left(\frac{\sqrt{2}}{\sigma_{\eta_a}^2}\right)^2 + \omega^2\right)^{-L/2} e^{-i\omega x}d\omega. \tag{17}$$

For *L*=2, we get [10, integral 3.389.5] a Laplacian random variable with PDF

$$f_{Re\{\lambda\}}(x) = \frac{1}{\sqrt{2}\sigma_{\eta_a}^2}\exp\left(-\frac{\sqrt{2}|x|}{\sigma_{\eta_a}^2}\right). \tag{18}$$

$|Re\{\lambda[k,a]\}|$ and $|Re\{\lambda[k,a]\}|$ are therefore independent exponential variables and the sum of them is therefore an Erlang-2 distributed variable. This means that the CDF of $|\lambda[k,a]|$ is upper bounded by the CDF of an Erlang-2 variable X_λ,

$$F_{X_\lambda}(x) = 1 - \left(1 + \frac{\sqrt{2}}{\sigma_{\eta_a}^2}x\right)\exp\left(-\frac{\sqrt{2}}{\sigma_{\eta_a}^2}x\right). \tag{19}$$

The CDF of $|\lambda[k,a]|$ is lower bounded by the CDF of the maximum, Y_λ, of two independent exponential variables,

$$F_{Y_\lambda}(y) = \left(1 - \exp\left(-\frac{\sqrt{2}}{\sigma^2_{\eta_a}} y\right)\right)^2 . \quad (20)$$

Finally, if L is large $Re\{\lambda\}$ and $Im\{\lambda\}$ can be approximated as independent Gaussian variables due to the central limit theorem. In this case $Re\{\lambda\}$ and $Im\{\lambda\}$ have zero mean and variance

$$\sigma^2_{Re\{\lambda\}} = \left(L(\sigma^2_{\eta_a})^2\right)/2 . \quad (21)$$

The normalized conventional timing metric is therefore Rayleigh distributed with CDF

$$P\{|\lambda[k, a]| \leq \alpha\} = F_{|\lambda|}(\alpha) \approx 1 - \exp\left(-\frac{\alpha^2}{2\sigma^2_{Re\{\lambda\}}}\right). \quad (22)$$

References

[1] J.J. van de Beek, M. Sandell and P. O. Börjesson, ”ML estimation of time and frequency offset in OFDM systems”, IEEE Trans. on Signal Processing, vol. 45, no. 7, July 1997, pp. 1800-1805.

[2] P. Moose, ”A technique for orthogonal frequency division multiplexing frequency offset correction”, IEEE Trans. on Communications, vol.42, no. 10, Oct. 1994, pp. 2908-14.

[3] T.M. Schmidl and D. C. Cox, ”Robust frequency and timing synchronization for OFDM”, IEEE Trans. on Communications, vol. 45, no. 12, Dec. 1997, pp. 1616-1621.

[4] U. Fawer, ”A coherent spread-spectrum diversity receiver with AFC for multipath fading channels”, IEEE Trans. on Communications, vol. 42, no. 2/3/4, Feb.-April 1994, pp. 1300-1311.

[5] N. Nishinaga, M. Nakagami and Y. Iwadare, ”A simple synchronization acquisition method for DS/SS system under carrier frequency offset”, IEICE Trans. on Fundamentals, vol. E80-A, no. 11, Nov. 1997, pp. 2162-2171.

[6] T.P. Holden and K. Feher, ”A spread spectrum based system technique for synchronization of digital mobile communication systems”, IEEE Trans. on Broadcasting, vol. 36, no. 3, Sept. 1990, pp. 185-194.

[7] A. Steingass, A.J. van Wijngaarden and W.G. Teich, ”Frame synchronization using superimposed sequences”, in Proc. IEEE International Symposium on

Information Theory, Ulm, Germany, June 29-July 4, 1997, p. 489.

[8] H. Meyr and M. Moeneclaey, Digital Communication Receivers. New York, U.S.A.: John Wiley & Sons, 1998.

[9] T. Pollet, M. Van Bladel and M. Monelcaey, "BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise", IEEE Trans on Communications, vol. 43, no. 2/3/4, Feb.-April 1995, pp. 191-193.

[10] I.S. Gradshteyn And I.M. Ryzhik, Table Of Integrals, Series, And Products. Alan Jeffrey, Editor, London, UK: Academic Press, 1994.

Paper F

This part has been submitted as:

Fredrik Tufvesson and Peter Hoeher, "Channel Estimation using Superimposed Pilot Sequences", submitted to IEEE Transactions on Communications, March 2000.

Channel Estimation using Superimposed Pilot Sequences*

Fredrik Tufvesson[‡] and Peter Hoeher[+]

[‡]Department of Applied Electronics, Lund University
Box 118, SE-221 00 Lund, Sweden
Fax: +46-46 12 99 48, E-mail: Fredrik.Tufvesson@tde.lth.se

[+]Information and Coding Theory Lab, University of Kiel
Kaiserstr. 2, D-24143 Kiel, Germany,
Fax: +49-431-77572-403, E-mail: ph@techfak.uni-kiel.de.

Index terms: Channel estimation, phase synchronization, equalizers,
Viterbi detection, sequences, Rayleigh channels,
Phase Shift Keying, Quadrature Amplitude Modulation

Abstract

For the purpose of various synchronization tasks (including carrier phase, time, frequency, and frame synchronization), one may add a known pilot sequence, typically a pseudo-noise sequence, to the unknown data sequence. This approach is known as a spread-spectrum pilot technique or as a superimposed pilot sequence technique. In this paper, we apply the superimposed pilot sequence technique for the purpose of channel estimation (CE). We propose a truly coherent receiver based on the Viterbi algorithm, and derive its pairwise error probability. This analytical result gives further insight and leads to an approximation of the bit error rate, which is confirmed by Monte Carlo simulations. Furthermore, we present a generic low-cost receiver structure based on reduced-state sequence estimation. Among the distinct advantages compared to conventional pilot-symbol-assisted CE are (i) no bandwidth expansion and (ii) a significantly improved performance in fast fading environments. The proposed Viterbi receiver may also be used as an alternative receiver for pilot-symbol-assisted CE, making this scheme more robust in fast-fading channels.

*Parts of this work were presented at the Int. Workshop on Multi-Carrier Spread-Spectrum, Oberpfaffenhofen, Germany, Sept. 1999 and IEEE GLOBECOM '99, Rio de Janeiro, Brazil, Dec. 1999.

1 Introduction

Consider digital data transmission over time-selective flat-fading channels. Assuming coherent demodulation, one of the main problems is carrier phase synchronization, both in terms of acquisition and tracking, particularly when the channel is fast and when a line-of-sight component is absent [1][2].

A popular technique to maintain coherent demodulation for a wide class of digital modulation schemes has been proposed by Moher and Lodge [3] [4], and is known as *pilot-symbol-assisted CE*[1]. The main idea of pilot-symbol-assisted CE is to multiplex known pilot symbols (also called training symbols) into an unknown data stream. Conventionally, the receiver firstly obtains tentative channel estimates at the positions of the pilot symbols by means of re-modulation, and then computes final estimates by means of interpolation. Aghamohammadi *et al.* and Cavers were among the first analyzing and optimizing pilot-symbol-assisted CE given different interpolation filters [5] [6]. Due to the pilot symbols the bandwidth slightly increases.

In this paper, we explore a related technique proposed a dozen years ago by Makrakis, Feher and Holden [7][8]. The technique was originally called spread-spectrum pilot technique and the idea is to linearly add a known pilot sequence to the unknown data sequence, see Section 2. Similar methods are applied in the field of automatic control, where intentional pertubation signals are used, e.g., for dual control, closed loop identification and extremum control of dynamical systems [9]. For communication systems Makrakis, Feher and Holden applied the technique to phase synchronization [7][8]. Later, the technique has been applied for the purposes of frame synchronization [10] and joint time and frequency synchronization of OFDM signals [11], respectively. A key feature of the scheme is that the pilot information and the data are transmitted in the same dimension on the same channel. Alternative techniques may be based on transmitting both sequences on separate channels. One can, e.g., use two orthogonal filters as suggested by Bejjani and Belfiore [12], or use an orthogonal signal dimension (e.g. the quadrature component), as done in the UMTS/W-CDMA uplink [13]. In both cases, however, the bandwidth efficiency would be reduced.

Within this paper, we assume synchronous symbol-by-symbol superposition of the pilot symbols, where the power of the pilot symbols is typically much less than the power of the data symbols. By construction, there is no increase in bandwidth when using a pseudo-noise pilot sequence. Therefore, we use the notion of a *superimposed pilot sequence technique* as suggested in [10], because "spread-spectrum" techniques may suggest a widening of the spectrum. As opposed to pilot-symbol-assisted CE, no interpolation is necessary. The superimposed pilot sequence technique is therefore more bandwidth *and* power

[1] Channel estimation may be seen as a generalization of carrier phase synchronization, since estimates of the quadrature components cover the carrier phase as well as reliability information, often called channel state information.

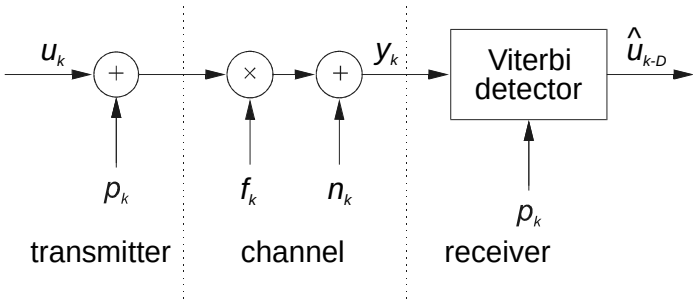

Figure 1: Block diagram of the transmission scheme.

efficient than pilot-symbol-assisted CE, particularly in fast fading conditions, as shown in this paper.

In this paper we focus on single-carrier systems and frequency-flat fading channels. Aspects of multi-carrier systems were presented in [14]. Our main contribution is the derivation and analysis of a receiver structure for the purpose of CE (phase synchronization). We derive the optimal metric and an additive metric which is suitable for recursive algorithms. We then analyze the performance both analytically and in terms of simulations. In Section 3 the receiver, and structured simplifications thereof, is presented. The proposed transmission scheme is then evaluated analytically in Section 4 and by simulations in Section 5. Finally, the conclusions are drawn in Section 6.

2 Transmission Scheme with Superimposed Pilot Sequence

A block diagram of the transmission scheme featuring the transmitter, the channel and the receiver is shown in Fig. 1. In the following, we use complex baseband notation and assume perfect time, frequency and frame synchronization. In the transmitter a known pilot sequence, typically a pseudo-noise (PN) sequence, is synchronously added to the unknown data sequence. Alternatively, a pilot sequence may be added to the phase of the data sequence, which may be the better choice for CPM. The chips of the known pilot sequence and the i.i.d. data symbols are denoted as $p_k \in \mathbb{C}$ and $u_k \in \mathbb{C}$, respectively, where k is the time index. We also assume that the chip duration is equal to the symbol duration. The receiver relies on a low correlation between the pilot sequence and the data sequence in order to resolve phase ambiguities. The data sequence is assumed to be M-ary and both sequences have zero mean: $E[u_k] = E[p_k] = 0$. Let the energy of the pilot symbols be $E[|p_k|^2] = \rho \cdot E_s$ and the energy of the data symbols be $E[|u_k|^2] = (1-\rho) \cdot E_s$, respectively.

Hence, the average energy of the composite symbols is E_s, if the sequences are mutually independent: $E[u_k \cdot p_k] = 0$. The average energy per information bit is $E_b = E_s / \log_2 M$.

The normalized power of the pilot symbols, ρ, corresponds to the amount of known information transmitted, and should be optimized. In pilot-symbol-assisted CE systems the related parameter is the spacing of the pilot symbols. In the remaining, ρ is assumed to be constant for all k, since our focus is on tracking. In order to improve the acquisition behavior, ρ may be larger for the first symbols within a data frame [11].

Note that in pilot-symbol-assisted CE systems the data and pilot symbols are orthogonal by means of time-division multiplexing, whereas for superimposed pilot sequences the cross-correlation between the data and pilot symbols typically is non-zero. However, the superimposed pilot sequence is known and therefore the disturbance can be made small. The loss due to the pilot sequence is over-compensated by the lack of interpolation, as we will show in Section 5. The transmission system may be interpreted as a two-user system, where one user transmits information unknown to the receiver, and the other user transmits known information. Correspondingly, the transmission technique is related to multi-user systems, and the synchronization/detection problem is related to multi-user detection.

Given our assumptions, the matched filter output samples, $y_k \in \mathbb{C}$, can be written as

$$y_k = (u_k + p_k) \cdot f_k + n_k, \tag{1}$$

where $f_k \in \mathbb{C}$ is a multiplicative (time-selective) fading process with $E[|f_k|^2] = 1$, and $n_k \in \mathbb{C}$ are zero-mean white Gaussian noise samples with one-sided power spectral density N_0. The channel under consideration has a Rayleigh fading amplitude and a "classical" Doppler spectrum with autocorrelation coefficients $r_l = \mathrm{J}_0(2\pi f_{D_{max}} T_s l)$, where $f_{D_{max}}$ denotes the maximum Doppler frequency, and T_s the symbol duration. Oversampling is not treated within this paper, but could be used to further improve the performance in fast fading.

In the transmitter, the major difference between the spread-spectrum pilot technique proposed by Feher *et al.* and the scheme described above is that in the original proposal the linear addition was done after D/A conversion, whereas we do digital baseband processing. Therefore, we can assure symbol-synchronous transmission, which is an important property for our novel receiver. As a side effect, the power density spectrum is unaffected by the superimposed pilot sequence when using a PN pilot sequence.

An important aspect is the design of the pilot sequence. If p_k is chosen to be constant for all k, the pilot sequence becomes a pilot tone. Given a PN data sequence, the performance would on the average be the same compared to a PN pilot sequence. However, simultaneous time and frequency synchronization would no longer be possible with a constant pilot sequence. Furthermore,

given a PN pilot sequence, we may extend our receiver to cope with frequency-selective fading as well. Finally, it is interesting to note that the concept of superimposed pilot sequences is applicable to most existing analog and digital transmission schemes as an "add-on" feature

3 Receiver Structure

The conventional receiver for the superimposed pilot sequence technique splits the received signal, and feeds it into a carrier recovery unit and a detection unit [7][8], respectively. The output signal of the carrier recovery unit is fed into the detection unit. In order to partly compensate the interference due to the pilot sequence, a simple cancellation scheme has been proposed.

The conventional receiver for the pilot-symbol-assisted CE technique operates similarly: the received signal is demultiplexed, and fed into a channel estimation unit (i.e., an interpolator) and a detection unit [3][4][5][6], respectively. Interference cancellation is not necessary by design.

As opposed to these receivers, we will now introduce a recursive receiver based on the Viterbi algorithm, which directly outputs estimates of the data symbols. This receiver is primarily designed for the superimposed pilot sequence technique, but is also applicable for the conventional pilot-symbol-assisted CE technique. Different channel estimates are computed for different hypothesis by means of the principles of per-survivor processing [15][16]. Interference cancellation is done inherently. The principle of the receiver is the following: The receiver starts with a hypothesis of the transmitted data sequence. Assuming the hypothesis is correct, the received symbols are then used together with the hypothesis to calculate channel estimates for each of the symbols in the block. Then the distance (metric) between the received sequence and the hypothesis affected by the corresponding channel estimates is calculated. The same procedure is then used for the next hypothesis. Finally, the receiver makes a decision in favor of the sequence with the lowest distance.

For the calculations it is convenient to use matrix notation. We denote the received sample vector by $\mathbf{y} = (y_1, y_2, ..., y_N)^T = \mathbf{Cf} + \mathbf{n}$, where $\mathbf{f}$ and $\mathbf{n}$ are vectors of the fading process and the noise, respectively, and $\mathbf{C}$ is a diagonal matrix with the transmitted symbols on its diagonal: $c_{k,k} = (u_k + p_k)$.

The channel estimates are calculated using the L latest received symbols. According to (1), a tentative channel estimate, $\hat{f}_{k-l}$, can be obtained by remodulation:

$$\hat{f}_k = \frac{y_k}{\tilde{u}_k + p_k}, \tag{2}$$

where l is an arbitrary integer number. Note that the first part of the denominator is based on the hypothesis, whereas the second part is known: the pilot symbols helps to resolve the phase ambiguity. Given hypotheses $\tilde{u}_k, \ldots, \tilde{u}_{k-L+1}$,

the channel estimate $\tilde{f}_k$ can be computed by linear filtering [15]:

$$\tilde{f}_k = \sum_{l=0}^{L-1} a_{k,l} \cdot \hat{f}_{k-l} = \mathbf{a}_k \tilde{\mathbf{C}}^{-1} \mathbf{y}, \tag{3}$$

where L is the filter order and $\mathbf{a}_k$ is a vector of the channel estimation coefficients for symbol number k.

If the channel is wide-sense stationary, the optimal filter coefficients, a_l, $0 \leq l \leq L-1$, are the solutions of the Wiener-Hopf equations:

$$(a_0, \ldots, a_{L-1})^T = \mathbf{a}_{opt}^T = \Phi^{-1} \boldsymbol{\theta}, \tag{4}$$

where $\theta_l = E[f_k \cdot \hat{f}_{k-l}^*] = r_l$ are the elements of the cross-correlation vector $\boldsymbol{\theta}$ assuming a correct hypothesis. Further, $\Phi_{l,j} = E[\hat{f}_{k-l} \cdot \hat{f}_{k-j}^*] = \sigma_n^2 / |\tilde{u}_{k-l} + p_{k-l}|^2 \cdot \delta_{l,j} + r_{j-l}$ are the elements of the autocorrelation matrix Φ, where $l, j = 1, \ldots, L$, $r_l = r_{-l}$ is the lth autocorrelation coefficient of the channel, and $\sigma_n^2 = N_0/E_s$ is the variance of the noise. The corresponding vector of channel estimates for a certain hypothesis can consequently be written as

$$\tilde{\mathbf{f}} = \mathbf{A} \tilde{\mathbf{C}}^{-1} \mathbf{y}. \tag{5}$$

For example, for a block length of 5 received symbols and $L = 3$, the channel estimation matrix has the form

$$\mathbf{A} = \begin{pmatrix} a_{1,0} & 0 & 0 & 0 & 0 \\ a_{2,1} & a_{2,0} & 0 & 0 & 0 \\ a_{3,2} & a_{3,1} & a_{3,0} & 0 & 0 \\ 0 & a_{4,2} & a_{4,1} & a_{4,0} & 0 \\ 0 & 0 & a_{5,2} & a_{5,1} & a_{5,0} \end{pmatrix}, \tag{6}$$

where $a_{i,j}$ are the filter coefficients. Note that the optimal filter coefficients are dependent on the hypothesis. For low ρ values the difference becomes small, however, and a static filter gives a good approximation.

The goal is to recursively compute the maximum-likelihood estimate of the transmitted sequence. The receiver then tests all possible hypotheses and decides for the sequence resulting in the lowest metric, i.e. the highest probability. It is, however, hard to compute the true maximum-likelihood estimate recursively, and therefore we have to make approximations to achieve an additive metric suitable for recursive receivers. We derive the metric in Appendix A. The optimal metric is given by

$$\tilde{\Lambda}_{opt} = \ln(\det(\tilde{\boldsymbol{\phi}})) + \mathbf{y}^H (\tilde{\mathbf{C}} \mathbf{R}_f^H \tilde{\mathbf{C}}^H + \sigma_n^2 \cdot \mathbf{I})^{-1} \mathbf{y}, \tag{7}$$

where $\boldsymbol{\phi}$ denotes the covariance matrix of the received samples, $\tilde{\boldsymbol{\phi}} = E[(\mathbf{y} - \bar{\mathbf{y}}) (\mathbf{y} - \bar{\mathbf{y}})^H]$. Assuming the hypothesis is correct this covariance matrix can be

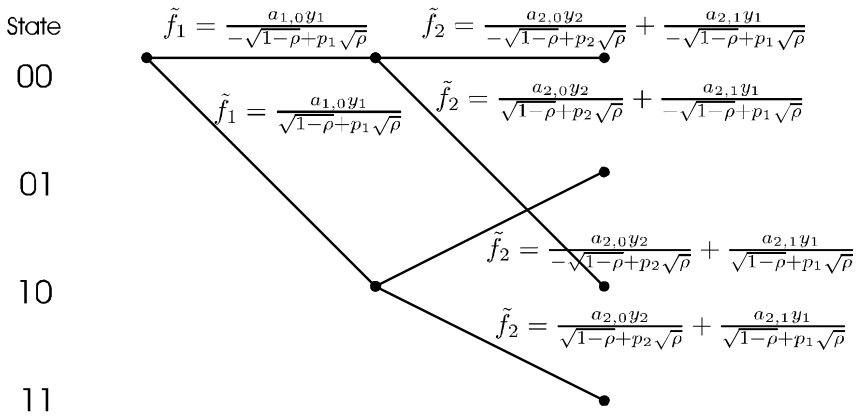

Figure 2: The channel estimates for the different hypotheses of the two first symbols.

calculated as

$$\begin{aligned}\boldsymbol{\phi} &= E[(\mathbf{I} - \tilde{\mathbf{C}}\mathbf{A}\tilde{\mathbf{C}}^{-1})\mathbf{y}\mathbf{y}^H(\mathbf{I} - \tilde{\mathbf{C}}\mathbf{A}\tilde{\mathbf{C}}^{-1})^H] \\ &= (\mathbf{I} - \tilde{\mathbf{C}}\mathbf{A}\tilde{\mathbf{C}}^{-1})(\tilde{\mathbf{C}}\mathbf{R}_f^H\tilde{\mathbf{C}}^H + \sigma_n^2 \cdot \mathbf{I})(\mathbf{I} - \tilde{\mathbf{C}}\mathbf{A}\tilde{\mathbf{C}}^{-1})^H \end{aligned} \quad (8)$$

where $\mathbf{R}_f$ is the covariance matrix of the channel containing the autocorrelation coefficients.

With the additive metric increments it is possible to use the Viterbi algorithm for detection. In Appendix A it is shown that a suitable approximation that yields additive metric is given by

$$\tilde{\Lambda} = (\mathbf{y} - \tilde{\mathbf{C}}\mathbf{A}\tilde{\mathbf{C}}^{-1}\mathbf{y})^H(\mathbf{y} - \tilde{\mathbf{C}}\mathbf{A}\tilde{\mathbf{C}}^{-1}\mathbf{y}). \quad (9)$$

The corresponding additive metric increment used in the Viterbi receiver is then given by

$$\tilde{\Lambda}_k(\tilde{c}_k, \tilde{c}_{k-1}, \ldots, \tilde{c}_{k-L}) = |y_k - \tilde{c}_k \cdot \tilde{f}_k|^2, \quad (10)$$

where $\tilde{f}_k$ is the channel estimate based on the current hypothesis. In Fig. 2 the channel estimates for the first two symbols in a four-state trellis is presented.

The channel estimates are based on the hypotheses and on the correlation between a specific channel estimate and the actual channel value is high only when the hypothesis is correct. The correlation can be calculated as

$$E\left[(\sum_{l=0}^{L-1} a_l \cdot \hat{f}_{k-l})f_k^*\right] = \sum_{l=0}^{L-1} a_l \frac{u_{k-l} + p_{k-l}}{\tilde{u}_{k-l} + p_{k-l}} E[f_{k-l}f_k^*]. \quad (11)$$

The fraction on the right hand side equals one only for the correct hypothesis thereby producing a correct channel estimate. The receiver can be interpreted as an estimator-correlating detector, where parts of the received signal are known and parts of the signal are unknown. In a first step, the unknown channel response (and thereby the received signal) is estimated by the received symbols and the hypothesis. In the next step this estimate is correlated with the received symbols. If the correlation between the channel estimate and the actual channel value is high, then the metric will be low and the hypothesis that corresponds to the transmitted data will be chosen.

The hypotheses $\tilde{u}_{k-1}, \ldots, \tilde{u}_{k-L+1}$ constitute the M^{L-1} states of a trellis, and $\tilde{u}_k$ belongs to the actual branch. The trellis has M branches/state. For the additive metric the familiar add-compare-select operation can be applied and the most probable sequence can be found by back-tracing or related operations. A rule of thumb for the decision delay is $D \approx L \ldots 4L$. Due to the *sequence* estimation, the receiver can cope with a low-power pilot sequence.

For large M and L, the complexity can be significantly reduced by applying the principles of reduced-state sequence estimation [17] [18]:

$$\tilde{f}_k = \sum_{l=0}^{K-1} a_l \cdot y_{k-l}/(\tilde{u}_{k-l} + p_{k-l}) + \sum_{l=K}^{L-1} a_l \cdot y_{k-l}/(\hat{u}_{k-l} + p_{k-l}), \qquad (12)$$

where the trade-off between complexity and performance can be adjusted by the parameter K, $0 \leq K \leq L$. $\tilde{u}_{k-1}, \ldots, \tilde{u}_{k-K+1}$ belong to a trellis with M^{K-1} states and M branches/state, whereas $\hat{u}_{k-K}, \ldots, \hat{u}_{k-L+1}$ belong to the path history (survivor sequence) of the corresponding state. Therefore, $K = L$ corresponds to maximum-likelihood sequence estimation and $K = 0$ is related to "parallel decision feedback estimation". For $M > 2$, further simplifications can be achieved by the principles of set-partitioning [17].

The novel receiver is suitable to process pilot-symbol-assisted CE too. In contrast to the conventional pilot-symbol-assisted channel estimator, all matched filter output samples, i.e. also the data symbols, are used for CE. Although not treated further, the proposed receiver may also be extended to accept a priori information and to deliver soft-outputs (which is necessary for iterative processing). Furthermore, the receiver may be generalized for frequency-selective channels.

4 Analytical Performance Evaluation

In this section we make an analytical analysis of the BER using superimposed pilots. First we derive the pairwise error probability that an erroneous hypothesis is chosen instead of the correct one. We then use this error probability to find the most probable error events. Using these we finally get an approximation of the overall BER by a Markov chain approach.

To make a decision the receiver calculates the cumulative metrics (9) for each of the hypotheses. The receiver makes a correct decision if the metric for the correct path is smaller than all of the other path metrics. Define the additive decision variable as $\hat{\Lambda} = \tilde{\Lambda} - \Lambda$. If $\hat{\Lambda} > 0$ for all hypotheses then the correct path will be chosen. The decision variable can be expressed as

$$\hat{\Lambda} = \mathbf{y}^H[(\mathbf{I} - \tilde{\mathbf{C}}\mathbf{A}\tilde{\mathbf{C}}^{-1})^H(\mathbf{I} - \tilde{\mathbf{C}}\mathbf{A}\tilde{\mathbf{C}}^{-1}) - (\mathbf{I} - \mathbf{C}\mathbf{A}\mathbf{C}^{-1})^H(\mathbf{I} - \mathbf{C}\mathbf{A}\mathbf{C}^{-1})]\mathbf{y}. \tag{13}$$

Rewrite the received samples as $\mathbf{y} = \mathbf{D}\mathbf{g}$ with

$$\mathbf{D} = \begin{bmatrix} u_1 + p_1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & u_2 + p_2 & & 0 & 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots & \vdots & & \ddots & \\ 0 & 0 & & u_N + p_N & 0 & 0 & & 1 \end{bmatrix} \tag{14}$$

and

$$\mathbf{g} = \begin{bmatrix} \mathbf{f} \\ \mathbf{n} \end{bmatrix}. \tag{15}$$

The decision variable can then be rewritten as

$$\begin{aligned} \hat{\Lambda} &= \mathbf{g}^H\mathbf{D}^H[(\mathbf{I} - \tilde{\mathbf{C}}\mathbf{A}\tilde{\mathbf{C}}^{-1})^H(\mathbf{I} - \tilde{\mathbf{C}}\mathbf{A}\tilde{\mathbf{C}}^{-1}) - (\mathbf{I} - \mathbf{C}\mathbf{A}\mathbf{C}^{-1})^H(\mathbf{I} - \mathbf{C}\mathbf{A}\mathbf{C}^{-1})]\mathbf{D}\mathbf{g} \\ &= \mathbf{g}^H\tilde{\mathbf{K}}\mathbf{g}. \end{aligned} \tag{16}$$

Note that the matrix $\tilde{\mathbf{K}}$ is symmetric and that $\mathbf{g}$ is a vector of correlated Gaussian samples. The covariance matrix, $\mathbf{R}_g$, of the zero-mean Gaussian variables $\mathbf{g}$ can be calculated as

$$\mathbf{R}_g = \begin{bmatrix} \mathbf{R}_f & \mathbf{0} \\ \mathbf{0} & \sigma_n^2 \cdot \mathbf{I} \end{bmatrix}, \tag{17}$$

where $\mathbf{R}_f$ is the covariance matrix of the fading process.

The expressions for the optimal decision variable are similar to the additive ones and the same calculations apply. The optimal decision variable is

$$\begin{aligned} \hat{\Lambda}_{opt} &= \mathbf{g}^H\mathbf{D}^H[(\tilde{\mathbf{C}}\mathbf{R}_f^H\tilde{\mathbf{C}}^H + \sigma_n^2 \cdot \mathbf{I})^{-1} - (\mathbf{C}\mathbf{R}_f^H\mathbf{C}^H + \sigma_n^2 \cdot \mathbf{I})^{-1}]\mathbf{D}\mathbf{g} \\ &= \mathbf{g}^H\tilde{\mathbf{K}}_{opt}\mathbf{g}, \end{aligned} \tag{18}$$

but to make correct decisions this has to exceed $\ln(\det(\boldsymbol{\phi})) - \ln(\det(\tilde{\boldsymbol{\phi}}))$ instead of zero.

4.1 Pairwise Error Probability

Many authors have studied the pairwise error probability resulting from quadratic forms like (16), e.g. in [19][20][21]. In the sequel we follow [21] to derive the

distribution of the decision variable and calculate the pairwise error probability. Define new Gaussian variables as $\mathbf{g}' = \mathbf{B}\mathbf{g}$, where $\mathbf{B}$ is a rank r matrix of size $2N \times r$ fulfilling $\mathbf{R}_g = \mathbf{B}\mathbf{B}^H$. In [21] it is shown that the decision variable can be regarded as a sum of r independent weighted χ^2-variables, where the weights are the eigenvalues of $\mathbf{B}^H\ \tilde{\mathbf{K}}\mathbf{B}$,

$$\hat{\Lambda} = \mathbf{g}^H \tilde{\mathbf{K}} \mathbf{g} = \sum_{n=1}^{N_s} \lambda_n \chi^2_{2,n}. \tag{19}$$

The vector of Gaussian samples, $\mathbf{g}$, is complex and the χ^2-variables thereby have two degrees of freedom. The mean of the decision variable gives a good indication of the error probability. It can be calculated as the trace [21]

$$E[\hat{\Lambda}] = tr(\tilde{\mathbf{K}}\mathbf{R}_g). \tag{20}$$

To calculate the probability density function of $\hat{\Lambda}$ we sort the weights so that the first m eigenvalues are positive and the last $N_s - m$ eigenvalues are negative. The decision variable can then be expressed as

$$\hat{\Lambda} = \sum_{n=1}^{m} \lambda_n \chi^2_{2,n} - \sum_{n=m+1}^{N_s} -\lambda_n \chi^2_{2,n.} \tag{21}$$

Assuming that all the eigenvalues are distinct, which they for all practical cases are, the probability density function of $\hat{\Lambda}$ can be calculated as [21]

$$f_{\hat{\Lambda}}(y) = \begin{cases} \sum_{n=1}^{m} \frac{(\lambda_n)^{N_s-2} e^{-y/(2\lambda_n)}}{2\left[\prod_{k=1,k\neq n}^{N_s} (\lambda_n - \lambda_k)\right]} & y > 0 \\ \sum_{n=m+1}^{N_s} \frac{(-\lambda_n)^{N_s-2} e^{-y/(2\lambda_n)}}{2\left[\prod_{k=1,k\neq n}^{N_s} (\lambda_k - \lambda_n)\right]} & y \leq 0. \end{cases} \tag{22}$$

The pairwise error probability for the additive decision variable, i.e. the probability that the decision variable (16) becomes negative so that we will make a decision in favor of the erroneous hypothesis instead of the correct one, is then

$$P_{add} = \int_{-\infty}^{0} f_{\hat{\Lambda}}(y)dy = \sum_{n=m+1}^{N_s} \frac{(-\lambda_n)^{N_s-1}}{\prod_{k=1,k\neq n}^{N_s} (\lambda_k - \lambda_n)}. \tag{23}$$

The corresponding error probability for the optimal decision variable depends

on the sign of the detection limit $x = \ln(\det(\boldsymbol{\phi})) - \ln(\det(\tilde{\boldsymbol{\phi}}))$,

$$P_{opt}(x) = \begin{cases} \sum_{n=m+1}^{N_s} \frac{(-\lambda_n)^{N_s-1}}{\prod_{k=1,k\neq n}^{N_s}(\lambda_k-\lambda_n)} + \sum_{n=1}^{m} \frac{(\lambda_n)^{N_s-1}}{\prod_{k=1,k\neq n}^{N_s}(\lambda_n-\lambda_k)}(1-e^{-\frac{x}{2\lambda_n}}) & x > 0 \\ \sum_{n=m+1}^{N_s} \frac{(-\lambda_n)^{N_s-1}}{\prod_{k=1,k\neq n}^{N_s}(\lambda_k-\lambda_n)}(1-e^{-\frac{x}{2\lambda_n}}) & x \leq 0. \end{cases} \tag{24}$$

Note that the eigenvalues λ_n are dependent on the transmitted data which means that the error probability varies depending on the transmitted data sequence.

4.2 Most Probable Errors

The pairwise error probabilities are, of course, dependent on the specific parameters used, e.g. the power of the pilot sequence, the signal to noise ratio (SNR) and the block length. As an example we will now study the theoretical pairwise error probabilities for a small block. The block is too small for a practical application, but it is instructive to analyze the performance when the total number of hypotheses, i.e. the number of different paths in the trellis, is small enough to make the example manageable. Assume that 8 BPSK symbols are sent ($M = 2$). The data and the pilot sequences are randomly chosen as $\mathbf{u} = \sqrt{1-\rho}(-1,-1,-1,-1,1,1,1)^T$ and $\mathbf{p} = \sqrt{\rho}(-1,-1,1,1,1,-1,1,1)^T$, respectively. For the example we have 5% pilot power, filter length 3 and a Doppler rate of 1% ($\rho = 0.05$, $L = 3$, $f_{D_{max}}T_s = 0.01$). In Fig. 3 we plot the pairwise error probabilities for all the 255 erroneous hypotheses when using the optimal decision variable and the optimal filter coefficients. The hypotheses are ordered by their theoretical path probability at $E_b/N_0 = 15$ dB. In Fig. 4 we plot the same probabilities for the additive metric when using the fixed filter coefficients, i.e. filter coefficients where the varying noise variance due to the different hypotheses is not taken into account. Although not shown here, the presented theoretical values are verified by simulations. As seen in the figures the difference in error probability is small between the optimal metric using the optimal filter coefficients and the additive metric using the fixed filter coefficients. The slope of the curves is somewhat lower for the additive metric, but for low and medium values of the SNR the penalty is small. At high SNR values however, the penalty for the additive metric becomes more evident for some specific hypotheses.

There are in principle three kinds of common errors for small blocks: Errors where the channel estimator makes a constant phase error of 180 degrees, errors where there is a 180 degree phase shift in the demodulated data sequence, and errors with a single phase shifted symbol somewhere in the middle of the sequence. The first group is avoided by increasing the block length or the pilot

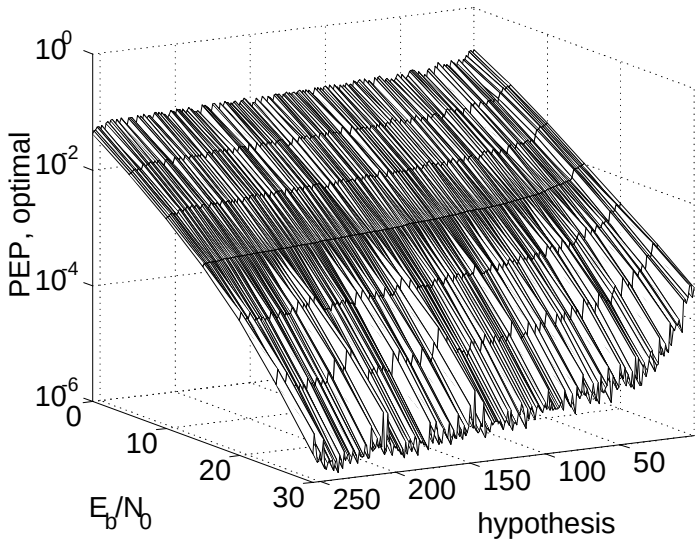

Figure 3: Pairwise error probabilities (PEP) for all 255 erroneous hypotheses in an 8-symbol block using the optimal metric and the optimal filter coefficients.

power. The errors with a 180 degree phase shift in the middle of the sequence are also avoided by increasing the block length. The metric for the erroneous paths in the trellis becomes high after a few symbols so for long blocks these hypotheses will not be chosen. However, also for longer blocks there are still common errors with a 180 degree phase shift in the beginning or near the end of the block. This is because there is no initial channel estimate for the first symbols and no known end-state for the last ones. If the first or last symbols are heavily disturbed by noise the channel estimates can be adjusted to these erroneous symbols and it takes some time before the channel estimate becomes correct again. Finally, the most common group of errors is where we have a single phase shifted symbol in the middle of the sequence. These errors can of course be corrected by using any kind of channel coding after the detector or by using joint detection and decoding. The errors with a single phase shifted symbol means that the channel estimate in the neighborhood of the error will be based on an incorrect hypothesis. But since only one of the L symbols used for the estimate is wrong, the channel estimation error anyhow becomes quite small.

For large or infinite block lengths the errors appear as single symbol errors or as a group of erroneous symbols. The latter is especially true for low Doppler frequencies where the fading dips in Rayleigh fading channels have duration

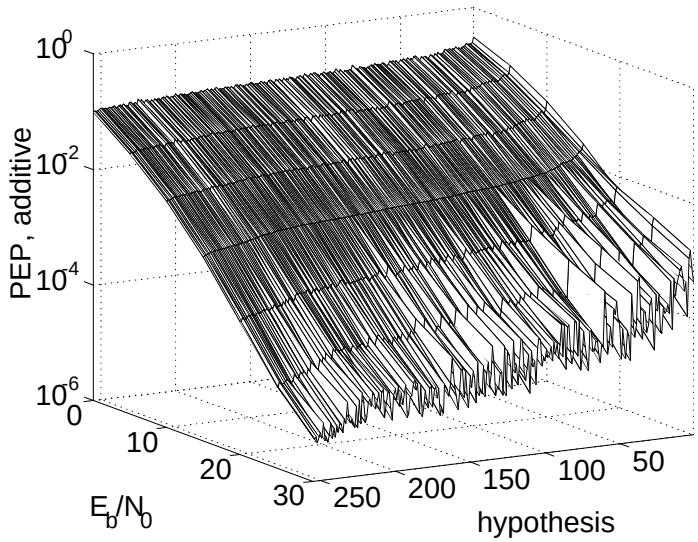

Figure 4: Pairwise error probabilities (PEP) for all 255 erroneous hypotheses in an 8-symbol block using the additive metric and fixed filter coefficients.

over several symbols. In a fading dip the probability is high that the wrong paths in the trellis will be chosen, which means that the channel estimates are less reliable there. However, if an erroneous path is chosen the metric becomes high after a few symbols and therefore the correct path will be reached again after a certain number of symbols. The receiver can be seen as a Markov chain where the transition probability and the probability of making a correct decision are dependent on the state, i.e. the particular symbols (that might be erroneous) used to form the channel estimate.

There is a high correlation between the mean of the decision variable and the corresponding pairwise error probability, and therefore studying the mean is an easy way to find the most probable error paths. The most probable erroneous hypotheses also have the lowest mean of the decision variable. There is also a strong dependence between the rank of the matrix $\mathbf{R}_g^H \tilde{\mathbf{K}}$ and the pairwise error probability. If the rank is low, the mean of the decision variable is low as well and there is a high probability for that specific error event.

4.3 Approximative BER Calculation

To calculate an approximative BER we use a Markov chain approach. For other detection schemes it is often possible to achieve a tight union bound using only

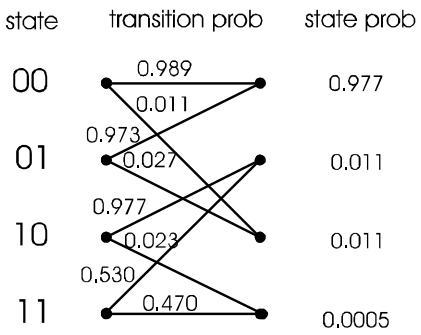

Figure 5: Transition probabilities in the trellis and the corresponding state probabilities.

the single error events, see [22] for BER calculations of trellis coded modulation. For the superimposed pilot technique, however, there is a dependence between the errors and a union bound does not converge. Instead we use the pairwise error probabilities to calculate the transition probabilities in the trellis, i.e. in the Markov chain. Assume that the error event has a limited duration. For a given transmitted sequence and a certain state in the trellis we first identify the two most probable hypotheses with and without an error at position k. We then use only these two sequences to calculate the probability that the decision is made in favor of the correct symbol and thereby the transition probabilities to the next states in the trellis. Assuming that the data sequence is $-1, \ldots, -1$, we can use these transition probabilities to calculate the probability of being in a certain state in steady state, i.e. in the middle of a long block. These state probabilities are then multiplied by the probability of making an erroneous decision for each of the states to get the approximative BER. For example, assume that $L = 3$, $\rho = 0.05$, $f_{D_{\max}} T_s = 0..01$ and $E_b/N_0 = 15$ dB. Further, assume that the transmitted pilot sequence is a sequence of ones mapped to $\mathbf{p} = \sqrt{\rho}(1, 1, 1, 1, 1, 1, 1)^T$ and the data sequence is the zero sequence mapped to $\mathbf{u} = \sqrt{1-\rho}(-1, -1, -1, -1, -1, -1, -1)^T$. For state 00 the most probable hypotheses with and without an error at position $k = 5$ is 0000100 and 0000000, respectively. Here we have two symbols to get into steady state, two symbols to reach the actual state, the erroneous or correct symbol and then two symbols to reach the correct state again. For state 01 the corresponding hypotheses are 0010100 and 0010000, respectively. The pairwise error probability between the sequences 0010100 and 0010000 when the data sequence is 0000000 is 0.027 for state 01 and therefore the probability of making an erroneous decision is 0.027 in this state, the transition probability to state 10 is 0.027 and the transition probability to state 00 is 0.973. In Fig. 5 the paths of the steady state trellis is presented with its corresponding transition and state probabilities. The

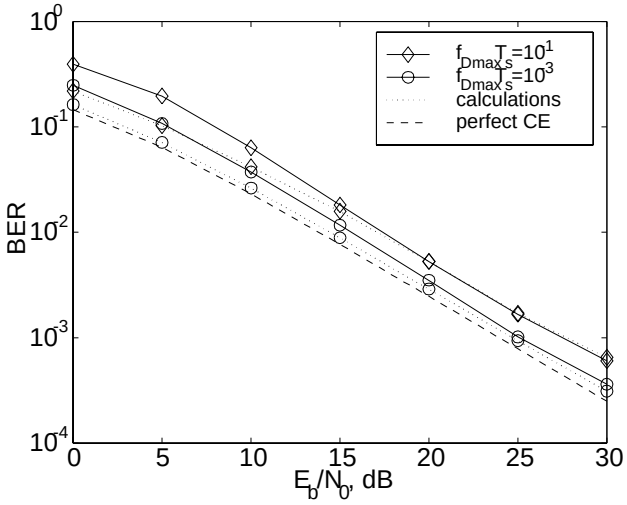

Figure 6: BER for different signal to noise ratios. The dotted lines are calculations assuming the zero data sequence and an m-sequence as pilot sequence.

state probabilities are achieved by the eigenvector of the transition matrix corresponding to the eigenvalue one [23]. An approximation of the overall BER is then the sum of the products between the state probabilities and the transition probabilities corresponding to a symbol error, in this case $P_b \approx 0.012$.

The method described above works well for high and medium SNR values. In Fig. 6 we plot the approximative BER using the method above together with simulated values of the BER for different SNR values. For the calculations and simulations we used $\rho = 0.05$ and $L = 6$. In Section 5 we motivate this specific choice further. The calculations were based on the assumption that the transmitted data was the zero sequence and the pilot sequence was an m-sequence. For low SNR the calculated values are somewhat lower than the simulated ones. Since we assumed that the length of the error event was limited, the two most probable hypotheses are the hypotheses returning back to the correct hypothesis directly after symbol k, the symbol of interest. This is true also for low SNR, but then the probability of other hypotheses, which we do not take into account, are high as well. Also, the BER is dependent on the relation between the data sequence and the pilot sequence. We assumed an m-sequence for the pilot sequence and the zero sequence for the data sequence. This assumption results in a low correlation between the pilot and data sequences and therefore the calculated values are somewhat lower than the simulated values. This dependence is especially large for large Doppler frequencies.

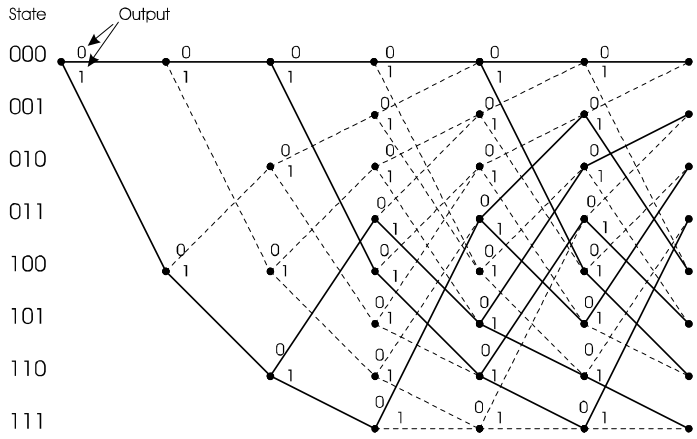

Figure 7: Possible state transitions in the uncoded (dashed line) and coded (solid line) case.

4.4 Influence of Channel Coding

If channel coding is used the number of possible paths in the trellis is reduced. This means a larger probability of choosing the correct path and thereby a larger possibility of using the correct hypotheses when calculating the channels estimates. Hence, besides the error correction capability, the code also implies that detection is improved. In Fig. 7 we show the possible paths in an eight-state trellis used without or with a convolutional code of rate $R = 1/2$, memory 2, and generator polynomial (5,7).

In the coded case all output combinations are not allowed and therefore there is, e.g., no feasible path between state 000 and state 100 in the second step. It is possible to use the same trellis for detection and for decoding. Instead of the conventional hard or soft decoding metric we use now the metric given by (9). Note the possible paths are the ones in the conventional four-state trellis for the (5,7)-code. For the four-state trellis we just use the code outputs as hypotheses and not the states directly.

5 Performance Evaluation by Simulations

We have verified our receiver given the following set-up: We used BPSK modulation for the data and pilot symbols (no staggering, no phase offset), i.e. real-valued symbols. For this particular modulation scheme, the performance could be improved by transmitting the pilot sequence as the quadrature component. However, here we want to demonstrate the feasibility of using a common channel instead of two independent channels. The pilot sequence was a long PN

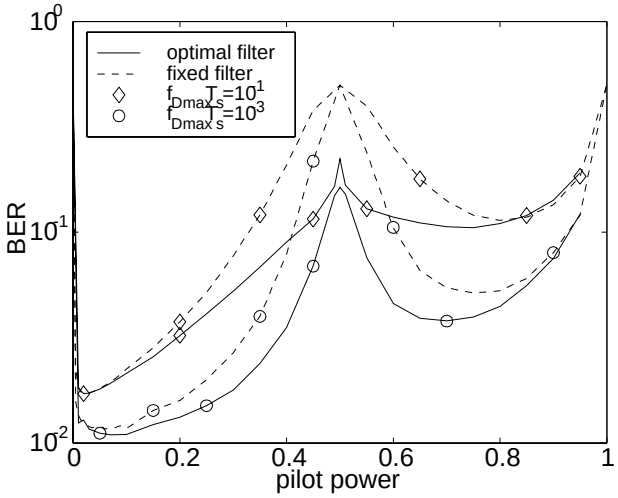

Figure 8: Resulting BER when using the optimal and fixed filter coefficients.

sequence known to the receiver, as opposed to a short sequence applied in [8]. Data and pilot symbols were generated by independent pseudo-random generators. We averaged our results over several thousands of possible pilot sequences. The block length was chosen to be 2000 symbols and the performance was studied for a wide range of different fading rates $f_{D_{max}}T_s$.

In the receiver, the proposed Viterbi based detector with a sufficiently long decision delay was applied. The Doppler spectrum and the SNR were assumed to be known in the channel estimation filter design. This assumption does not appear to be critical as indicated in related work. The BER was chosen as performance criteria. Focus was on the tracking phase, i.e. we did not take the first part of the block into account.

In a first set of Monte Carlo simulations we optimized the normalized power, ρ, of the superimposed pilot symbols. In Fig. 8 the resulting BER is plotted for different pilot power values when the optimal filter coefficients (where the channel estimation filter coefficients are adjusted to the power of the specific symbols) or the fixed filter coefficients are used. The signal to noise ratio per information bit was $E_b/N_0 = 15$ dB and the filter order was chosen to be $L = 6$, which corresponds to $2^{L-1} = 32$ states. For the interesting power range, i.e. where the BER and the pilot power is low, the performance is similar for both filters. The fixed filter coefficients can therefore be used, resulting only in a small performance degradation compared to the optimal filter coefficients. In Fig. 9 a detailed plot of the low pilot powers is shown. It appears that the BER is not sensitive with respect to ρ over a wide range. The optimum normalized

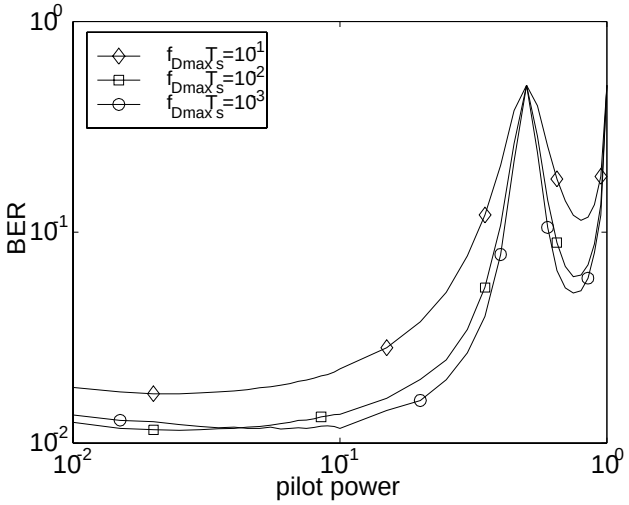

Figure 9: Optimization of the power, ρ, of the superimposed pilot sequence.

power is about $\rho = 0.02 \ldots 0.05$ for all Doppler frequencies of interest: The power of pilot symbols should only be about 2 % to 5 % of the power of the data symbols. In the following, $\rho = 0.05$ is applied.

In a second set of simulations we investigated the influence of the filter length L, and the results are presented in Fig. 10. Again, $E_b/N_0 = 15$ dB. It appears from the curves that a filter length of about $L = 6$ is sufficient for all Doppler frequencies of interest. Although not shown here, the mean squared error of the filter, which can be calculated analytically, shows the same principle behavior as the BER curve in the figure.

Given these optimizations, we simulated the BER versus E_b/N_0 for the scheme under investigation. The result is shown in Fig. 6. As a benchmark, the BER performance of 2-PSK on a flat Rayleigh fading channel given perfect channel estimation is plotted as well. For fading rates up to 10^{-2} the loss is less than 2 dB, and even for fast fading ($f_{D_{max}}T_s = 0.1$) no error floor is visible in the interesting range. This is in contrast to conventional pilot-symbol-assisted CE. For high Doppler frequencies pilot-symbol-assisted CE shows an error floor and the superimposed pilot sequence technique significantly outperforms this method.

In Fig. 11 we compare the BER for different Doppler values when using conventional pilot-symbol-assisted CE and superimposed pilot sequences. As a reference we also give the BER for perfect channel estimation. In both cases the pilot power is 5%. This means for the pilot-symbol-assisted scheme that every 20th symbol is a pilot symbol and that the energy per information bit is

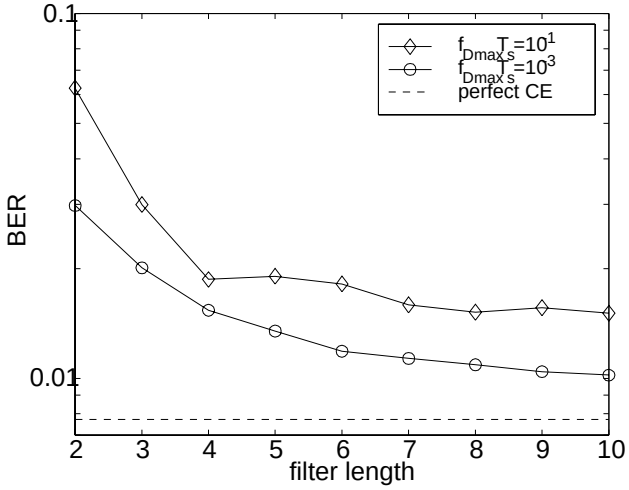

Figure 10: Influence of the channel estimation filter length.

decreased by 5%. As expected, this method shows large performance degradation at Doppler values above 0.025 due to the violation of the Nyquist sampling theorem. There the receiver is not able to track all channel changes and the BER performance is degraded. The superimposed pilot scheme, however, uses the L latest received symbols for channel estimation and therefore do not suffer from a too large distance between the pilot symbols.

6 Conclusions

In this paper, we explored the superimposed pilot sequence technique for the purpose of channel estimation. Due to the redundancy, truly coherent demodulation is achieved. The main contribution was the derivation and analysis of a receiver based on the Viterbi algorithm. We derived both the optimal metric, and a sub-optimal metric suitable for recursive detectors. The performance of the proposed receiver was then analyzed both analytically and by simulations. For the analytical analysis we derived the pairwise error probabilities and based on these we made approximations of the overall BER.

Compared to pilot-symbol-assisted CE, which is currently state-of-the-art, distinct advantages are as follows:

- No bandwidth expansion
- Better power efficiency in fast fading environments

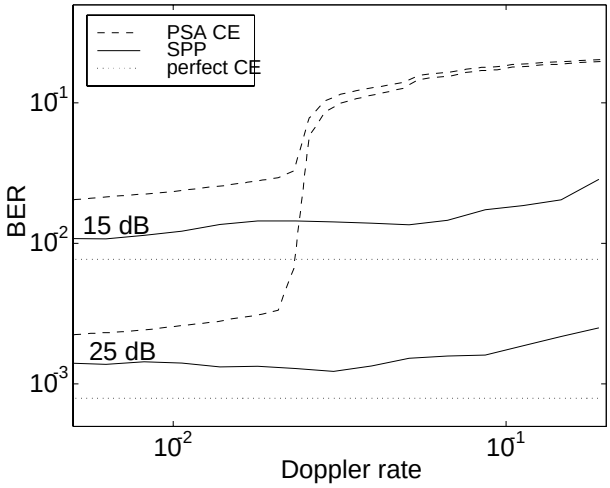

Figure 11: BER for perfect channel estimation, superimposed pilot sequences (SPP) or pilot-symbol-assisted CE (PSA CE). In the two latter cases 5% of the power are spent on pilots.

- The technique is more universally applicable: The same sequence may be used for time, phase, frequency, and frame synchronization and as a unique word without additional overhead. The concept of superimposed pilot sequences may be used as an "add-on" in existing systems.

- Pilot-symbol-assisted CE, with isolated pilot symbols, fails when the pilot symbols are affected by frequency-selective fading

- Data symbols are equally sensitive against transmission errors when acquisition is not considered.

The proposed Viterbi receiver is, however, more complex than a single interpolation filter due to the per-survivor processing. Therefore, simplified structures based on reduced-state sequence estimation were proposed as well. As outlined in the analysis, the performance depends on the block length and on the choice of the PN sequence.

Future work may be devoted to optimizations of the pilot sequence, an evaluation of M-ary modulation schemes, multi-carrier modulation schemes (e.g., OFDM with 2-D pilot array), and receiver structures with oversampling.

A Derivation of the metric

The probability density function of the received symbol vector is given by

$$f(\mathbf{y} \mid \mathbf{C}) = \frac{1}{(2\pi)^{N/2} \det^{1/2}(\boldsymbol{\phi})} \exp(-\frac{1}{2}(\mathbf{y}-\bar{\mathbf{y}})^H \boldsymbol{\phi}^{-1}(\mathbf{y}-\bar{\mathbf{y}})), \tag{25}$$

where $\bar{\mathbf{y}}$ denotes the mean of the vector $\mathbf{y}$ and $\boldsymbol{\phi}$ denotes the covariance matrix of the received samples, $\boldsymbol{\phi} = E[(\mathbf{y}-\bar{\mathbf{y}})(\mathbf{y}-\bar{\mathbf{y}})^H]$. The mean $\bar{\mathbf{y}}$ is not known a priori, but given the hypothesis $\tilde{\mathbf{C}}$ the mean is given by the channel estimates as

$$\bar{\mathbf{y}} = E[\mathbf{y} \mid \tilde{\mathbf{C}}] = \tilde{\mathbf{C}}\tilde{\mathbf{f}} = \tilde{\mathbf{C}}\mathbf{A}\tilde{\mathbf{C}}^{-1}\mathbf{y}. \tag{26}$$

The covariance matrix can now be calculated as

$$\begin{aligned} \phi &= E[(\mathbf{I} - \tilde{\mathbf{C}}\mathbf{A}\tilde{\mathbf{C}}^{-1})\mathbf{y}\mathbf{y}^H(\mathbf{I} - \tilde{\mathbf{C}}\mathbf{A}\tilde{\mathbf{C}}^{-1})^H] \\ &= (\mathbf{I} - \tilde{\mathbf{C}}\mathbf{A}\tilde{\mathbf{C}}^{-1})(\tilde{\mathbf{C}}\mathbf{R}_f^H\tilde{\mathbf{C}}^H + \sigma_n^2 \cdot \mathbf{I})(\mathbf{I} - \tilde{\mathbf{C}}\mathbf{A}\tilde{\mathbf{C}}^{-1})^H \end{aligned} \tag{27}$$

where $\mathbf{R}_f$ is the covariance matrix of the channel. Substituting (26) and (27) into (25) and taking the natural logarithm yields the log-likelihood function

$$\begin{aligned} \ln f(\mathbf{y} \mid \tilde{\mathbf{C}}) &= -\frac{1}{2}\ln((2\pi)^N \det(\boldsymbol{\phi})) - \frac{1}{2}(\mathbf{y} - \tilde{\mathbf{C}}\mathbf{A}\tilde{\mathbf{C}}^{-1}\mathbf{y})^H\tilde{\boldsymbol{\phi}}^{-1}(\mathbf{y} - \tilde{\mathbf{C}}\mathbf{A}\tilde{\mathbf{C}}^{-1}\mathbf{y}) \\ &= -\frac{1}{2}\ln((2\pi)^N \det(\boldsymbol{\phi})) - \frac{1}{2}\mathbf{y}^H(\tilde{\mathbf{C}}\mathbf{R}_f^H\tilde{\mathbf{C}} + \sigma_n^2 \cdot \mathbf{I})^{-1}\mathbf{y} \end{aligned} \tag{28}$$

Maximizing this is equivalent to minimizing

$$\begin{aligned} \tilde{\Lambda}_{opt} &= \ln(\det(\boldsymbol{\phi})) + (\mathbf{y} - \tilde{\mathbf{C}}\mathbf{A}\tilde{\mathbf{C}}^{-1}\mathbf{y})^H\tilde{\boldsymbol{\phi}}^{-1}(\mathbf{y} - \tilde{\mathbf{C}}\mathbf{A}\tilde{\mathbf{C}}^{-1}\mathbf{y}) \\ &= \ln(\det(\tilde{\boldsymbol{\phi}})) + \mathbf{y}^H(\tilde{\mathbf{C}}\mathbf{R}_f^H\tilde{\mathbf{C}}^H + \sigma_n^2 \cdot \mathbf{I})^{-1}\mathbf{y} \end{aligned} \tag{29}$$

which is the optimal metric.

The goal is now to derive suitable metric increments such that the Viterbi algorithm can be used. If the power of the pilot sequence is small, then the difference between the determinants for the different hypotheses is small and therefore this term is neglected. Further, assuming that the inverse of the covariance matrix is diagonally dominant, then it can be approximated as an identity matrix and we get the desired additive metric as

$$\tilde{\Lambda} = (\mathbf{y} - \tilde{\mathbf{C}}\mathbf{A}\tilde{\mathbf{C}}^{-1}\mathbf{y})^H(\mathbf{y} - \tilde{\mathbf{C}}\mathbf{A}\tilde{\mathbf{C}}^{-1}\mathbf{y}). \tag{30}$$

References

[1] J. G. Proakis. *Digital communications*. Prentice-Hall, 3rd edition, 1995.

[2] H. Meyr, M. Moneclaey, and S. Fechtel. *Digital Communication Receivers*. Wiley, New York, USA, 1998.

[3] M. L. Moher and J. H. Lodge. TCMP – A modulation and coding strategy for Rician-fading channels. *IEEE J. Select. Areas Commun.*, 7(9):1347–1355, Dec. 1989.

[4] M. Moher and J. Lodge. A time diversity modulation strategy for the satellite-mobile channel. In *13th Biennial Symp. Commun.*, Kingston, Canada, June 1986.

[5] A. Aghamohammadi, H. Meyr, and G. Ascheid. A new method for phase synchronization and automatic gain control of linearly-modulated signals on frequency-flat fading channels. *IEEE Trans. Commun.*, 39(1):25–29, Jan. 1991.

[6] J. K. Cavers. An analysis of pilot-symbol assisted modulation for Rayleigh-fading channels. *IEEE Trans. Vehic. Technol.*, 40(4):686–693, Nov. 1991.

[7] D. Makrakis and K. Feher. A novel pilot insertion-extraction method based on spread spectrum techniques. In *Miami Technicon*, Miami, USA, 1987.

[8] T. Holden and K. Feher. A spread spectrum based system technique for synchronization of digital mobile communication systems. *IEEE Trans. Broadc.*, pages 185–194, Sept. 1990.

[9] K. J. Åström and B. Wittenmark. *Adaptive Control*. Addison-Wesley, Reading, Massachusetts, second edition, 1995.

[10] A. Steingass, A. Wijngaarden, and W. Teich. Frame synchronization using superimposed sequences. In *Proc. IEEE Intern. Symp. Inf. Theory*, page 489, Ulm, Germany, June 1997.

[11] F. Tufvesson, M. Faulkner, P. Hoeher, and O. Edfors. OFDM time an frequency synchronization by spread spectrum pilot technique. In *Eighth Communication Theory Mini Conference in conjunction with IEEE ICC '99*, pages 115–119, Vancouver, Canada, June 1999.

[12] E. Bejjani, J.-C. Belfiore, and P. Leclair. Coherent detection for transmission over severly time and frequency dispersive multipath fading channels. In *Proc. IEEE Intern. Symp. Inf. Theory*, page 212, Whistler, Canada, Sept. 1995.

[13] E. Dahlman, P. Beming, J. Knutsson, F. Ovesjö, M. Persson, and C. Roobol. WCDMA - the radio interface for future mobile multimedia communications. *IEEE Trans. Vehic. Technol.*, 47(4):1105–1118, Nov. 1998.

[14] P. Hoeher and F. Tufvesson. Channel estimation with superimposed pilot sequence applied to multi-carrier systems. In *International Workshop on Multi-Carrier Spread-Spectrum and Related Topics*, pages 295–302, Oberpfaffenhofen, Germany, Sept. 1999.

[15] J. Lodge and M. Moher. Maximum likelihood sequence estimation of CPM signals transmitted over Rayleigh flat-fading channels. *IEEE Trans. Commun.*, 38(6):787–794, June 1990.

[16] R. Raheli, A. Polydoros, and C.-K. Tzou. Per-survivor processing: A general approach to MLSE in uncertain environments. *IEEE Trans. Commun.*, 43(2/3/4):354–364, Feb/Mar/Apr 1995.

[17] M. V. Eyuboglu and S. U. Qureshi. Reduced-state sequence estimation with set partitioning and decision feedback. *IEEE Trans. Commun.*, 36(1):13–20, Jan. 1988.

[18] A. Duel-Hallen and C. Heegard. Delayed decision-feedback sequence estimation. *IEEE Trans. Commun.*, 37(5):428–436, May 1989.

[19] M. J. Barret. Error probability for optimal and suboptimal quadratic receivers in rapid Rayleigh fading channels. *IEEE J. Select. Areas Commun.*, SAC-5(2):302–304, Feb. 1987.

[20] C. Schlegel. Error probability calculation for multibeam Rayleigh channels. *IEEE Trans. Commun.*

[21] A. Mathai and S. B. Provost. *Quadratic forms in random variables: Theory and applications*. Marcel Dekker, New York, U.S.A., 1992.

[22] J. K. Cavers and P. Ho. Analysis of the error performance of trellis-coded modulation in Rayleigh-fading channels. *IEEE Trans. Commun.*, 40(1):74–83, Jan. 1992.

[23] G. Strang. *Linear Algebra and Its Applications*. Academic Press, 3rd edition, 1988.

Paper G

This part is a revised version of:

Fredrik Tufvesson and Ove Edfors, "A Comparaitve Analysis Between Different Detection Schemes in Flat Rayleigh-fading Channels", Proceedings of IEEE Vehicular Technology Conference, Tokyo, Japan, pp. 1205-1209, May 2000.

A Comparative Analysis Between Different Detection Schemes in Flat Rayleigh-fading Channels

Fredrik Tufvesson and Ove Edfors
Department of Applied Electronics, Lund University,
Box 118, SE-221 00 Lund, Sweden
Phone: +46-46 222 30 21, Fax: +46-46 12 99 48,
E-mail: Fredrik.Tufvesson@tde.lth.se

Abstract

We compare the performance of different detection schemes in terms of resulting pairwise error probability in Rayleigh-fading channels. We introduce a unified analysis framework used to compare sensitivity to Doppler and noise. The analyzed methods include: pilot symbol assisted modulation, differential detection, multiple-symbol differential detection and detection using superposed pilot symbols. We also introduce an improved hybrid detection scheme, combining the strengths of pilot symbol assisted modulation and superposed pilot symbols.

1 Introduction

It is well known that coherent detection results in good performance. It also enables the use of multilevel signalling but requires reliable phase and/or channel estimation. Differential detection, on the other hand, is robust and requires no channel estimation, but there is a performance loss compared to coherent detection and it is mostly used for phase shift keying (PSK). Many new schemes have been presented to enhance the performance of both coherent and differential methods. In this paper we make an analytical comparison between some of these detection schemes. The analysis is based on a unified framework where we calculate the pairwise probability that an erroneous sequence is detected in favor of the correct one. The analyzed methods include: conventional differential detection (Diff), multiple-symbol differential detection (MSDD), pilot symbol assisted modulation (PSAM), detection using superposed pilot symbols (SPP), and a novel coherent hybrid detection scheme (Hybr).

The MSDD was introduced for AWGN channels in [1] and extended to fading channels in [2] and [3]. The idea behind this technique is to simultaneously

detect a group of phase shifts in order to have more than one symbol as a reference. In this way the reference is less affected by noise and the performance is thereby enhanced. The analysis in [3] uses expressions for the pairwise error probabilities and a union bound to calculate an approximative bit-error rate. In [4] a detector based on the Viterbi algorithm was presented, which we use in our analysis. This Viterbi-based detector is essentially the same as the decision-directed coherent detector introduced in [5], except that in the latter one symbol-by-symbol decisions are made instead of sequence decisions.

PSAM was originally proposed in [6] and later analyzed in [7]. In this scheme known pilot symbols are multiplexed into the data sequence. The received pilot symbols makes it possible to observe the influence of the channel at the pilot positions. The channel estimates for the data symbols are then achieved by interpolation. Since the pilot symbols contain no data there is a small overhead causing a bandwidth expansion and an energy loss. However, the method has shown to be robust in fading environments with only a small loss compared to ideal coherent detection. In addition, it also allows multilevel signalling.

The use of superposed pilot sequences was introduced in [8] for carrier phase tracking. As opposed to PSAM, the pilot sequence is added to the data sequence (superposed) and thereby transmitted in the same frequency band and at the same time. Since the pilot sequence and the data sequence are superposed there is no increase in bandwidth. In [9] this technique was proposed for coherent detection of data symbols. The SPP scheme is robust against high Doppler frequencies and it enables multilevel signalling. However, the complexity grows with the number of signalling alternatives.

The strengths of PSAM and superposed pilot symbols can be combined into a new hybrid detection scheme where not only known pilot symbols, but also hypotheses of the unknown data symbols, are used for channel estimation. In [10] a similar scheme was presented using decisions of the unknown data symbols for consecutive phase corrections of the channel estimate. Our new scheme uses *hypotheses* of the data symbols as virtual pilot symbols and it is therefore less sensitive to decision errors.

The contributions of this paper include both the new hybrid detection scheme and a unified analysis framework for pairwise error events, that can be applied to all of the above detection schemes. Though beyond the scope of this paper, the pairwise error probabilities presented can be used to calculate bounds for the resulting bit error rate (BER) using the most probable error sequences in the same way as in [3].

2 Detection Schemes

The analyzed detection methods, including the differential ones, require some kind of "channel estimate" for detection. Conventional differential detection

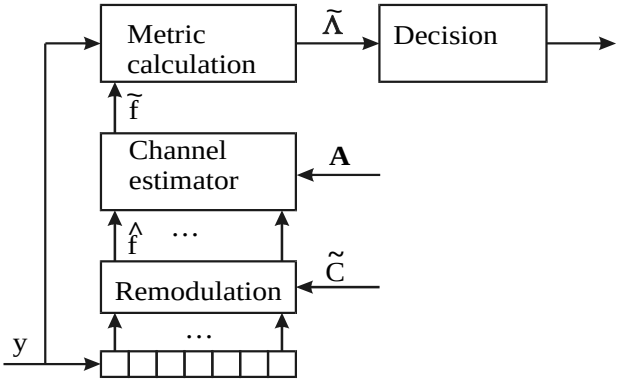

Figure 1: Block scheme of the detector used for the analysis.

uses e.g. the last received symbol as a channel estimate. In PSAM the received pilot symbols are used to calculate channel estimates. When using SPP the channel estimates are based on the known pilot sequence together with hypotheses of the unknown data sequence. The general receiver used for the analysis of all detection methods is presented in Fig.1.

In the following, we use complex baseband notation and assume perfect time, frequency and frame synchronization. The transmitted data symbols are denoted as $c_k \in \mathbb{C}$. The symbol c_k is either a data symbol, u_k, a pilot symbol, p_k, or a combination of both, $c_k = \sqrt{1-\rho_k} \cdot u_k + \sqrt{\rho_k} \cdot p_k$, where ρ_k is the amount of energy spent on pilot information at time k. Let y_k denote the received symbol samples,

$$y_k = c_k \cdot f_k + n_k, \tag{1}$$

where $f_k \in \mathbb{C}$ is a multiplicative (time-selective) fading process with $E[|f_k|^2] = 1$, and $n_k \in \mathbb{C}$ are zero-mean white Gaussian noise samples with one-sided power spectral density N_0. Without loss of generality, the average energy of the received symbols is $E_s = E[|c_k f_k|^2] = 1$. It is convenient to use matrix notation and we denote the received sample vector by $\mathbf{y} = (y_1, y_2, ..., y_N)^T = \mathbf{Cf} + \mathbf{n}$, where $\mathbf{f}$ and $\mathbf{n}$ are vectors of the fading process and the noise, respectively, and $\mathbf{C}$ is a diagonal matrix with the transmitted symbols on its diagonal.

The goal is to compute the maximum-likelihood estimate of the transmitted sequence. The receiver tests all possible hypotheses and decides for the sequence resulting in the lowest metric. For Constant-envelope modulation, the optimal metric is given [3] as

$$\tilde{\Lambda} = \mathbf{y}^H \tilde{\mathbf{C}} (\Phi + \mathbf{I})^{-1} \tilde{\mathbf{C}}^H \mathbf{y}, \tag{2}$$

where $\tilde{\mathbf{C}}$ is a diagonal matrix of the hypothesis to be tested (defined later on) and Φ is the covariance matrix of the channel. Since it is advantageous to have

an additive metric increment for each new symbol one often use approximations instead. The noise is assumed to be Gaussian, and therefore a suitable metric increment is

$$\tilde{\Lambda}_k(\tilde{c}_k; \tilde{c}_{k-P_0}, \ldots, \tilde{c}_{k-P_L}) = |y_k - \tilde{c}_k \cdot \tilde{f}_k|^2, \tag{3}$$

where $\tilde{c}_k$ is a hypothesis for the k-th data symbol and $\tilde{f}_k$ is a channel estimate. The latter is derived from the data hypothesis $[\tilde{u}_{k-P_0}, \ldots, \tilde{u}_{k-P_L}]$ and/or pilot symbols $[p_{k-P_0}, \ldots, p_{k-P_L}]$. The parameter L is the number of symbols used for the channel estimate and P_l denotes the time indices of the received symbols used for the estimate.

In order to make decisions, the receiver derives channel estimates and using these it tests all possible hypotheses against the received symbols. The channel estimates are calculated as follows: A tentative channel estimate is obtained by re-modulation, $\hat{f}_{k-P_l} = y_{k-P_l}/\tilde{c}_{k-P_l}$. Given the hypothesis and the pilot symbols, the channel estimate $\tilde{f}_k$ can be computed by linear filtering [11]:

$$\tilde{f}_k = \sum_{l=0}^{L-1} a_{k,P_l} \cdot \hat{f}_{k-P_l} = \sum_{l=0}^{L-1} a_l \cdot \frac{y_{k-P_l}}{\tilde{c}_{k-P_l}}. \tag{4}$$

The filter coefficients $a_{k,m}$ are dependent on the detection method used. For conventional differential detection only the last symbol is used for the channel estimate so $L = 1, P_0 = 1$, $a_1 = 1$, and all other filter coefficients are zero. If a Wiener estimator is used, as in e.g. PSAM or SPP, the filter coefficients are obtained as the solution of the Wiener-Hopf equations.

The channel estimate at time k, based on a certain hypothesis, is given by the previous received symbols as

$$\tilde{f}_k = \mathbf{a}_k \tilde{\mathbf{C}}^{-1} \mathbf{y}, \tag{5}$$

where $\mathbf{a}_k$ is a vector of the channel estimation coefficients $a_{k,m}$, e.g. the Wiener filter coefficients, and $\tilde{\mathbf{C}}$ is a diagonal matrix containing the hypothesis of the transmitted symbols on its diagonal, i.e. $\tilde{c}_{k,k} = \sqrt{1-\rho_k} \cdot u_k + \sqrt{\rho_k} \cdot p_k \tilde{u}_k$. The corresponding vector of channel estimates for a certain hypothesis can consequently be calculated as

$$\tilde{\mathbf{f}} = \mathbf{A} \tilde{\mathbf{C}}^{-1} \mathbf{y}. \tag{6}$$

The elements in the channel estimation matrix $\mathbf{A}$ and the remodulation matrix $\tilde{\mathbf{C}}^{-1}$ are determined by the detection method. These are described below.

If symbol-by-symbol detection is used then the metric increment (3) is used for the decisions and if sequence detection is used then a sum of these is used instead. Adding the metric increments and substituting the channel estimates by (6) leads to the cumulative metric

$$\tilde{\Lambda} = \mathbf{y}^H (\mathbf{I} - \tilde{\mathbf{C}} \mathbf{A} \tilde{\mathbf{C}}^{-1})^H (\mathbf{I} - \tilde{\mathbf{C}} \mathbf{A} \tilde{\mathbf{C}}^{-1}) \mathbf{y}. \tag{7}$$

Below we shortly describe the analyzed detection methods and give their respective channel estimation matrices and hypothesis matrices.

2.1 Pilot Symbol Assisted Modulation

When using PSAM we transmit known pilot symbols at certain time slots and only these symbols are used to calculate the channel estimates. The channel estimation matrix has the filter coefficients, a_{k,P_l}, at positions where the pilot symbols are located, all other elements are zero, i.e.

$$\mathbf{A}_{PSAM} = \begin{bmatrix} a_{1,P_0} & 0 & \cdots & 0 & a_{1,P_2} & 0 & \cdots \\ a_{2,P_0} & 0 & \cdots & 0 & a_{2,P_2} & 0 & \cdots \\ \vdots & & & & \vdots & & \\ a_{N,P_0} & 0 & \cdots & 0 & a_{N,P_2} & 0 & \cdots \end{bmatrix}.$$

The hypothesis matrix is diagonal and has the known pilot symbol values at their corresponding positions and the values for the data hypothesis to be tested at the other positions, i.e.

$$\tilde{\mathbf{C}}_{PSAM} = diag(p_1, \tilde{u}_2, \dots, \tilde{u}_{k-1}, p_k, \tilde{u}_{k+1}, \dots).$$

2.2 Superposed Pilot Symbol Detection

When using superposed pilot symbols the known pilot sequence is superposed on the unknown data sequence. The hypotheses are therefore a combination of both known and unknown values and the hypothesis matrix have the form

$$\tilde{\mathbf{C}}_{SPP} = diag\left(\sqrt{1-\rho}(\tilde{u}_1, \dots, \tilde{u}_N) + \sqrt{\rho}(p_1, \dots, p_N)\right)$$

whereas the channel estimation matrix have the form

$$\mathbf{A}_{SPP} = \begin{bmatrix} a_{1,0} & 0 & 0 & \cdots & 0 \\ a_{2,1} & a_{2,0} & 0 & \cdots & 0 \\ \vdots & \vdots & & & \\ 0 & 0 & a_{N,L-1} & \cdots & a_{N,0} \end{bmatrix}.$$

2.3 Multiple-Symbol Differential Detection

In multiple-symbol differential detection several differentially encoded symbols are detected at the same time. The decision is taken over N symbols and the receiver aims to find a data sequence that aligns the received phases not only between consecutive symbols, but also between non-neighboring ones. The references are calculated using hypotheses in the same way as for superposed pilot symbols, so $\mathbf{A}_{MSDD} = \mathbf{A}_{SPP}$. There is no known information transmitted so therefore the hypothesis matrix consists of the unknown data symbols only, $\tilde{\mathbf{C}}_{MSDD} = diag(\tilde{u}_1, \dots, \tilde{u}_N)$.

2.4 Hybrid Detection

The hybrid detector uses both pilot symbols and hypotheses to form the channel estimates. The pilot symbols ensure that there is less phase ambiguity and the hypotheses helps the channel estimator to track the channel variations between pilot symbols. In this way it is possible to, e.g., manage large Doppler frequencies without having too much overhead due to pilot symbols. Here again, the channel estimation matrix has the same form as for superposed pilot symbols, $\mathbf{A}_{Hybr} = \mathbf{A}_{SPP}$, but the hypothesis matrix is the same as for PSAM, i.e., $\tilde{\mathbf{C}}_{Hybr} = \tilde{\mathbf{C}}_{PSAM}$.

2.5 Differential Detection

In conventional differential detection the decisions are made symbol-by-symbol. The probability of correct decisions can be calculated in the unified framework by the channel estimation matrix $\mathbf{A}_{DD} = \begin{bmatrix} 1 & 0 \end{bmatrix}$ and the hypothesis matrix

$$\tilde{\mathbf{C}}_{DD} = \begin{bmatrix} \tilde{u}_{k-1} & 0 \\ 0 & \tilde{u}_k \end{bmatrix}.$$

As opposed to the other methods, the decisions are made symbol by symbol when using conventional differential detection. Therefore, the analysis presented in next section is used to study only single-error events for this scheme.

3 Detector Performance

In this section we derive analytical expressions for the probability of pairwise error events. The analysis is to some extent similar to the one in [3], but we also provide closed form expressions for the distribution of the decision variable.

The receiver makes a correct decision if the metric for the correct path is smaller than all other path metrics. Define the decision variable $\hat{\Lambda} = \tilde{\Lambda} - \Lambda$,where $\tilde{\Lambda}$ and Λ are the cumulative metrics of the hypothesis and correct path, respectively. If $\hat{\Lambda} > 0$ for all hypotheses then the correct path will be chosen. By rewriting the received samples as $\mathbf{y} = \mathbf{Dg}$ with $\mathbf{D} = \begin{bmatrix} \mathbf{C} & \mathbf{I} \end{bmatrix}$ and $\mathbf{g} = \begin{bmatrix} \mathbf{f} & \mathbf{n} \end{bmatrix}^T$, the decision variable can be expressed

$$\begin{aligned} \hat{\Lambda} &= \mathbf{y}^H[(\mathbf{I} - \tilde{\mathbf{C}}\mathbf{A}\tilde{\mathbf{C}}^{-1})^H(\mathbf{I} - \tilde{\mathbf{C}}\mathbf{A}\tilde{\mathbf{C}}^{-1}) \\ &\quad -(\mathbf{I} - \mathbf{C}\mathbf{A}\mathbf{C}^{-1})^H(\mathbf{I} - \mathbf{C}\mathbf{A}\mathbf{C}^{-1})]\mathbf{y} \\ &= \mathbf{g}^H\tilde{\mathbf{K}}\mathbf{g}. \end{aligned} \tag{8}$$

The covariance matrix of the zero-mean Gaussian variables in $\mathbf{g}$ is

$$\mathbf{R}_g = \begin{bmatrix} \mathbf{R}_f & \mathbf{0} \\ \mathbf{0} & \sigma_n^2 \cdot \mathbf{I} \end{bmatrix}, \tag{9}$$

where $\mathbf{R}_f$ is the covariance matrix of the fading process. For 2-D isotropic scattering $\mathbf{R}_f$ has elements $r_{i,j} = J_0(2\pi f_{D\,\max} T_s \cdot (i-j))$, where $J_0(\cdot)$ is the zeroth order Bessel function, $f_{D\,\max}$ is the maximum Doppler frequency and T_s is the symbol time.

3.1 Pairwise Error Probability

Many authors have studied the pairwise error probability resulting from quadratic forms like (8), e.g. in [12][13]. Following them we define new Gaussian variables $\mathbf{g}' = \mathbf{B}\mathbf{g}$, where $\mathbf{B}$ is a matrix of size $2N \times r$ fulfilling $\mathbf{R}_g = \mathbf{B}\mathbf{B}^H$ and $|\mathbf{B}| \neq 0$, where r is the rank of $\mathbf{R}_g$. In [12] it is shown that the decision variable can be regarded as a sum of r independent weighted χ^2-variables, where the weights are the eigenvalues of $\mathbf{B}^H\,\tilde{\mathbf{K}}\mathbf{B}$. The vector of Gaussian samples is complex and the χ^2-variables thereby have 2 degrees of freedom.

By sorting the weights so that the first m eigenvalues are positive and the last $N_s - m$ eigenvalues are negative the decision variable can be expressed as

$$\hat{\Lambda} = \sum_{n=1}^{m} \lambda_n \chi^2_{2,n} - \sum_{n=m+1}^{N_s} -\lambda_n \chi^2_{2,n.} \tag{10}$$

Assuming that all eigenvalues are distinct, the probability density function of $\hat{\Lambda}$ can be calculated as [12]

$$f_{\hat{\Lambda}}(y) = \begin{cases} \sum\limits_{n=1}^{m} \dfrac{(\lambda_n)^{N_s-2} e^{-y/(2\lambda_n)}}{2\left[\prod\limits_{k=1,k\neq n}^{N_s} (\lambda_n - \lambda_k)\right]} & y \geq 0 \\ \sum\limits_{n=m+1}^{N_s} \dfrac{(-\lambda_n)^{N_s-2} e^{-y/(2\lambda_n)}}{2\left[\prod\limits_{k=1,k\neq n}^{N_s} (\lambda_k - \lambda_n)\right]} & y < 0. \end{cases} \tag{11}$$

The pairwise error probability, i.e. the probability that the decision variable becomes negative so that the receiver makes a decision in favor of the erroneous hypothesis s instead of the correct one, is then given by

$$P_s = \sum_{n=m+1}^{N_s} \frac{(-\lambda_n)^{N_s-1}}{\prod\limits_{k=1,k\neq n}^{N_s} (\lambda_k - \lambda_n)}. \tag{12}$$

4 Comparison based on Error Events

The analysis tools provided in the previous section can be used to determine approximate bit- and frame-error rates for the various detections methods. However, here we only study the pairwise error probabilities (12) for some selected error events. This will reveal some of the inherent strengths and weaknesses of the different detection methods.

First we exemplify the probability of error events of different lengths in a block-fading Rayleigh channel ($f_D T_s = 0$). This shows the ambiguity of antipodal sequences for superposed pilots in coherent systems and the influence of a memory in the detector for differential systems. Secondly, we exemplify the robustness of these methods against variations in average SNR and Doppler effects. This singles out the poor performance of PSAM and differential detection at large Doppler spreads and high SNRs, when compared to the other methods.

We investigate pairwise error probabilities for runs of k consecutive trellis stages using BPSK. The blocks are of length $N = 51$, which is assumed to be long enough to approximate the steady state behavior, and the channel estimation filter lengths are $L = 5$. The k stages of deviation from the zero-path of the trellis are placed in the middle of the block and we have two kinds of interesting error events: error sequences where k is small so that the erroneous sequence merges with the correct sequence after k symbols and error sequences where the last (half) part is erroneous. The latter represents a single symbol error after differential detection. We separate the coherent and differential schemes in the examples. For the coherent methods every deviation from the true sequence will result in a symbol error. For the differential methods only the transitions between the correct and erroneous sequence cause symbol errors after differential decoding. This means that an error event of k consecutive errors only produces two final symbol errors.

In Fig. 2 we illustrate the pairwise error probability for error events of length 1, 2, 3, 4, and 26 (long) in a block-fading Rayleigh channel and an E_b/N_0 of 10 dB[1]. We have 5% pilot energy for the superposed pilot scheme and every 10th symbol as a pilot symbol in the other two coherent schemes. The schemes show rather similar behavior for short events, but separate for long events. The poorer performance of the SPP and hybrid schemes for long error events is explained by the inherent ambiguity of antipodal sequences since there is a possibility that the wrong hypothesis is used for channel estimation.

For the MSDD scheme, there are only two symbol errors for short error events and one symbol error for the long event. An interesting behavior for this scheme is that two symbol errors close to each other have a higher probability than two bit errors separated from each other. This is due to the memory of the MSDD, where the negative influence of earlier symbol errors decreases with time.

In Fig. 3 we present the pairwise error probability of events causing a single bit error in the transmission block. For the coherent methods this corresponds to an error event of length one, while for the differential schemes it is the equivalent of a long error event. Among the coherent schemes in Fig. 3 (top), the difference is relatively small for the block fading channel ($f_d T_s = 0$). However, when the Doppler frequency increases ($f_d T_s = 0.05$), the limitation of PSAM

[1] In the original publication a curve for conventional differential detection was also presented. In this revised version the curve is omitted and the text has been changed accordingly.

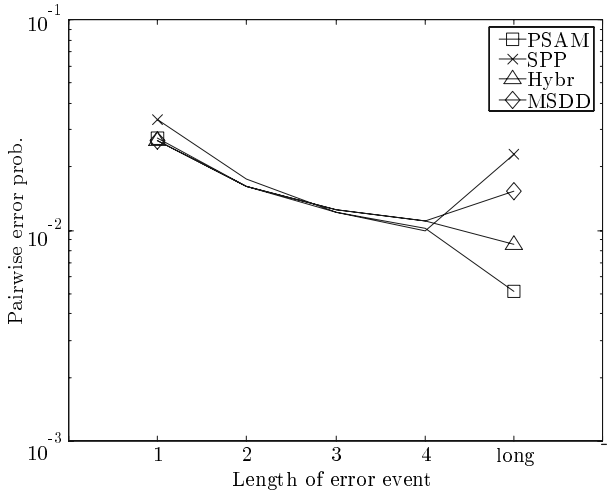

Figure 2: Pairwise error probability for error sequences of different lengths.

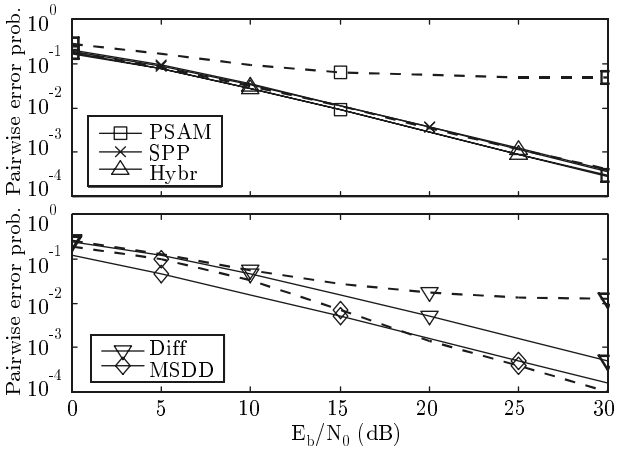

Figure 3: Pairwise error probability for different SNR values. Solid lines represent no Doppler ($f_D T_s = 0$) whereas dashed lines represent fast fading ($f_D T_s = 0.05$).

becomes apparent. Due to the limitations inflicted by the sampling theorem, the channel estimation in PSAM becomes poor (only one of 10 symbols is a pilot) and the resulting error floor is quite severe. With superposed pilots we are not close to the limit of the sampling theorem at the higher Doppler frequencies, and the pairwise error probability is hardly affected. This clearly shows the strength of superposed pilots at high Doppler frequencies, keeping in mind that the amount of energy spent on pilot information is even less in this case compared to PSAM.

For the differential schemes in Fig. 3 (bottom) we observe the increased performance by introducing a memory in the differential detection. This is true for both the block-fading channel ($f_d T_s = 0$) and when the channel is fading within the block ($f_d T_s = 0.05$). At high Doppler frequency and low SNRs, the error probability is dominated by the additive noise and the two detection methods have rather similar performance. However, when the SNR increases, the drawback of conventional differential detection becomes apparent. Due to phase changes in the channel between two consecutive symbols, the differential detector experience a quite severe error floor, whereas the multiple-symbol differential detector with its memory can compensate for these.

Using a single analytical expression for pairwise error events, we have managed to capture many of the inherent characteristics of a whole class of transmission/detection schemes. By identifying significant error events it is possible to derive a (truncated) union bound for the error rates, and thereby estimating the overall performance. This is, however, out of scope for the current paper and left for further research.

5 Conclusions

We have analyzed the performance and behavior of different detection schemes in Rayleigh fading channels. Based on a unified analytical framework we have calculated the pairwise error probability for certain coherent and non-coherent detection methods. By studying the variation of the pairwise error probabilities we have compared the robustness of these methods against variations in average SNR and Doppler effects. As expected, the schemes using a group of symbols close to the symbol to be detected to derive a "channel estimate" are shown to be very robust in fast fading environments and show no error floor. Conventional differential detection, on the other hand, suffers from a noisy "channel estimate" whereas PSAM may suffer from an obsolete channel estimate in fast fading channels. The latter can be avoided by increasing the number of pilot symbols, but then the overhead increases. In this paper we have also introduced a new detection scheme that combines the strengths of using pilot- *and* data symbols to derive channel estimates. The pilot symbols ensures that there is less phase ambiguity while the use of unknown data symbols makes it possible to track channel variations between the pilot symbols.

The presented analysis is general and can easily be applied to other detection schemes. For a further analysis, the expressions of the pairwise error probability can be used to get approximative expressions of the resulting bit and frame error rates by using union bound techniques.

Acknowledgement

The authors would like to thank Prof. Peter Hoeher at University of Kiel, Germany, for inspiring this work.

References

[1] Dariush Divsalar and Marvin K. Simon. Multiple-symbol differential detection of MPSK. *IEEE Trans. Commun.*, 38(3):300–308, March 1990.

[2] Dariush Divsalar and Marvin K. Simon. Maximum-likelihood differential detection of uncoded and trellis-coded amplitude phase modulation over AWGN and fading channels – Metrics and performance. *IEEE Trans. Commun.*, 42(1):76–89, January 1994.

[3] Paul Ho and Dominic Fung. Error performance of multiple-symbol differential detection of PSK signals transmitted over correlated Rayleigh fading channels. *IEEE Trans. Commun.*, 40(10):1566–1569, October 1992.

[4] Giorgio M. Vitetta and Desmond P. Taylor. Viterbi decoding of differentially encoded PSK signals transmitted over Rayleigh frequency-flat fading channels. *IEEE Trans. Commun.*, 43(2/3/4):1256–1259, 1995.

[5] Arne Svensson. Coherent detector based on linear prediction and decision feedback for DQPSK. *Electron. Lett.*, 30(20):1642–1643, September 1994.

[6] Michael L. Moher and John H. Lodge. TCMP – A modulation and coding strategy for Rician-fading channels. *IEEE J. Select. Areas Commun.*, 7(9):1347–1355, December 1989.

[7] James K. Cavers. An analysis of pilot-symbol assisted modulation for Rayleigh-fading channels. *IEEE Trans. Vehic. Technol.*, 40(4):686–693, November 1991.

[8] T.P. Holden and K. Feher. A spread spectrum based system technique for synchronization of digital mobile communication systems. *IEEE Trans. Broadc.*, pages 185–194, September 1990.

[9] Peter Hoeher and Fredrik Tufvesson. Channel estimation with superimposed pilot sequence. In *Proc. Globecom*, pages 2162–2166, Rio de Janeiro, Brazil, December 1999.

[10] Garrick T. Irvine and Peter J. McLane. Symbol-aided plus decision-directed reception for PSK/TCM modulation on shadowed mobile satellite fading channels. 10(8):1289–1299, October 1992.

[11] J.H. Lodge and M.L. Moher. Maximum likelihood sequence estimation of CPM signals transmitted over Rayleigh flat-fading channels. *IEEE Trans. Commun.*, 38(6):787–794, June 1990.

[12] A.M. Mathai and Serge B. Provost. *Quadratic forms in random variables: Theory and applications*. Marcel Dekker, New York, U.S.A., 1992.

[13] Christian Schlegel. Error probability calculation for multibeam Rayleigh channels. *IEEE Trans. Commun.*