Kanaalschatting en synchronisatie voor multiplexering
met orthogonale frequentie-opdeling en gekendesymboolopvulling

Channel Estimation and Synchronization
for Orthogonal Frequency Division Multiplexing
with Known Symbol Padding

Dieter Van Welden

Promotor: prof. dr. ir. H. Steendam
Proefschrift ingediend tot het behalen van de graad van
Doctor in de Ingenieurswetenschappen

Vakgroep Telecommunicatie en Informatieverwerking
Voorzitter: prof. dr. ir. H. Bruneel
Faculteit Ingenieurswetenschappen en Architectuur
Academiejaar 2011 - 2012

ISBN 978-90-8578-496-8
NUR 959
Wettelijk depot: D/2012/10.500/22

Dankwoord

Voila, vandaag, op het moment dat ik dit schrijf, ben ik begonnen aan het laatste en misschien ook wel het belangrijkste onderdeel van dit werk: het dankwoord. Het is vanzelfsprekend dat ik dit doctoraat nooit tot een succesvol einde kon brengen zonder de hulp van enkele mensen.

Vooreerst wil ik mijn promotor Heidi Steendam bedanken voor de geboden kans om gedurende zes jaar te werken aan dit doctoraat. Bovendien heeft ze er steeds voor gezorgd dat onze wetenschappelijke resultaten van voldoende kwaliteit waren zodat ze konden gepubliceerd worden. Daarnaast wil ik ook Marc Moeneclaey bedanken. Zijn aandeel in dit werk is minstens even groot. Het is verbazingwekkend wat er allemaal op een schijnbaar eenvoudige manier kan afgeleid worden met een blaadje papier.

Verder wil ik ook Lennert Jacobs en Jabran Bhatti bedanken. Reeds jarenlang zitten we in hetzelfde kantoor en behoren we tot dezelfde onderzoeksgroep. Wie had er ooit gedacht dat er zoveel tijd kon gespendeerd worden aan het onderzoeken van drieletterige afkortingen? Beide collega's (ondertussen eerder vrienden) waren dan ook vaste reisgezellen op de legendarische buitenlandse trips gedurende de voorbije jaren. Ook de lunchtijden 's middags in ons favoriete restaurant waren onvergetelijk. Daar discussieerden we

uren over muziek en wisselden tips uit over de meest essentiële obscure platen (uiteraard vinyl, wat anders?) in de meest uiteenlopende genres.

Daarnaast wil ik uiteraard ook de andere collega's en ex-collega's van DIGCOM vermelden, die elk wel hun steentje hebben bijgedragen op de een of andere manier: Henk Wymeersch, Frederik Vanhaverbeke, Nele Noels, Dieter Duyck, Julie Neckebroek, Nico Aerts, Iancu Avram, Samuel Van de Velde, Jeroen Van Hecke. Een speciale vermelding is er voor Frederik Simoens die, vooral in de beginjaren, een grote hulp was met zijn kennis en inzicht. Naast Julie, Lennert en Jabran bedank ik ook de drie andere collega's uit ons kantoor, Bart Feyaerts, Sebastian Gruenwedel en Ewout Vansteenkiste.

Verder wil ik ook Annette Nevejans, Alice Verheylesonne, Patrick Schaillée en Sylvia Moeneclaey van het secretariaat bedanken. Zij staan steeds klaar om de administratie in orde te brengen. Phillipe Serbruyns en Davy Moreels bedank ik voor de uitstekende technische ondersteuning. Uiteraard ben ik ook alle andere collega's van de vakgroep TELIN niet vergeten, maar het is wat moeilijk om iedereen persoonlijk vermelden.

Tot nu toe heb ik enkel mensen vermeld die vooral professioneel gezien iets te maken hadden met mijn doctoraat. Daarnaast wil ik uiteraard ook nog mijn familie en vrienden bedanken. Zij zorgden voor de nodige dosis ontspanning tijdens de voorbije jaren. Eén van de eerste personen die ik ooit ben tegen gekomen op deze wereld was ongetwijfeld Ellen. In het begin dacht ik: daar scheelt iets mee. Ondertussen weet ik het wel zeker, maar onder het motto 'there will be no next time' zorgt ze voor een welgekomen afwisseling van de serieuze wetenschappelijke wereld en tovert ze steevast een glimlach op mijn gezicht. Dan is er ook nog Pol, partner in crime voor legendarische skitrips (en andere). Hij combineert de stielkennis van de vakman met het artistieke van de kunstenaar, wat interessante en leerrijke conversaties oplevert. Als je iemand zoekt om fruitbomen te snoeien, dan is Pol daar ongetwijfeld de geknipte persoon voor. Verder wil ik ook Jens bedanken. Het is steeds verfrissend om van gedachten te wisselen met deze filosoof/manager uit de Lierdese underground scene (initialen JM, toeval of niet?). Verder wil ik ook Peter bedanken. Deze man heeft ongetwijfeld alles al eens gedaan (zij het misschien allemaal niet zo heel lang). Ik apprecieer vooral zijn open geest en gevoel voor humor die soms tot verrassende inzichten leiden. Ook Kirsten wil ik speciaal vermelden. Tegenwoordig zien we elkaar niet zo veel meer want ze is ondertussen een fiere moeder geworden, maar als we nog eens afspreken is het nog altijd even wijs als vroeger. Tot slot is er nog Marijke. Haar ken ik sinds we samen studeerden voor ingenieur. Het is steeds een plezier om met haar op stap te gaan. Dat leidt soms tot de meest hilarische situaties (gestolen gsms tijdens de Gentse feesten, dronken aanbidders, ...). Uiteraard zijn er nog vele vrienden die ik niet persoonlijk vermeld heb, maar daarom ben ik ze niet minder dankbaar: de spelers van de Racing, de Breininge Party Crew, vrienden uit jh alfa, ...

Mijn familie mag ik natuurlijk ook niet vergeten. Zij hebben mij steeds gesteund tijdens de voorbije jaren. Ik denk dan in de eerste plaats aan mijn moeder en mijn zus. Zonder jullie steun was het nooit gelukt. Dat geldt uiteraard ook voor de rest van de familie. Dan wil ik nog één persoon speciaal bedanken, mijn vader. Pa, zonder jou was ik nooit geraakt waar ik nu ben. Bedankt voor alles! Jammer dat je het eindresultaat van al die jaren niet meer kan zien.

Dieter Van Welden
9 december 2011

Contents

Abbreviations

AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BPSK	Binary Phase Shift Keying
CFO	Carrier Frequency Offset
CP	Cyclic Prefix
CRB	Cramer-Rao Bound
DFT	Discrete Fourier Transform
EM	Expectation-Maximization
FFT	Fast Fourier Transform
GCRB	Gaussian Cramer-Rao Bound
i.i.d.	independent and identically distributed
IBI	Inter Block Interference

ICI	Inter Carrier Interference
IDFT	Inverse Discrete Fourier Transform
IFFT	Inverse Fast Fourier Transform
ISI	Inter Symbol Interference
KSP	Known Symbol Padding
ML	Maximum Likelihood
MLSD	Maximum Likelihood Sequence Detector
MSE	Mean Squared Error
OFDM	Orthogonal Frequency Division Multiplexing
PEP	Pairwise Error Probability
PAM	Pulse Amplitude Modulation
PSK	Phase Shift Keying
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase Shift Keying
SNR	Signal-to-Noise Ratio
TDS	Time Domain Synchronous
ZF	Zero Forcing
ZP	Zero Padding

Symbols

$\mathbf{a}_{\mathrm{c}}$	Vector of $M-\nu$ frequency domain pilot symbols
$\mathbf{a}_{\mathrm{d}}$	Sequence of data symbols
$\mathbf{a}_{\mathrm{g}}$	Vector of ν time domain pilot symbols
$\angle(x)$	Argument of the complex number x in the interval $[-\pi,\ \pi[$
x^*	Complex conjugate of x
$(f \star g)(t)$	Convolution of the functions f and g at instant t
E_{b}	Average energy per information bit
E_{s}	Average symbol energy
$\mathbb{E}_X[Y]$	Expectation of Y with respect to the probability density function of X
$n!$	Factorial of the integer n: $n! = n.(n-1)\ldots 1$
$\mathbf{F}$	$N \times N$ FFT matrix

$\lfloor x \rfloor$	Floor of x resulting in the largest integer not greater than x
$\mathbf{h}$	Channel impulse response vector
$\mathbf{I}_n$	Identity matrix of order n
$\Im\{\mathbf{X}\}$	Imaginary part of $\mathbf{X}$
i	Block index
L	Length of the channel impulse response vector
$\log x$	Natural logarithm of x
$\lvert x \rvert_K$	Modulo-K reduction of x yielding a result in the interval $[0,\ K[$
M	Total number of transmitted pilot symbols per OFDM block + guard interval
m_s	Number of bits per symbol
N	Total number of carriers
N_b	Number of information bits per information word
N_c	Number of code bits per code word
N_d	Number of symbols in the transmitted symbolstream
N_0	Noise power spectral density
$\lVert \mathbf{x} \rVert$	Norm of $\mathbf{x}$
$\mathcal{O}(f(t))$	A function $g(t)$ is $\mathcal{O}(f(t))$ if and only if there exists a positive real number c and a real number t_0 such that $\lvert g(t) \rvert \leq c \lvert f(t) \rvert$, $\forall t \geq t_0$
$f(t) \propto g(t)$	The function $f(t)$ is proportional to the function $g(t)$
$\Re\{\mathbf{X}\}$	Real part of $\mathbf{X}$
ρ_c	Code rate ($\rho_\mathrm{c} = N_\mathrm{b}/N_\mathrm{c}$)
S_d	Set of data carriers
S_p	Set of pilot carriers
T	Sample period
$\mathrm{tr}(\mathbf{X})$	Trace of the square matrix $\mathbf{X}$
x	Scalar variable

$\mathbf{x}$	Column vector
$\mathbf{X}$	Matrix
$\mathbf{X}^T$	Transpose of $\mathbf{X}$
$\mathbf{X}^H$	Conjugate transpose of $\mathbf{X}$
$\mathbf{0}_{m \times n}$	$m \times n$ matrix of zeros
α_n	Index of the n-th pilot carrier
β_n	Index of the n-th data carrier
$\delta(t)$	Dirac function
δ_k	Discrete Dirac function
ν	Length of the guard interval
$\zeta_{\mathcal{C}}$	Set of legitimate code words
$\zeta_{\mathcal{D}}$	Set of legitimate data symbol sequences

Samenvatting

De voorbije jaren hebben we een gigantische toename gezien van het aantal draadloze digitale communicatiediensten. Deze nieuwe diensten zorgen voor een steeds toenemende vraag naar hogere datadebieten. Bovendien willen meerdere gebruikers terzelfdertijd informatie versturen over hetzelfde draadloze kanaal. Dit alles gecombineerd met de reeds bestaande schaarste van het beschikbare radiospectrum leidt tot het gebruik van steeds hogere draaggolffrequenties en bandbreedtes. In traditionele singlecarriermodulatietechnieken, zorgen hogere bandbreedtes voor steeds korter wordende zendpulsen. Na het verzenden van het informatiesignaal, dat bestaat uit een som van die korte zendpulsen, over een dispersief kanaal, zien we dat de verzonden pulsen wijd verspreid zijn in de tijd. Dit komt omdat de duur van de zendpulsen veel korter is dan de duur van het impulsantwoord van het kanaal. Aan de ontvangerzijde, zien we dat de ontvangen pulsen overlappen en daardoor interfereren met elkaar. De ontvanger moet dan die intersymboolinterferentie (ISI) ongedaan maken om de verstuurde informatie te kunnen reconstrueren. Typisch wordt de interferentie weggewerkt door gebruik te maken van een egalisatiefilter. Deze aanpak vraagt een hoge rekencomplexiteit, zeker bij hoge datadebieten.

Om het complexe egalisatieprobleem bij singlecarriermodulatie te vermijden, zoeken we onze toevlucht tot multicarriermodulatie. De te versturen datastroom wordt opgesplitst in verschillende parallelle datastromen, die alle een lager debiet hebben dan de oorspronkelijke datastroom. Elke datastroom wordt dan gemoduleerd op een sinusoïdale drager met een specifieke draaggolffrequentie. In dit werk beschouwen we een speciaal geval van multicarriermodulatie: alle dragers worden toegewezen aan dezelfde gebruiker. In dat geval spreken we van 'orthogonal frequency division multiplexing' (OFDM). Door het verlagen van het debiet van het verstuurde signaal, wordt de ISI sterk verminderd omdat de zendpulsen nu enkel overlappen gedurende een kleine fractie van hun totale duur. Het multipathkanaal wordt omgezet in een aantal vlakke kanalen. De verstuurde pulsen overlappen nog steeds gedurende een korte tijd, dus moet de zender een 'guard interval' voorzien tussen twee OFDM-blokken om interblokinterferentie (IBI) te vermijden. Het signaal dat uitgezonden wordt gedurende het guard interval kan niet gebruikt worden voor de detectie van de verstuurde data. Als gevolg daarvan kan de zender zelf beslissen wat er verstuurd wordt tijdens het guard interval. In dit werk beschouwen we known symbol padding (KSP) voor het vullen van het guard interval. In deze techniek wordt er een sequentie van gekende pilootsymbolen verstuurd tijdens het guard interval.

Om een betrouwbare communicatie op te zetten tussen zender en ontvanger, moet de ontvanger gesynchroniseerd zijn met de zender. Bovendien moet de ontvanger het impulsantwoord van het kanaal kennen. De lokale oscillatoren aan de zenderzijde en aan de ontvangerzijde kunnen onmogelijk een draaggolf genereren met exact dezelfde draaggolffrequentie, met als gevolg dat er steeds een draaggolffrequentieafwijking bestaat tussen zender en ontvanger. Het verstuurde signaal heeft ook een bepaalde tijd nodig om de weg af te leggen van de zender naar de ontvanger. Dit zorgt voor een tijdsvertraging tussen zender en ontvanger. Om te kunnen synchroniseren met de zender, moet de ontvanger de tijdsvertraging en de draaggolffrequentieafwijking schatten. Om dan het egalisatieproces te kunnen uitvoeren moet de ontvanger bovendien ook een schatting maken van het impulsantwoord van het kanaal.

De eerste vier hoofdstukken van dit werk vormen de basis voor de overige hoofdstukken. Het KSP-OFDM systeem dat we wensen te bestuderen, wordt geïntroduceerd in hoofdstuk twee. De verschillende stappen die nodig zijn aan de zenderzijde om een bitsequentie om te zetten in een gepaste vorm die kan verstuurd worden over het kanaal en de corresponderende stappen aan de ontvangerzijde om de verstuurde bitsequenctie te detecteren, komen aan bod. Verder worden er verschillende detectiestrategieën om de verstuurde informatie te reconstrueren uit het ontvangen signaal besproken, waarbij we veronderstellen dat de ontvanger perfect gesynchroniseerd is met de zender en dat het impulsantwoord van het kanaal gekend is. De performantie van de verschillende detectiealgoritmes wordt uitgedrukt aan de hand van de gemid-

delde bitfoutprobabiliteit (bit error rate of BER).

In een KSP-OFDM systeem worden sommige dragers enkel gebruikt om pilootsymbolen te versturen, die dan gebruikt worden bij uitvoeren van het synchronisatieproces en de kanaalschatting. Hier in dit werk, stellen we verschillende synchronisatie- en kanaalschattingsalgoritmes voor die gebruik maken van zowel de pilotsymbolen die verstuurd worden tijdens het guard interval als de pilootsymbolen op de pilootdragers. Hoofdstuk vijf behandelt het invoegen van pilootdragers in het KSP-OFDM systeem en geeft ook een overzicht van verschillende algemene strategieën om parmeterschattingsalgoritmen af te leiden.

In hoofdstukken zes tot acht worden de schattingen van het impulsantwoord van het kanaal, de tijdsvertraging en de draaggolffrequentie afwijking behandeld. Eerst komt de kanaalschatting aan de beurt. We beginnen met het berekenen van de Gaussiaanse Cramer-Rao-grens en de limiet voor respectievelijk een hoge en een lage signaal-ruisverhouding (signal-to-noise ratio of SNR). Deze performantiegrenzen geven een idee over welke gemiddelde kwadratische fout (mean squared error of MSE) een praktische kanaalschatter kan bereiken. Daarna bespreken we twee bestaande kanaalschattingsalgoritmes die kunnen toegepast worden in een KSP-OFDM systeem. De MSEs van beide kanaalschatters vertonen een horizontale asymptoot (error floor) voor matige tot hoge waarden van de SNR. Deze error floors worden veroorzaakt door de interferentie afkomstig van de onbekende datasymbolen. Wanneer alle dragers gebruikt worden voor het versturen van pilootsymbolen, dan geven beide algoritmes hetzelfde resultaat. De performantie van deze tijdsdomein (TD) 'all pilots' schatter dient als waardemeter voor de performantie van iteratieve 'decision aided' algoritmes. Wij hebben een frequentiedomein (FD) 'data aided' kanaalschatter geïntroduceerd. Het algoritme schat het impulsantwoord van het kanaal op basis van de fast Fourier transform (FFT) uitgangen van het ontvangen signaal die corresponderen met de posities van de pilootdragers. Deze schatter heeft geen last van de interferentie van de onbekende datasymbolen, omdat in het FD de datadragers perfect kunnen gescheiden worden van de pilootdragers. Er is dan ook geen error floor voor de MSE van deze schatter. Indien alle dragers als pilootdragers gebruikt worden, dan bekomen we de FD 'all pilots' schatter. De kwaliteit van de schatting van het impulsantwoord van het kanaal kan nog verder verbeterd worden door het toepassen van zogenaamde 'decision aided' schattingsalgoritmes. Dit zijn iteratieve algoritmes die itereren tussen kanaalschatting en het detecteren / decoderen van de datasymbolen. In dit werk stellen we er 3 verschillende voor. Het eerste decision aided algoritme werkt in het TD en gebruikt de harde decisies van de datasymbolen als extra pilootsymbolen. Een tweede TD algoritme is gebaseerd op het expectation maximization (EM) algoritme. In dit algoritme gebruikt de kanaalschatter zachte informatie in verband met de datasymbolen om het kanaal te schatten. Deze zachte informatie wordt

aangereikt door de detector / decoder. Het derde algoritme werkt in het FD en is een uitbreiding van FD data aided schatter. In dit algoritme worden alle carriers beschouwd om het kanaal te schatten. Naast de pilootsymbolen worden ook de harde decisies van de datasymbolen gebruikt. De numerieke resultaten tonen aan dat de TD decision aided algoritmes dezelfde performantie halen als de TD all pilots schatter voor de MSE op voorwaarde dat de SNR voldoende hoog is. Het FD decision aided algoritme werkt iets slechter en bereikt de performantie van de FD all pilots schatter. Voor ongecodeerde transmissie, wordt de BER van een ontvanger die het kanaal perfect kent, bereikt. Voor gecodeerde transmissie is er een kleine degradatie.

In hoofdstuk zeven wordt het probleem van de tijdsvertraging aangepakt. Het hoofdstuk begint met een studie van de invloed van een tijdsvertraging op de performantie van een KSP-OFDM systeem. Het blijkt dat er een interval van toelaatbare tijdsvertragingen bestaat. Een tijdsvertraging die binnen het interval ligt, veroorzaakt geen degradatie. Tijdsvertragingen buiten dit interval, veroorzaken IBI en interferentie tussen de verschillende dragers (inter carrier interferentie of ICI). Daarna bespreken we enkele bestaande algoritmes die oorspronkelijk voorgesteld werden in een time domain synchronous (TDS) OFDM systeem. Dit is een OFDM systeem zonder pilootdragers, er worden enkel pilootsymbolen verstuurd tijdens het guard interval. Na enkele kleine aanpassingen kunnen deze algoritmes ook toegepast worden in een KSP-OFDM systeem. Ze werden echter afgeleid in de veronderstelling dat de transmissie over een vlak kanaal gebeurde. In een dispersief kanaal is hun werking daarom ook minder goed. Wij hebben enkele TD data aided schattingsalgoritmes afgeleid, waarbij we een laag SNR regime veronderstellen. Deze algoritmes maken gebruik van zowel de pilootsymbolen in het guard interval als de pilootsymbolen verstuurd op de pilootdragers. De kansfunctie kan benaderd worden door de eerste twee termen van haar Taylorreeks. Voor het eerste algoritme, wordt de kansfunctie uitgemiddeld over de onbekende datasymbolen. De gezamenlijke schatting van de tijdsvertraging en het impulsantwoord van het kanaal valt uiteen in een lineaire zoekmethode voor de tijdsvertraging, waarna de schatting van het impulsantwoord van het kanaal analytisch kan berekend worden. Het tweede TD data aided algoritme is heel gelijkaardig aan het eerste, maar nu worden de datasymbolen gewoon verwaarloosd. De schatting van de tijdsvertraging wordt opnieuw gevonden door middel van een lineaire zoekmethode. Beiden vertonen een degelijke performantie. Daarna hebben we een alternatief met lage complexiteit voorgesteld. Dit algoritme is gelijkaardig aan de voorgaande maar het beschouwt enkel de pilootsymbolen in het guard interval om de tijdsvertraging te schatten. De performantie is echter niet zo goed, zeker niet voor systemen met korte guard intervals. Ten slotte wordt er een FD data aided algoritme voorgesteld. Dit algoritme schat de tijdsvertraging op basis van de FFT uitgangen van het ontvangen signaal die corresponderen met de posities van de pilootdragers. Deze

schatter heeft de beste performantie van alle tijdsvertragingsschatters die aan bod komen in dit werk.

In hoofdstuk acht wordt de schatting van de draaggolffrequentieafwijking behandeld. Een niet gecompenseerde draaggolffrequentieafwijking kan desatreuze gevolgen hebben voor de BER, want de draaggolffrequentieafwijking veroorzaakt ICI en een attenuatie van de nuttige signaalbijdrage. Opnieuw worden er eerst enkele bestaande algoritmes voor TDS-OFDM besproken. Twee algoritmes veronderstellen opnieuw een vlak kanaal, en vertonen bijgevolg een niet zo goede performantie voor een multipath fading kanaal. Een derde algoritme, het algoritme dat de draaggolffrequentie afwijking schat op basis van het guard interval, werkt wel goed en kan dus toegepast worden in een KSP-OFDM systeem. Wij hebben een TD data aided schatter voorgesteld die robust is tegen tijdsvertragingen. De performantie van deze TD data aided schatter is echter wat slechter dan de performantie van de guard interval gebaseerde schatter. Daarna hebben we de FD pilot aided schatter afgeleid. Dit algoritme schat de draaggolffrequentieafwijking op basis van de FFT uitgangen van het ontvangen signaal die corresponderen met de posities van de pilootdragers. Dit algoritme vertoon de beste performantie van alle in dit werk besproken data aided algoritmes. Maar, zelfs voor de FD pilot aided schatter is er nog steeds een error floor zichtbaar voor de MSE bij hoge SNR. Uiteindelijk hebben we ook een iteratief FD decision aided algoritme voorgesteld. Dit algoritme gebruikt in elke iteratie de harde decisies van de data symbolen zodat alle dragers in aanmerking komen om de draaggolffrequentie afwijking te schatten. De error floor is volledig verdwenen na één iteratie voor de hier beschouwde SNR waarden.

Hoofdstuk negen is het laatste hoofdstuk. Hierin herhalen we de belangrijkste verwezenlijkingen van dit werk. Bovendien geven we nog enkele suggesties voor toekomstig onderzoek.

Summary

In recent years, we have seen a huge increase in the number of wireless digital communication services. The scarcity of the radio spectrum and the demand for higher data rates combined with a lot of different users who want to transmit simultaneously over the same wireless channel, have resulted in transmission at ever increasing carrier frequencies and with ever increasing bandwidths. In single carrier modulation, higher bandwidths result in shorter durations of the transmitted pulses. After transmission of the data signal over a dispersive channel, the transmitted pulses are widely spread in time because the duration of the transmitted pulses is much smaller compared to the duration of the impulse response of the channel. At the receiver, the received pulses overlap and interfere with each other. To detect the transmitted data, the receiver needs to make use of an equalizer to remove the inter symbol interference (ISI). This approach results in prohibitively high computational complexity, especially at high data rates.

To reduce the computational requirements for the equalizer in single carrier communications, multicarrier modulation is proposed. The data stream is split up in several parallel data streams, each having a lower rate than the original data stream. Each data stream is then modulated on a sinusoidal

carrier with a different carrier frequency. In this work we consider a special case of multicarrier modulation where all the different carriers belong to one user: orthogonal frequency division multiplexing (OFDM). Lowering the rate of the transmitted signal, results in a reduction of the ISI caused by the multipath fading channel because the transmit pulses only overlap during a short fraction of their total duration. The multipath fading channel is transformed into several parallel flat fading channels. The transmitted pulses still overlap for a short time, so the transmitter needs to introduce a guard interval between two OFDM blocks to avoid inter block interference (IBI). The transmitted signal during the guard interval cannot be used for data detection, so the transmitter can choose what to transmit during the guard interval. In this work, the guard interval is filled with a known pilot sequence. This technique is called known symbol padding (KSP).

To enable a reliable communication between transmitter and receiver, it is important that the transmitter and the receiver are synchronized, and that the receiver knows the channel impulse response. The local oscillators at the transmitter and receiver side do not generate a carrier wave with exactly the same carrier frequency, which causes a carrier frequency offset. The transmitted signal needs a certain time to travel from the transmitter to the receiver, so there exists a time delay offset between transmitter and receiver. For the synchronization process, the receiver needs to estimate the time delay offset and the carrier frequency offset. Furthermore, to perform the equalization in order to detect the transmitted data, the receiver needs to estimate the channel impulse response.

The first four chapters of this work serve as the basis for the remaining chapters. The considered KSP-OFDM system is introduced. The different steps to transform a bit sequence into a suitable form that can be transmitted over the channel at the transmitter side and their corresponding steps at the receiver side are briefly discussed. Furthermore, some different approaches to recover the transmitted data sequence from the received signal are reviewed under the assumption that the receiver is perfectly synchronized with the transmitter and that the channel impulse response is known. The performance of the different data detection algorithms is expressed in terms of the bit error rate (BER). This is the probability that a bit is erroneously detected by the receiver.

For the synchronization and the channel estimation some carriers are reserved as pilot carrier in a KSP-OFDM system. In this work we propose several synchronization and channel estimation algorithms that exploit both the time domain pilot symbols, which are transmitted during the guard interval, and the pilot symbols transmitted on the pilot carriers. Chapter 5 deals with the insertion of the pilot symbols and gives an overview of some general strategies to derive parameter estimation algorithms.

Chapters 6 to 8 deal with channel estimation and the estimation of the time

delay offset and the carrier frequency offset. First we take care of the channel estimation problem for KSP-OFDM. We start with the derivation of the Gaussian Cramer-Rao bound (GCRB) and a high and low signal-to-noise ratio (SNR) limit. This lower bounds give an idea of what could be achieved by a practical channel estimator in terms of the mean squared error (MSE). Then, in the next step we review two existing channel estimators. The MSE's of both estimators show an error floor for moderate to high SNR values. These error floors are caused by the interference from the unknown data symbols. When all of the available carriers are used to transmit pilot symbols, both algorithms result in the same estimate. The time domain (TD) all pilots estimator serves as a benchmark for the performance of iterative decision aided channel estimators. We have proposed a frequency domain (FD) data aided (DA) channel estimation algorithm. This estimator estimates the channel impulse response based on the fast Fourier transform (FFT) outputs of the received signal which correspond to the pilot carrier positions. This estimator does not suffer from the interference from the unknown data symbols, because in the FD, the pilot carriers and the data carriers can be perfectly separated. When all carriers are used as pilot carriers, we obtain the FD all pilots estimator. To further improve the quality of the obtained pilot based channel estimate, we have proposed three decision directed channel estimators. They are iterative algorithms that iterate between data detection / decoding and channel estimation. A first algorithm operates in the TD and uses hard decisions on the data symbols as extra pilot symbols to obtain a new channel estimate. A second TD algorithm is based on the expectation maximization (EM) algorithm. In this approach the channel estimator uses soft information about the unknown data symbols to update the channel estimate. The soft information is provided by the detector / decoder. The third one is an extension of the FD DA channel estimator. It considers the hard decisions on the data symbols as pilot symbols. The numerical results show that the TD decision aided estimation algorithms reach the performance of the TD all pilots estimator for moderate to high SNR, while the FD decision aided algorithm reaches the performance of the FD all pilots estimator. For an uncoded transmission, the BER performance of a receiver with perfect channel knowledge is achieved. For the coded case there is a small degradation.

In chapter 7, we have derived algorithms to estimate the time delay offset. First of all, we have investigated the influence of a timing error on the performance of a KSP-OFDM system. It turns out that there is a range of tolerable time delay offsets which do not cause any degradation. Timing errors outside of this range, cause IBI and inter carrier interference (ICI). Then some existing algorithms, which are proposed for time domain synchronous (TDS) OFDM, are described. TDS-OFDM is a kind of OFDM system without any pilot carriers, there are only pilot symbols transmitted during the guard interval. With some small adaptions these algorithms can be also applied for KSP-OFDM.

However they were developed for a flat fading channel, so their performance is degraded in a dispersive channel. We have derived some TD pilot aided estimation algorithms while assuming the low SNR regime. These algorithms exploit both the pilot symbols in the guard interval as the pilot symbols transmitted on the pilot carriers. The likelihood function is approximated by the first two terms of its Taylor series expansion. In a first algorithm, the likelihood function is first averaged over the unknown data symbols. The joint estimation problem of the channel impulse response and the time delay offset boils down to a one-dimensional search for the time delay offset followed by an analytical computation of the channel impulse response estimate. The second TD pilot aided estimator ignores the data symbols. The estimate of the time delay offset is also found by a one-dimensional search. They both show a good performance but have a high computational complexity. Then, a low complexity algorithm based on the pilot symbols in the guard interval only is proposed. It does not show such a great performance especially for short guard intervals. Then finally an estimation algorithm is proposed that estimates the time delay offset based on the FFT outputs of the received signal at pilot carrier positions. It turns out, that among all the time delay offset estimators discussed in this work, the FD pilot aided estimator shows the best performance.

Chapter 8 deals with the estimation of the CFO. An uncompensated CFO causes ICI and attenuates the useful signal component, resulting in a severely BER degradation. First we have looked into existing estimators (which were proposed for TDS-OFDM). Two algorithms assume a flat fading channel and do not perform well in a multipath fading channel. A third one, the guard interval based estimator shows a good performance and can be applied in a KSP-OFDM system. We propose a TD pilot aided estimator, which is robust against a time delay offset, but it performs a bit worse than the guard interval based estimator. Then, the FD pilot aided estimator is proposed. It estimates the CFO based on the FFT outputs of the received signal corresponding to the pilot carrier positions. This estimator shows the best performance among all the pilot aided estimators that are discussed in this work, but there is still an error floor in terms of the MSE. Finally an FD decision aided estimator is proposed which considers both the pilot carriers and the data carriers for the estimation of the CFO. The algorithm uses the hard decisions on the data symbols. The error floor is completely removed after one iteration for the considered range of SNR values.

In the concluding chapter 9, the main achievements of this work are summarized. Furthermore we have formulated some suggestions for future research.

1 Introduction

1.1 Background

During the last few years, we have seen an explosion of the number of applications of wireless digital communications (mobile cell phones, tablet pcs, wireless home networks, ...). As a result, there is a demand for constantly increasing data rates, and the available radio spectrum becomes more and more saturated. Due to the scarcity of the available spectrum and the increasing demand in bandwidth, these new wireless technologies have to resort to ever increasing carrier frequencies.

In conventional single carrier modulation, to achieve large bandwidths the transmit pulses become very short. Transmission over a multipath fading channels results in lots of interference and the receiver structures become very complex because they need to apply an equalizer which has a high computational load to be able to restore the transmitted data symbols. Therefore, there was need for an alternative. Luckily this alternative has been discovered already some years ago, i.e. multicarrier modulation. This technique transforms the multipath fading channel in several parallel flat fading channels. The equalization in a flat fading channel is very easy because only a one tap

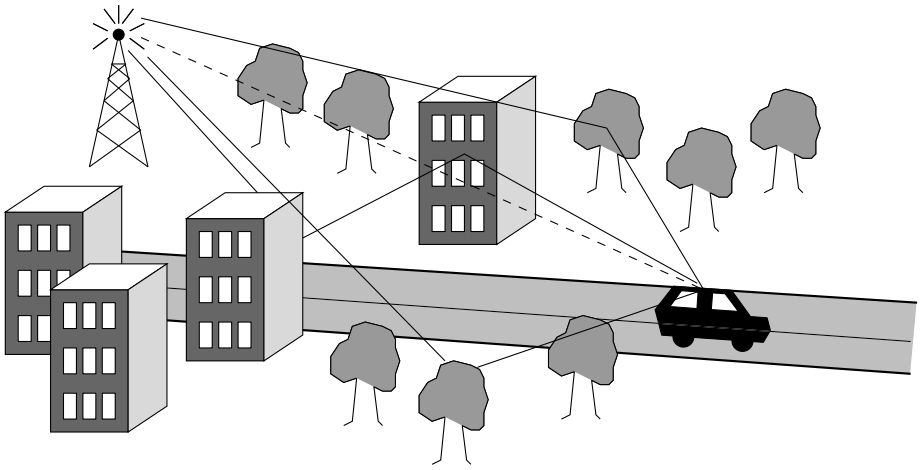

Figure 1.1: Typical scattering environment.

equalizer is necessary.

Multicarrier modulation shows attractive features to be the supporting technology for these new wireless digital communication services. As a result multicarrier modulation has been a hot research topic for some years and it still is.

1.2 Multipath Fading Channels

Radio wave propagation through wireless channels is typically characterized by multipath propagation where transmitted radio waves are reflected and / or absorbed by obstacles (buildings, cars, vegetation, ...) in the environment [1, 2]. An example of a typical scattering environment is shown in figure 1.1. The different lines between the transmitter and the receiver represent the different rays that are picked up by the receiver. When it is not blocked, the line of sight ray (dotted line) arrives first at the receiver because it follows the shortest path between transmitter and receiver. The other rays are reflected from the obstacles present in the environment and exhibit an attenuation, a phase shift and a time delay.

In practical situations, it often happens that several different paths arrive at approximately the same time and combine constructively or destructively at the receiver [3]. This is called a cluster. Each cluster is characterized by an arrival time and an attenuation. Because the multipath fading channel consists of different paths that arrive with different delays, information that is transmitted over such a channel is spread in time; the channel is called a dispersive channel.

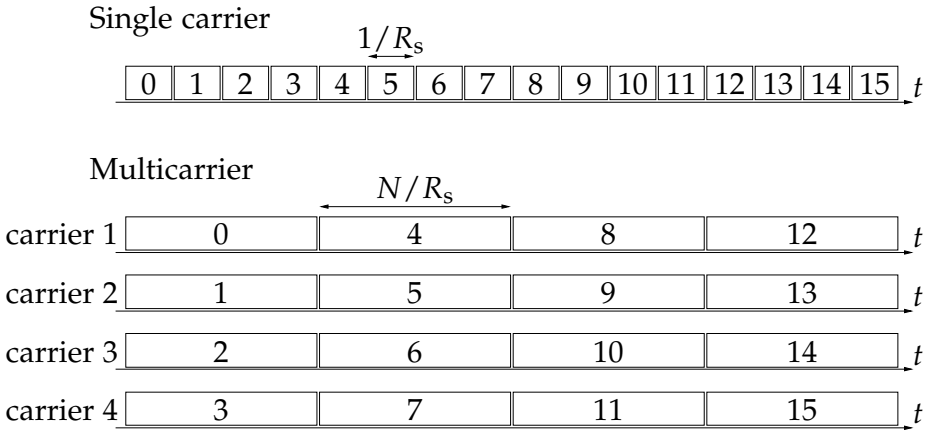

Figure 1.2: The single carrier signal and the corresponding multicarrier signal, $N = 4$.

1.3 Multicarrier Modulation

In conventional single carrier systems the symbols are transmitted at a rate R_s [1,2,4]. Each data symbol is represented by a pulse which has a duration that is proportional to $1/R_s$. The pulses are modulated on a sinusoidal carrier with fixed frequency. The resulting signal is transmitted over the dispersive channel. When the duration of the transmitted pulses is small compared to the duration of the impulse response of the channel, the transmitted pulses are widely spread in time. As a result the received pulses will overlap and interfere with each other: the channel causes ISI. Typically, the receiver makes use of an equalizer to remove the ISI. The major drawback of this approach is its high computational complexity at high data rates [2,4,5].

To tackle the equalization problem encountered in single carrier communications, multicarrier modulation was proposed [6,7]. The data stream at rate R_s is split in N parallel data streams, each having a lower rate equal to R_s/N. Each data stream is modulated on a sinusoidal carrier wave with a different carrier frequency. Figure 1.2 [4, Figure 1.1] shows the single carrier signal and the multicarrier signal (for $N = 4$ carriers) in the time domain. The number of carriers N is selected high enough so that the resulting symbol duration N/R_s is much larger than the duration of the channel impulse response. Lowering the rate R_s/N by increasing N reduces the effect of the ISI introduced by the multipath fading channel, as the transmitted pulses only overlap during a short fraction of their total duration. This is illustrated in figure 1.3 [4, Figure 1.2]. Further, as the duration of the symbols is much larger than the impulse response duration, the multipath channel is transformed into N flat fading subchannels. The data symbols on the different carriers are therefore only

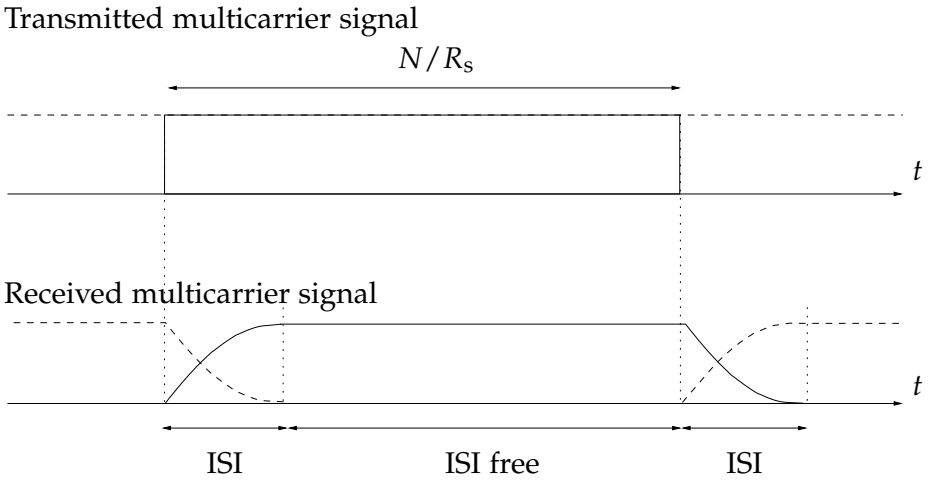

Figure 1.3: The transmitted and received multicarrier signal after transmission over a multipath fading channel.

scaled and rotated. The receiver uses only the ISI-free part of the signal for further processing, to avoid the detrimental effects of the ISI [8]. As the N data streams are all transmitted at the same time, an appropriate choice for the carrier frequencies has to be made in order to avoid interference between symbols on different carriers. It is known that harmonic sinusoids are orthogonal on a period corresponding to the fundamental frequency [7], so they are good candidates. To avoid ISI, the period of the fundamental sinusoid is chosen smaller than the duration of the ISI free zone from figure 1.3. Because each carrier exhibits flat fading, it suffices to perform at the receiver one-tap equalization at all the N carriers.

In this contribution we are considering a special case of multicarrier modulation named orthogonal frequency division multiplexing (OFDM). It is a single user modulation technique. As the name suggests, the data stream that has to be transmitted is divided in a number of substreams which are all multiplexed together in one new signal where each substream is assigned its own carrier to be transmitted on. At the receiver, the modulation of the data symbols on the different carriers can computationally efficient be performed by using the inverse fast Fourier transform (IFFT). Similarly, the receiver applies the fast Fourier transform (FFT) to demodulate the multicarrier signal.

These advantageous properties make multicarrier modulation a strong candidate for communication over multipath fading channels. In recent years, OFDM has been applied in numerous wired and wireless applications. For the transmission over twisted pair cables, the broadband digital subscriber line (DSL) standards heavily rely on OFDM: the asymmetric DSL (ADSL) and the very high-bit-rate DSL (VDSL) standards implement discrete multitone mod-

ulation (DMT), which is basically the same as OFDM [9–11]. Furthermore the broadcast of digital audio and video also use OFDM as the underlying modulation standard. See the digital audio broadcasting (DAB) standard [12] for audio transmission, and the digital video broadcasting (DVB) and digital terrestrial television broadcasting (DTTB) standards [13–16] for the implementation of digital television. Also several standards for high speed wireless local area networks (LAN) and metropolitan area networks are OFDM based, e.g. the IEEE 802.11 standard [17], WiMAX (IEEE 802.16) [18], ... Power line communications, which are systems for carrying data on a conductor also used for electric power transmission, are also heavily relying on the OFDM technique [19, 20]. Amongst others, they can be used to set up LANs. Moreover, OFDM has been proposed for LTE [21] and its successor LTE advanced [22] which are standards to increase the speed and capacity of mobile telephone networks.

1.4 Known Symbol Padding

As already mentioned in the previous section, the transmitted pulses corresponding to the OFDM blocks slightly overlap in the time domain (see figure 1.3). So only the middle part, that is free of ISI, can be used to detect data. This also means that we can choose what to transmit during the time interval between two of such ISI-free parts. This interval is called the guard interval.

In the literature, the most popular technique to fill the guard interval is called the cyclic prefix (CP) technique [7, 23]. The last samples of each OFDM block are copied and transmitted before the actual OFDM block. The main advantage of the CP is that for the samples in the ISI-free part it transforms the linear convolution of the transmitted signal and the channel impulse response into a circular convolution. This results in low complexity data detection and equalization algorithms. A drawback of this approach is the loss of the multipath diversity: if one of the N subchannels exhibits a deep fade, all the information transmitted over that subchannel might be lost [24].

A second technique is called zero padding (ZP). In this case, each two OFDM blocks are separated by an empty guard interval [23, 24]. In contrast with CP-OFDM, ZP-OFDM guarantees symbol recovery, regardless of the presence of channel fades [25]. Furthermore, a received ZP-OFDM signal offers the possibility to be transformed into a CP-OFDM signal, which means it can benefit from the low complexity equalization algorithms for CP-OFDM systems. We can conclude that ZP offers a higher flexibility than CP: equalization complexity can be traded off for symbol detectability [25].

Both the CP and the ZP technique do not offer much freedom in what to transmit during the guard interval. As a result, the guard interval can hardly be used for channel estimation and is sometimes insufficient for synchronization purposes [26]. To have more control about the content of the

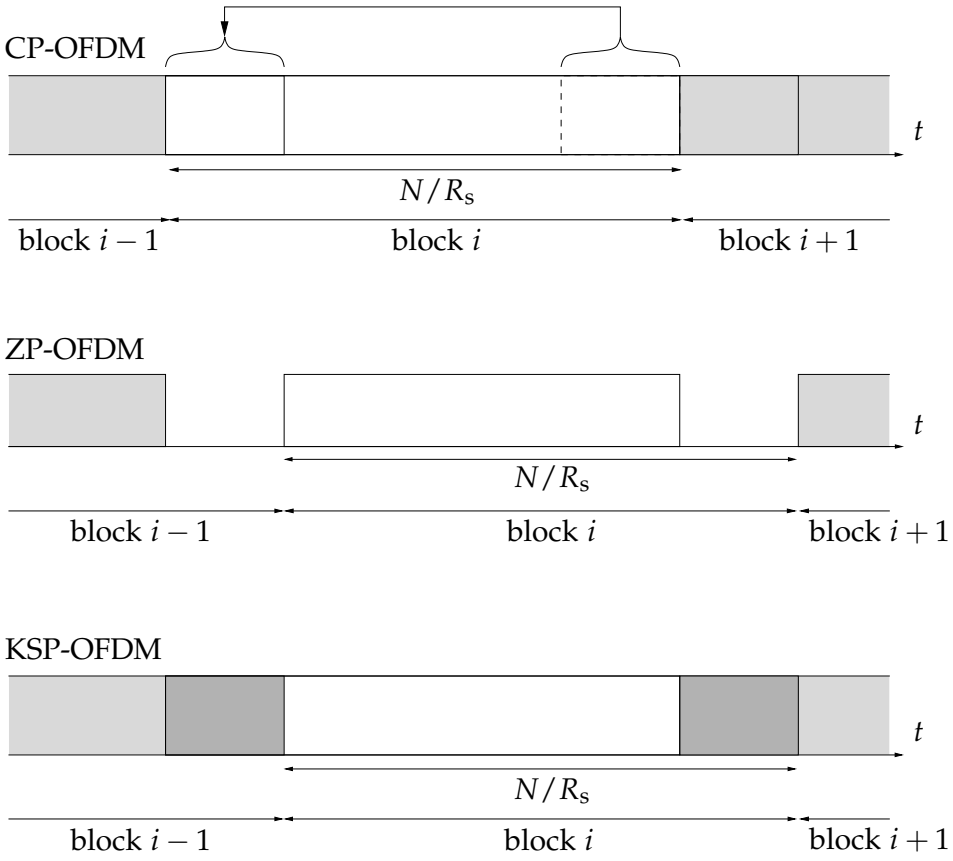

Figure 1.4: The structure of the transmitted signal for CP-OFDM, ZP-OFDM and KSP-OFDM respectively.

guard interval, recently a third guard interval technique was proposed: the known symbol padding (KSP) technique [23, 26, 27]. The guard interval is now filled with a sequence of known symbols, which can be optimized to have beneficial properties for synchronization and/or channel estimation. In the literature, there is already an extensive knowledge about sequence design for synchronization purposes available (pseudo noise sequences, ...), which can be exploited to optimize the training sequence in the guard interval. A KSP-OFDM system can be seen as a ZP-OFDM system where an extra training sequence is transmitted during the guard interval. As a result, KSP-OFDM exhibits the same (beneficial) properties in terms of symbol detectability as a ZP-OFDM. In this work we will mainly focus on KSP-OFDM.

Figure 1.4 illustrates the three discussed guard interval techniques.

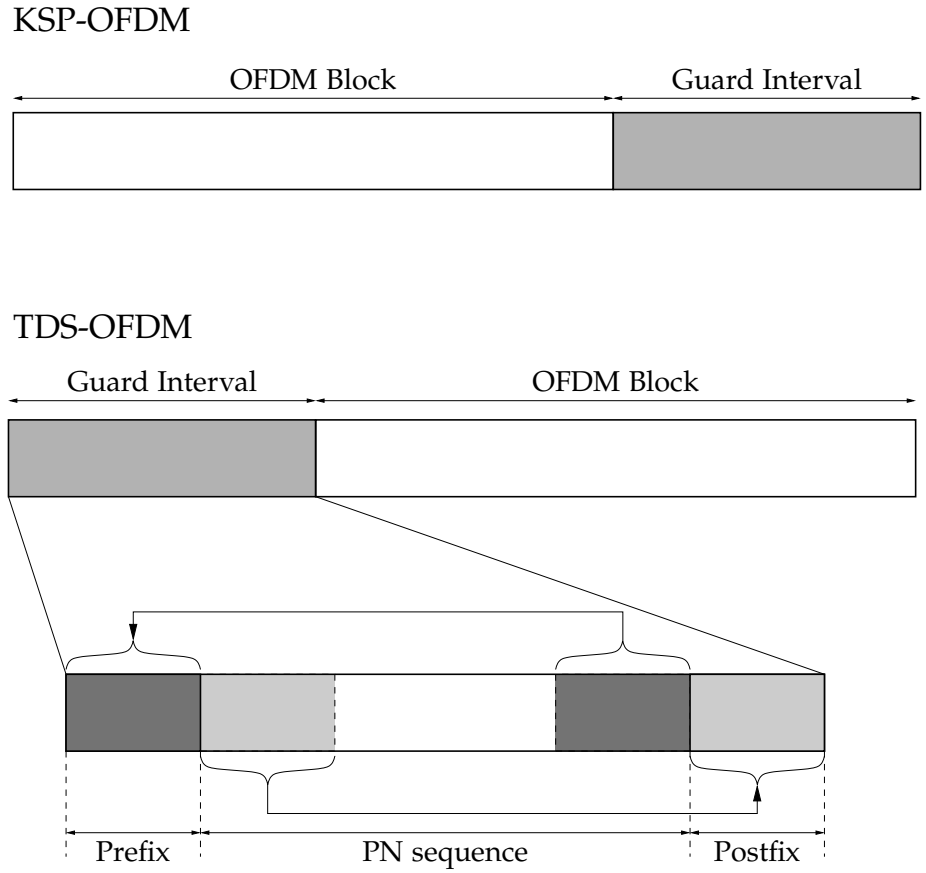

Figure 1.5: The structure of the transmitted signal for KSP-OFDM and TDS-OFDM.

1.4.1 Time Domain Synchronous OFDM

Time domain synchronous (TDS) OFDM is a technique that is very similar to KSP-OFDM and that is extensively used in the Chinese digital television standard [16]. The guard interval also consists of a known pilot sequence. The structure of one transmitted block is shown in 1.5.

TDS-OFDM is similar to KSP-OFDM but not identical. There are some differences. First of all the structure of one transmitted block is different: in a KSP-OFDM system a transmitted block consists of the IFFT outputs followed by the known time domain pilot sequence, while in a TDS-OFDM system the known time domain pilot sequence is transmitted first and then followed by the IFFT outputs.

Secondly, the known time domain pilot sequence exhibits a special structure in a TDS-OFDM system (see figure 1.5). It consists of a prefix, a main part

and a postfix. A pseudo noise sequence of length N_{PN} forms the main part. The last N_{pre} samples of this pseudo noise sequence are copied and placed before the actual pseudo noise sequence and they are the prefix. Same goes for the postfix: the first N_{post} samples of the pseudo noise sequence are copied and appended at the end of the pseudo noise sequence. As a result, we obtain a guard interval consisting of $N_{\mathrm{pre}} + N_{\mathrm{PN}} + N_{\mathrm{post}}$ samples.

Finally, the last difference is whether there are pilot carriers available or not. In a general OFDM system some carriers are used to transmit additional pilot symbols in the frequency domain, as we will see later on. They are used for synchronization and channel estimation. This is also the case in a KSP-OFDM system, but not in a TDS-OFDM system.

1.5 Motivation

Two important aspects in setting up a reliable communication between transmitter and receiver involve the synchronization of the receiver with the transmitter and the channel estimation. In the literature, there are hardly any publications which propose synchronization algorithms or channel estimation algorithms for a KSP-OFDM system. Therefore, we investigate and propose several algorithms for the estimation of the frequency offset and the time delay in this work . Furthermore we develop some algorithms to obtain an estimate of the channel impulse response. With the rise of the very powerful error correcting codes, iterative processing became very important in signal processing to lower the complexity of the complicated decoding process, while still achieving nearly optimum performance. The decoding has been combined with iterative parameter estimation algorithms, exhibiting very good results even in high noise environments. So a part of this work consists of finding iterative synchronization and channel estimation algorithms that can be combined with one of the powerful error correction schemes.

1.6 Outline

The remaining part of this work is organized as follows. In the next chapter, a typical KSP-OFDM system is introduced. The building blocks of the transmitter and the receiver are discussed. Furthermore, a signal model is developed, which is then used in the remaining part of this work.

In chapter 3, some basic principles from the estimation and detection theory that are useful for digital communications are reviewed. They form the basis for the derivation of algorithms to perform the estimation of the unknown synchronization parameters and the detection of the transmitted data.

In chapter 4, the detection of the data symbols for a KSP-OFDM system is treated. Several approaches are reviewed. Two low complexity methods are

compared with the optimal, but computationally intractable detection algorithm. Both uncoded and coded transmission is considered.

The chapters 5-8 treat the estimation of synchronization parameters and the channel impulse. In chapter 5, a general introduction to the different types of estimation strategies is given. One of the important aspects is the insertion of pilot symbols in the transmitted OFDM symbol. Chapter 6 deals with channel estimation. First of all, some theoretical lower bounds on the achievable performance of practical channel estimators are derived. Then some existing algorithms are studied. In the last part of the chapter, we propose several estimation algorithms starting from a rather simple pilot aided algorithm and then moving on to more advanced algorithms. The chapter contains some numerical results to illustrate the achievable performance of the different proposed algorithms. Chapter 7 treats the problem of time delay estimation. First the effect of a possible time delay offset is investigated on the performance of a KSP-OFDM system and the required accuracy of practical estimation algorithms is derived. Then some existing algorithms that are proposed in a TDS-OFDM context are reviewed. In the remaining part different algorithms are proposed and compared with each other. Their performance is extensively illustrated by means of some numerical results. Chapter 8 deals with the estimation of unknown frequency offsets. Some existing methods are discussed first and then some new approaches are proposed. The here in this work developed frequency offset estimation algorithms range from very simple pilot aided estimators to more advanced iterative frequency offset estimation algorithms. At the end of the chapter some numerical results are shown to give an idea of the achievable performance of the proposed frequency offset estimation algorithms.

Finally chapter 9 sums up the most important achievements of this work. Furthermore, the chapter provides some ideas for future research.

2 System Description

In this chapter we will give a thorough description of a typical KSP-OFDM communication system. The structure is similar to a standard multicarrier system, which can be found in [2]. Figure 2.1 shows the three main parts

- the transmitter, described in section 2.1
- the channel, described in section 2.2
- the receiver, described in section 2.3

In the last part of this chapter (see section 2.4), we introduce the time delay offset and the carrier frequency offset which are besides the channel impulse response the synchronization parameters that have to be estimated by the receiver before the data detection can be performed.

2.1 Transmitter

The main task of the transmitter consists of transforming the information bit stream into a physical wave form, which can be transmitted over the physical channel. We consider a bandpass communication system, which means

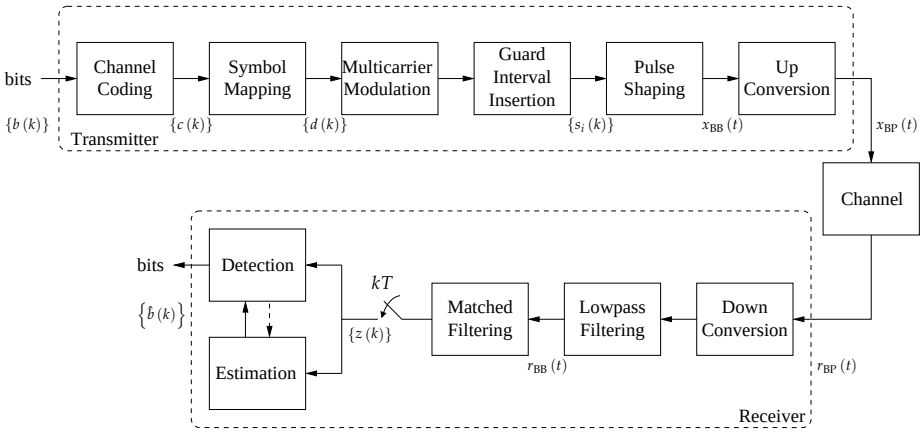

Figure 2.1: Building blocks of a KSP-OFDM system.

that the transmitter is only allowed to transmit its information in a certain frequency band of $2B$ around a carrier frequency f_c. The process of transforming the bit stream in to a bandpass signal $x_{BP}(t)$ comprises several steps:

- channel coding
- symbol mapping
- multicarrier modulation
- guard interval insertion
- pulse shaping
- up conversion

which will be described in the following subsections.

2.1.1 Channel Coding

To be able to set up reliable communication in sometimes severe conditions, the information bit stream is first sent through a channel encoder. The bit stream is made more robust against bit errors by adding some redundant information in a systematic way. The information bit stream is divided in groups of N_b bits, which are called the information words. Each information word $\mathbf{b}$ is then encoded into a new bit sequence of N_c bits ($N_c > N_b$), which is called the code word $\mathbf{c}$, according to an encoding rule $\mathbf{c} = \mathcal{C}(\mathbf{b})$. When the information bits themselves are included in the coded bits, one speaks of a *systematic* code. In this case, the code word can be separated in the N_b information bits and the $N_c - N_b$ parity-check bits. When the code word does

not contain the original information bits, the code is called *non systematic.* The ratio $\rho_c = N_b/N_c$ is called the code rate and gives an idea of the amount of redundancy that is introduced by the code: the lower ρ_c, the higher the amount of redundancy. The set of legitimate code words is denoted as $\zeta_{\mathcal{C}}$.

An important class of error correcting codes are the *linear* codes. A code is linear when the sum of two code words is another code word. The code words of a linear code are obtained by modulo-2 multiplication of the information words with a binary generator matrix $\mathbf{G}$:

$$\mathbf{c} = \mathbf{G}\mathbf{b}. \tag{2.1}$$

The matrix $\mathbf{G}$ defines the properties of the code. Well known examples of linear codes are convolutional codes, turbo codes and low-density parity-check codes [28–30].

2.1.2 Symbol Mapping

The second step of the transmission process is the conversion of the coded bitstream $\mathbf{c}$, consisting of N_c bits, into a sequence of N_d complex-valued data symbols $\mathbf{a} = (a(0), a(1), \ldots)$ (see also [31, subsection 2.2.2]). First the coded bits are grouped in $N_d = N_c/m_s$ groups of m_s bits; the k-th group is defined as $\mathbf{c}_k = (c_k(0), \ldots, c_k(m_s - 1))$. Then each group of bits $\mathbf{c}_k$ is mapped to one of the $M_s = 2^{m_s}$ complex numbers of the constellation $\Omega_{M_s} = \{\omega_0, \ldots, \omega_{M_s-1}\}$. We call M_s the order of the constellation. In this work, only normalized constellations are considered: $\frac{1}{M_s}\sum_{i=0}^{M_s-1}|\omega_i|^2 = 1$. The mapping function is denoted as:

$$\mathcal{M}_s : \{0,1\}^{m_s} \to \Omega_{M_s} : \; a(k) = \mathcal{M}_s(\mathbf{c}_k).$$

The mapping of the code word $\mathbf{c}$ on to the sequence of data symbols $\mathbf{a}$ will be denoted as $\mathcal{M}$. The set of legitimate data symbol sequences is denoted as $\zeta_{\mathcal{D}}$. A data symbol sequence $\mathbf{a}$ is legitimate if and only if there exists an information word $\mathbf{b}$ for which $\mathbf{a} = \mathcal{M}(\mathcal{C}(\mathbf{b}))$. Further we assume that the first order and second order moments of the coded data symbols are equal to the corresponding moments of uncoded data symbols. For uncoded transmission, the symbol mapping is expressed as: $\mathbf{a} = \mathcal{M}(\mathbf{b})$, because the coded bits $\mathbf{c}$ are equal to the information bits $\mathbf{b}$. As a result, the first order and second order moments of the data symbols are equal to:

$$\begin{aligned} \mathbb{E}[a(k)] &= 0 \\ \mathbb{E}\left[a(k_1)^* a(k_2)\right] &= \delta_{k_1-k_2} \\ \mathbb{E}[a(k_1)\, a(k_2)] &= 0 \end{aligned} \tag{2.2}$$

In practice, the most frequently used constellations are

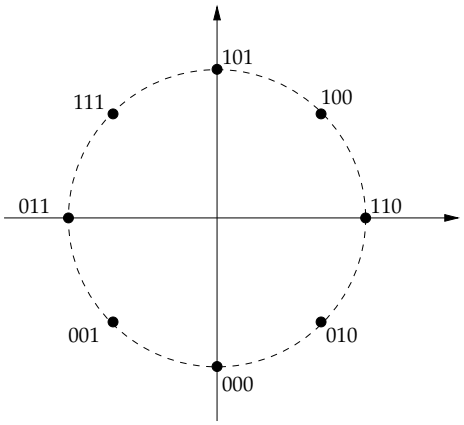

Figure 2.2: 8-PSK constellation with Gray Mapping.

- M_s-ary Pulse Amplitude Modulation (M_s-PAM) with

$$\Omega_{M_s} = \sqrt{\frac{3}{M_s^2 - 1}} \{\pm 1,\ \pm 3,\ \ldots,\ \pm (M_s - 1)\}.$$

- M_s-ary Quadrature Amplitude Modulation (M_s-QAM) with

$$\Omega_{M_s} = \left\{\omega : \Re\{\omega\}, \Im\{\omega\} \in \sqrt{M_s}\text{-PAM}\right\}.$$

  The real part of ω, denoted as $\Re(\omega)$, and the imaginary part of ω, denoted as $\Im\{\omega\}$ are both members of a $\sqrt{M_s}$-PAM constellation. Note that we only consider rectangular QAM constellations, which means that M_s has to be a power of 4.

- M_s-ary Phase Shift Keying (M_s-PSK)

$$\Omega_{M_s} = \left\{e^{j2\pi\frac{k}{M_s}} :\ k = 0,\ \ldots,\ M_s - 1\right\}.$$

The mapping of the M_s possible sequences of m information bits to the M_s possible constellation points, can be done in $M_s!$ different ways, where $n!$ is the factorial of the non negative integer n, i.e. $n! = n \cdot (n-1) \cdot \ldots 1$. The choice of the mapping function affects the performance of the communication system [2]. For uncoded transmission, the preferred mapping function is the one in which adjacent constellation points only differ in one bit. This is called *Gray Mapping*. During the demodulation process of the signal, the most likely errors caused by noise will result in the selection of an adjacent constellation point. When we consider Gray Mapping, selecting erroneously an adjacent constellation point results in only a single bit error in the sequence of m bits.

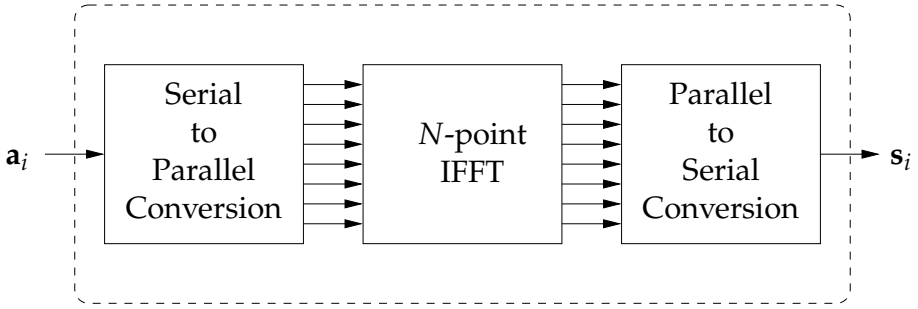

Figure 2.3: Multicarrier modulation.

Hence, of all mapping functions, Gray mapping will minimize the probability of an erroneous bit. Figure 2.2 shows an example of Gray Mapping for the 8-PSK constellation.

After the mapping of the bits to the complex symbols, the resulting sequence of complex numbers is multiplied with $\sqrt{E_s}$, where E_s is the average symbol energy.

2.1.3 Multicarrier Modulation

The multicarrier modulation building block (see also [2]) from figure 2.1 is depicted more in detail in figure 2.3. The available bandwidth is divided into N subchannels. We call the parameter N the number of carriers of the OFDM system. The different subchannels are made orthogonal by selecting the carrier spacing equal to the per carrier symbol rate [4].

The stream of symbols to be transmitted is divided in blocks of N symbols $\mathbf{a}_i = (a_i(0), \ldots, a_i(N-1))^T$, where i denotes the block index. The N time-domain samples $\mathbf{s}_i = (s_i(0), \ldots, s_i(N-1))^T$ to be transmitted during the i-th block are easily generated by applying the N symbols of block $\mathbf{a}_i$ to the IFFT. The samples of $\mathbf{s}_i$ are defined as

$$s_i(k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} a_i(n)\, e^{j2\pi \frac{kn}{N}},\ k = 0, \ldots, N-1. \tag{2.3}$$

This can also be formulated using the matrix notation:

$$\mathbf{s}_i = \mathbf{F}^H \mathbf{a}_i, \tag{2.4}$$

where $\mathbf{F}$ is the $N \times N$ FFT matrix, which is defined as

$$(\mathbf{F})_{k,n} = \frac{1}{\sqrt{N}} e^{-j2\pi \frac{kn}{N}}. \tag{2.5}$$

Note that $\mathbf{F}$ is a unitary matrix: $\mathbf{F}\mathbf{F}^H = \mathbf{F}^H\mathbf{F} = \mathbf{I}_N$ where $\mathbf{I}_N$ is the $N \times N$ identity matrix.

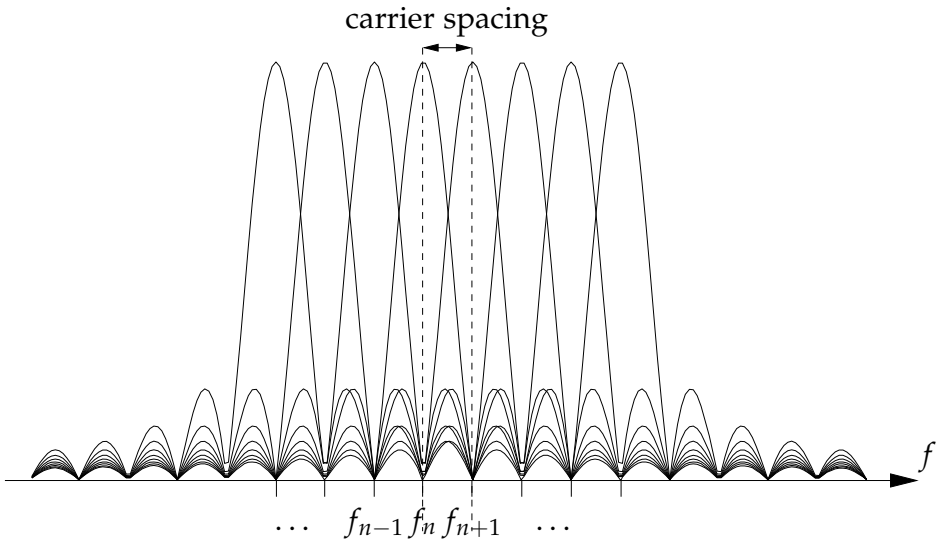

Figure 2.4: Spectrum of a part of the OFDM carriers.

As we already know from chapter 1, we want to achieve a rate equal to R_s, which means that we have a time frame of duration N/R_s to transmit each OFDM block. A part of this time frame is reserved for the guard interval, which is inserted between each two OFDM blocks to avoid interference between them. We divide the available time frame in $N+\nu$ intervals of equal duration T: N intervals to transmit the samples $s_i(k)$ and ν samples for the guard interval. We call T the sample period, which is given by

$$T = \frac{1}{N+\nu}\frac{N}{R_s}. \tag{2.6}$$

The rate per carrier is equal to $1/NT$, which means that the carrier spacing also has to equal $1/NT$ to guarantee orthogonality [4,7]. The different carriers are located at the frequencies $f_n = n/(NT)$, $n = 0, \ldots, N-1$. Figure 2.4 shows a part of the spectrum of the OFDM signal. It is given by a sum of frequency shifted sinc functions, that are spaced by the carrier spacing $1/(NT)$. We see that the different carriers spectrally overlap, but, as at frequency f_n, the contributions from other carriers with $f_{n'} \neq f_n$ are zero, the carriers do not interfere at the carrier frequencies f_n.

2.1.4 Guard Interval Insertion

At the end of each OFDM block $\mathbf{s}_i$, a guard interval consisting of ν samples is appended [23]. Those samples are denoted as $\mathbf{a}_g = \left(a_g(0), \ldots, a_g(\nu-1)\right)^T$ and consist of known pilot symbols which have the same energy per symbol

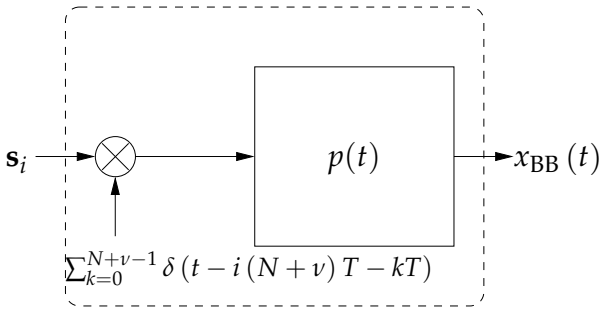

Figure 2.5: Pulse Shaping.

as the symbols modulated on the carriers:

$$\mathbb{E}\left[\left|a_{\mathrm{g}}(l)\right|^2\right] = \mathbb{E}\left[\left|a_i(n)\right|^2\right] = E_{\mathrm{s}}. \tag{2.7}$$

As already mentioned in chapter 1, the purpose of the guard interval is twofold in a KSP-OFDM system [32–35]. First of all, it is inserted between two OFDM blocks to avoid inter block interference. Secondly, the actual content of the guard interval, which is a known training sequence, can be exploited for synchronization tasks and channel estimation.

The energy of the total block $\mathbf{s}_i$ is normalized so that the energy transmitted per OFDM block is equal to NE_{s} which yields for the $N+\nu$ time domain samples $\mathbf{s}_i = (s_i(0), \ldots, s_i(N+\nu-1))^T$

$$\mathbf{s}_i = \sqrt{\frac{N}{N+\nu}} \begin{pmatrix} \mathbf{F}^H \mathbf{a}_i \\ \mathbf{a}_{\mathrm{g}} \end{pmatrix}. \tag{2.8}$$

The average symbol energy E_{s} can be related to the average energy per information bit E_{b} by

$$E_{\mathrm{b}} = \frac{\text{total number of transmitted OFDM blocks}}{\text{total number of transmitted information bits}} NE_{\mathrm{s}}. \tag{2.9}$$

In the special case that one information word of N_{b} bits is transmitted in one OFDM block E_{b} becomes

$$E_{\mathrm{b}} = \frac{NE_{\mathrm{s}}}{N_{\mathrm{b}}} = \frac{NE_{\mathrm{s}}}{\rho_{\mathrm{c}} m_{\mathrm{s}} N_{\mathrm{d}}}. \tag{2.10}$$

2.1.5 Pulse Shaping

The discrete time sequence $\mathbf{s}_i$ (2.8) has to be transformed into a continuous time baseband signal $x_{\mathrm{BB}}(t)$ to be able to be transmitted. This process is called *pulse shaping* and is shown in figure 2.5. The sequence of samples $\mathbf{s}_i$

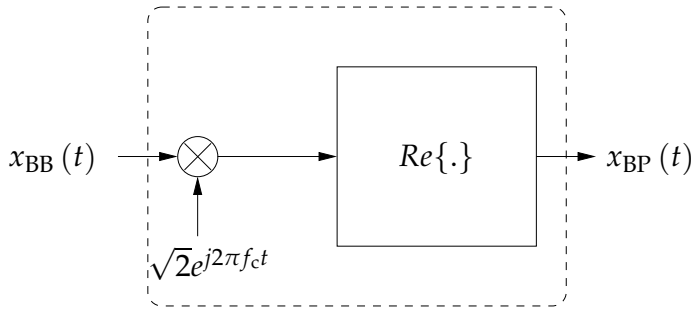

Figure 2.6: Up conversion.

is multiplied with the transmit clock signal which consists of a sequence of Dirac impulses occurring at time instants $i(N+\nu)T + kT$. The parameter T is called the *sample period*. The resulting sequence of weighted Dirac impulses is then applied to the transmit filter with transfer function $P(f)$ and impulse response $p(t)$. This filter is band limited and has a one sided bandwidth B: $P(f) = 0$ when $|f| > B$. The transmit pulse is considered to have unit energy:

$$\int_{-\infty}^{\infty} |p(t)|^2 \, dt = \int_{-B}^{B} |P(f)|^2 \, df = 1. \tag{2.11}$$

The resulting base band signal $x_{BB}(t)$ is then given by

$$x_{BB}(t) = \sum_{i} \sum_{k=0}^{N+\nu-1} s_i(k) \, p(t - i(N+\nu)T - kT). \tag{2.12}$$

In this work the considered transmit pulse is a square root raised cosine pulse with respect to the symbol period T which means that

$$\int_{-\infty}^{\infty} p(t - kT) \, p^*(t) \, dt = \delta_k. \tag{2.13}$$

2.1.6 Up Conversion

Every communication system is only allowed to transmit in a frequency band centered around a certain carrier frequency f_c. So the baseband signal has to be upconverted to the allocated frequency band. This operation is performed by multiplying the baseband signal $x_{BB}(t)$ with the sinusoidal $e^{j2\pi f_c t}$. This yields for the bandpass signal $x_{BP}(t)$ (see also figure 2.6)

$$x_{BP}(t) = \sqrt{2}\Re\left\{x_{BB}(t) \, e^{j2\pi f_c t}\right\}. \tag{2.14}$$

The presence of the factor $\sqrt{2}$ is necessary to ensure that the power of the bandpass signal $x_{BP}(t)$ is equal to the power of the baseband signal $x_{BB}(t)$.

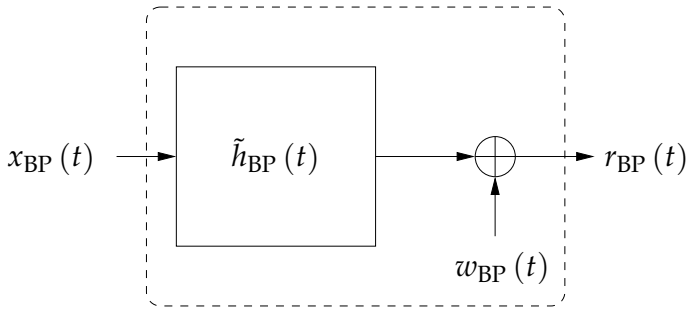

Figure 2.7: Channel.

This guarantees that the transmitted energy per symbol does not change after the up conversion. Note that the bandpass signal $x_{\mathrm{BP}}(t)$ is a real-valued signal while the baseband signal $x_{BB}(t)$ is complex-valued. The bandpass signal is then transmitted over the channel.

2.2 Channel

We consider a multipath fading channel with $\tilde{L}$ paths. Each path l is characterized by a real valued path gain γ_l and a path delay τ_l. The bandpass impulse response $\tilde{h}_{\mathrm{BP}}(t)$ is given by[1]

$$\tilde{h}_{\mathrm{BP}}(t) = \sum_{l=0}^{\tilde{L}-1} \gamma_l \delta(t - \tau_l). \tag{2.15}$$

The channel also adds real-valued additive white Gaussian noise (AWGN), denoted as $w_{\mathrm{BP}}(t)$ with zero mean and power spectral density equal to $N_0/2$ with double sided bandwidth $2B$ around the carrier frequency f_{c}. The total bandpass signal $r_{\mathrm{BP}}(t)$ that arrives at the receiver is given by

$$r_{\mathrm{BP}}(t) = \sum_{l=0}^{\tilde{L}-1} \gamma_l x_{\mathrm{BP}}(t - \tau_l) + w_{\mathrm{BP}}(t). \tag{2.16}$$

We can express the bandpass signal $r_{BP}(t)$ as a function of an equivalent base band signal $r_{\mathrm{BB}}(t)$:

$$r_{\mathrm{BP}}(t) = \sqrt{2}\Re\left\{ r_{\mathrm{BB}}(t)\, e^{j2\pi f_c t} \right\}. \tag{2.17}$$

[1]Note that in general, the parameters $\tilde{L}$, $\{\gamma_l\}$ and $\{\tau_l\}$ depend on t, but here we assume block fading which means that the channel parameters do not change much over several OFDM symbol durations.

When we substitute $x_{BP}(t)$ by (2.14) and $r_{\mathrm{BP}}(t)$ by (2.17) in (2.16), we find that the baseband representation of (2.16) is given by

$$r_{\mathrm{BB}}(t) = \sum_{l=0}^{\tilde{L}-1} \gamma_l e^{-j2\pi f_c \tau_l} x_{\mathrm{BB}}(t-\tau_l) + w_{\mathrm{BB}}(t) \tag{2.18}$$

where $w_{\mathrm{BB}}(t)$ is complex-valued AWGN with independent real and imaginary parts each a having a zero mean and a power spectral density equal to $N_0/2$ in the frequency band $[-B,\ B]$. The base band representation of $\tilde{h}_{\mathrm{BP}}(t)$ can be expressed as

$$\tilde{h}(t) = \sum_{l=0}^{\tilde{L}-1} \gamma_l e^{-j2\pi f_c \tau_l} \delta(t-\tau_l). \tag{2.19}$$

Note that $\tilde{h}(t)$ is in fact an infinite bandwidth signal, but since $x_{\mathrm{BB}}(t)$ is limited to the frequency band $[-B,\ B]$, the convolution $\left(\tilde{h} \star x_{\mathrm{BB}}\right)(t)$ is also bandlimited and yields the same result as the convolution $\left(\tilde{h}_{\mathrm{BB}} \star x_{\mathrm{BB}}\right)(t)$ with $\tilde{h}_{\mathrm{BB}}(t)$ defined as

$$\tilde{h}_{\mathrm{BB}}(t) = \sum_{l=0}^{\tilde{L}-1} \gamma_l e^{-j2\pi f_c \tau_l} 2B\mathrm{sinc}\left(2B(t-\tau_l)\right). \tag{2.20}$$

Defining $\tilde{h}_l$ as the complex-valued path gain of the l-th channel path:

$$\tilde{h}_l = \gamma_l e^{-j2\pi f_c \tau_l}, \tag{2.21}$$

we can express the channel impulse response $\tilde{h}(t)$ as

$$\tilde{h}(t) = \sum_{l=0}^{\tilde{L}-1} \tilde{h}_l \delta(t-\tau_l) \tag{2.22}$$

We assume that the different paths are generated by different scatterers, so they have negligible correlations [3]. As a result, the amplitudes $\left\{\tilde{h}_l\right\}_{l=0}^{\tilde{L}-1}$ can be considered as statistically independent random variables. Furthermore it is assumed that there is no line of sight path. In that case, the Rayleigh fading model is frequently used to describe the multipath fading, which means that the real and imaginary parts of each complex amplitude $\tilde{h}_l$ are modeled by independent and identically distributed zero-mean Gaussian variables [2,3].

2.3 Receiver

The most important purpose of the receiver is the reconstruction of the transmitted bits from the received signal. To achieve this goal, several steps have to be completed:

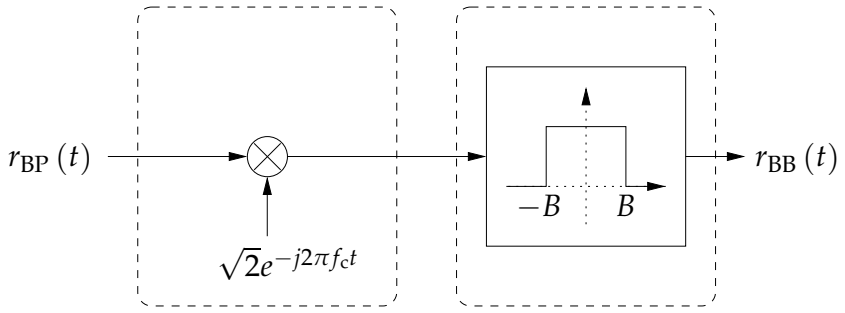

Figure 2.8: Down Conversion and Lowpass Filtering.

- down conversion
- lowpass filtering
- matched filtering
- sampling
- parameter estimation
- data detection

Most of these steps are the reverse operations of the ones performed at the transmitter side. Only the *parameter estimation* step has no counterpart at the transmitter side. During this step, the channel and synchronization parameters (e.g. time delay offset, frequency offset and channel impulse response) are estimated.

2.3.1 Down Conversion and Lowpass Filtering

The first steps consist of transforming the received bandpass signal $r_{\mathrm{BP}}(t)$ to a base band signal $r_{\mathrm{BB}}(t)$ as shown in figure 2.8. The bandpass signal $r_{\mathrm{BP}}(t)$ is first multiplied by the sinusoidal $e^{-j2\pi f_c t}$, which is in fact the inverse operation of the up conversion which is performed in the transmitter. The spectrum of $r_{\mathrm{BP}}(t)$ is shifted to the left by this operation and has a contribution around 0 Hz and $-2f_c$. Sending the resulting signal through an ideal low pass filter with single sided bandwidth equal to B removes the high frequency components around $-2f_c$ and results in the desired base band signal $r_{\mathrm{BB}}(t)$, which is defined as (2.18). The extra factor $\sqrt{2}$ guarantees that the energy of the useful signal in $r_{\mathrm{BB}}(t)$ remains unaltered.

2.3.2 Matched Filtering and Sampling

The base band signal $r_{\mathrm{BB}}(t)$ is sent through a filter with an impulse response matched to the impulse response $p(t)$ of the pulse shaping filter from the transmitter. The impulse response of the matched filter is given by $p^*(-t)$. Then the resulting signal $z(t)$ is sampled at a rate $1/T$, resulting in the samples $z(k)$ which are given by

$$z(k) = \int_{-\infty}^{+\infty} r_{\mathrm{BB}}(t)\, p(t - kT)\, dt. \tag{2.23}$$

Using (2.18) and (2.12), we can rewrite (2.23) as

$$z(k) = \sum_{i=-\infty}^{+\infty} \sum_{l=0}^{\tilde{L}-1} \sum_{k'=0}^{N+\nu-1} \alpha_l s_i\left(k'\right) g\left(\left(k-k'\right) T - i\left(N+\nu\right) T - \tau_l\right) + w(k), \tag{2.24}$$

where $g(t) = \int_{-\infty}^{+\infty} p(t+u)\, p^*(u)\, du$ and $w(k) = \int_{-\infty}^{+\infty} w_{\mathrm{BB}}(t)\, p(t - kT)\, dt$. We define $h(t)$ as the overall pulse shape: $h(t)$ consists of the cascade of the transmit pulse $p(t)$, the channel impulse response $\tilde{h}(t)$ and the impulse response of the matched filter $p^*(-t)$. This yields for $h(t)$:

$$h(t) = \sum_{l=0}^{\tilde{L}-1} \tilde{h}_l g(t - \tau_l). \tag{2.25}$$

Rewriting (2.24) using (2.25) results in

$$z(k) = \sum_{i=-\infty}^{+\infty} \sum_{k'=0}^{N+\nu-1} s_i\left(k'\right) h\left(\left(k-k'\right) T - i\left(N+\nu\right) T\right) + w(k). \tag{2.26}$$

We assume that $h(t)$ has only a limited duration: $h(t) = 0$ for $t < 0$ and $t \geq LT$. The vector $\mathbf{h} = \left(h(0), \ldots, h\left(\left(L-1\right) T\right)\right)^T$ collects the samples of the impulse response.

2.3.3 Estimation and Detection

After the analog to digital conversion of the received signal, the samples are fed to the estimation unit and the detection/decoding unit. The estimation task of the receiver consists of providing estimates of unknown synchronization parameters (which are introduced in section 2.4) and the channel coefficients to the detection/decoding unit. The detector/decoder will then use those estimates to try to recover the transmitted information bits. Some possible estimation algorithms will be treated in chapters 6-8.

There exist several possible detection strategies for KSP-OFDM. Since the detection is an important task of the receiver, we dedicate chapter 4 to this subject. Sometimes, information from the detector/decoder is fed back to the estimator unit to improve the quality of the estimates. The receiver iterates between estimation and detection/decoding.

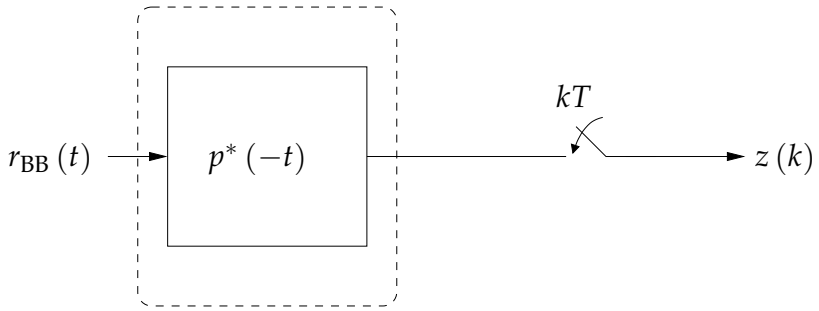

Figure 2.9: Matched filtering and sampling.

2.4 Imperfectly Synchronized Receiver

In general, the transmitter and receiver are not perfectly synchronized. The main parameters that have to be estimated by the receiver in an OFDM system are the starting point of the FFT window (or the time delay offset), the frequency offset due to the inaccuracies of the transmitter and receiver oscillators, and the channel impulse response [36]. The channel impulse response has already been introduced in section 2.2. In this section we will discuss:

- the time delay offset
- the carrier frequency offset

2.4.1 Time Delay Offset

At the receiver, we consider a block of $2(N+\nu)+L-2$ received signal samples $\mathbf{z}' = (z'(0), \ldots, z'(2(N+\nu)+L-3))^T$ (see figure 2.10). The samples of $\mathbf{z}'$ are given by

$$z'(k) = z(k-k_0),\ k = 0, \ldots, 2(N+\nu)+L-3 \tag{2.27}$$

where $z(k)$ is defined in (2.26). Every transmitted OFDM block has a duration of $N+\nu$ samples, so after transmission over a channel with an impulse response consisting of L samples, every transmitted OFDM block contributes to $N+\nu+L-1$ successive samples of the received signal. Because of its length, the vector $\mathbf{z}'$ contains the total contribution from only one OFDM block (along with partial contributions from adjacent blocks). Without loss of generality we can give this block the index $i = 0$. To be able to detect the transmitted information during this block, the receivers needs to know which samples in $\mathbf{z}'$ correspond to this transmitted OFDM block, and more precisely the starting point k_0 of this block in the received signal vector $\mathbf{z}'$. The parameter k_0 is called the *time delay offset*. Chapter 7 is dedicated to the estimation of k_0.

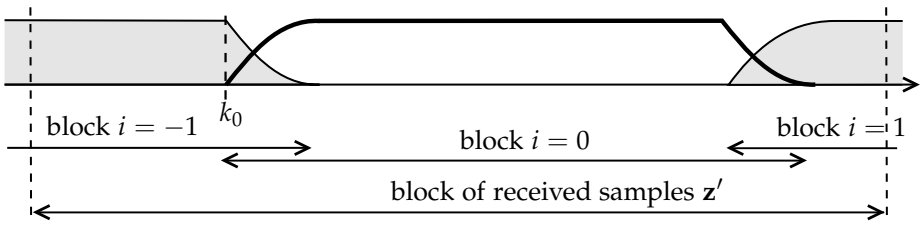

Figure 2.10: Definition of the received signal vector $\mathbf{z}'$.

2.4.2 Carrier Frequency Offset

The transmitter and the receiver have both an oscillator that operates at frequency f_c for the up conversion and down conversion respectively. In theory both oscillators should be equal, but in practice it is impossible to create two free running oscillators with exactly the same central frequency and the same phase, so there exists a frequency offset and a phase difference between transmitter and receiver. The phase difference is not much of an issue as it can be incorporated in the channel impulse response. Secondly, the movement of the transmitter and/or the receiver results in a Doppler shift which also adds to the frequency offset. The difference between the carrier frequency at the receiver side and the transmitters side is called the *carrier frequency offset* (CFO). We can model the carrier frequency $f_{c,r}$ at the receiver side as

$$f_{c,r} = f_{c,t} - \Delta_f \tag{2.28}$$

where $f_{c,t}$ is the carrier frequency at the transmitter and Δ_f is the CFO. We define ϵ as the frequency offset normalized by the carrier spacing:

$$\epsilon = \Delta_f NT. \tag{2.29}$$

Up until now we assumed that $f_{c,r} = f_{c,t}$ in section 2.3. The received signal after down conversion and lowpass filtering is no longer given by (2.18). The down conversion can be expressed as:

$$\begin{aligned} \sqrt{2} r_{BP}(t) e^{-j2\pi f_{c,r} t} &= \left(r_{BB}(t) e^{j2\pi f_{c,t} t} + r_{BB}^*(t) e^{-j2\pi f_{c,t} t} \right) e^{-j2\pi \left(f_{c,t} - \Delta_f\right) t} \\ &= r_{BB}(t) e^{j2\pi \Delta_f t} + r_{BB}^*(t) e^{-j2\pi \left(2 f_{c,t} - \Delta_f\right) t} \end{aligned} \tag{2.30}$$

After lowpass filtering, the received signal is given by

$$r(t) = r_{BB}(t) e^{j2\pi \Delta_f t}, \tag{2.31}$$

where we assume that Δ_f is much smaller than the one sided bandwidth B of the low pass filter[2]. As described in subsection 2.3.2, the signal $r(t)$ is then

[2]In a practical situation, the CFO is very small so the assumption is valid.

passed through a matched filter and sampled. We assume that $e^{j2\pi\Delta_f t}$ does not change much during the duration of $p(t)$ so

$$\int_{-\infty}^{+\infty} p(t+u)e^{j2\pi\Delta_f(t+u)}p^*(u)\,du = g(t)\,e^{j2\pi\Delta_f t}. \tag{2.32}$$

This yields for $z(k)$:

$$z(k) = \sum_{i=-\infty}^{+\infty}\sum_{k'=0}^{N+\nu-1} s_i(k')\,h\left(\left(k-k'\right)T - i\left(N+\nu\right)T\right)e^{j2\pi\frac{ek}{N}} + w(k). \tag{2.33}$$

2.5 Chapter Summary

In this chapter, we have introduced the basic parts of a communication system: the transmitter, the channel and the receiver. At the transmitter side we have described how to transform the information bits into a continuous time bandpass signal that can be transmitted over the channel. The receiver transforms the received signal to a discrete time representation and tries to detect the transmitted bits. The estimation and detection process are only briefly discussed. In the last part of this chapter, the received signal model was altered so that a frequency offset and a time delay also can be incorporated. The detection process itself, will be treated in depth in chapter 4, and more details about the parameter estimation can be found in chapters 6-8. But first we will give some theoretical background about estimation and detection.

3

Estimation and Detection Theory

In this chapter we review some basics of estimation and detection theory, which are part of the statistical inference theory. Some common statistical techniques that are widely used in the field of digital communications are introduced. The basic goal of those techniques is to obtain a value for a set of parameters from a set of observations which are contaminated with noise.

First, we introduce the concepts of maximum likelihood decision and estimation, which are the starting points for many estimation and detection algorithms. Then, some performance measures are defined and finally, the expectation-maximization (EM) algorithm is discussed.

3.1 Problem Description

We start from an observation vector $\mathbf{r} = [r(1), \ldots, r(M)]^T$ which depends on a specific realization of a vector of parameters $\mathbf{u} = [u_1, \ldots, u_K]^T$. The pa-

rameter space of $\mathbf{u}$ is denoted as $\boldsymbol{\mathcal{U}} = \mathcal{U}_1 \times \ldots \times \mathcal{U}_K$, where $\mathcal{U}_k$ is the parameter space of the parameter u_k. The main goal is to obtain a value of the unknown parameter vector $\mathbf{u}$. The statistical relationship between $\mathbf{u}$ and $\mathbf{r}$ is described by the *conditional probability density function* $p(\mathbf{r}|\mathbf{u})$. When an observation $\mathbf{r}$ is given and $\mathbf{u}$ has to be estimated, $p(\mathbf{r}|\mathbf{u})$ can be viewed as a function of $\mathbf{u}$. In that case, $p(\mathbf{r}|\mathbf{u})$ is called the *likelihood function* of $\mathbf{u}$ and the natural logarithm of the likelihood function $\ln p(\mathbf{r}|\mathbf{u})$ is called the *log likelihood function* of $\mathbf{u}$. We denote the obtained value as $\hat{\mathbf{u}}(\mathbf{r}) = [\hat{u}_1(\mathbf{r}), \ldots, \hat{u}_K(\mathbf{r})]^T$. When the components of $\mathbf{u}$ are *continuous* values or *discrete* values, we call $\hat{\mathbf{u}}(\mathbf{r})$ the *estimate* of $\mathbf{u}$ or the *decision* of $\mathbf{u}$ respectively.

Sometimes the observation vector $\mathbf{r}$ depends on other unknown parameters $\mathbf{v}$ besides $\mathbf{u}$. We call the components of $\mathbf{v}$ latent or nuisance parameters. In a communication system, the channel parameters are typical examples of nuisance parameters for the detection of the transmitted data symbols, while the data symbols themselves are nuisance parameters for the estimation of the channel parameters. In this case the observation $\mathbf{r}$ corresponds to the statistical model $p(\mathbf{r}|\mathbf{u},\mathbf{v})$, which is the conditional probability density function of $\mathbf{r}$ given $\mathbf{u}$ and $\mathbf{v}$. To obtain $\hat{\mathbf{u}}(\mathbf{r})$, we need to compute the likelihood function of $\mathbf{u}$ by averaging $p(\mathbf{r}|\mathbf{u},\mathbf{v})$ over the a priori distribution $p(\mathbf{v})$ of $\mathbf{v}$:

$$p(\mathbf{r}|\mathbf{u}) = \mathbb{E}_{\mathbf{v}}\left[p(\mathbf{r}|\mathbf{u},\mathbf{v})\right]. \tag{3.1}$$

In the last part of this chapter, we will introduce the expectation-maximization (EM) algorithm which comes in handy to deal with nuisance parameters.

3.2 Bayesian versus Non-Bayesian Approach

The Bayesian Estimation Theory associates a joint a priori probability density function $p(\mathbf{u}) = p(u_1, \ldots, u_K)$ to the unknown parameters $\mathbf{u}$ and treats them as random variables. Basically, all parameters (also the nuisance parameters) are assumed to be random variables which results in a consistent approach. This is in contrast with the non-Bayesian approach which assumes that the unknown parameters are deterministic, while the nuisance parameters are usually treated as random variables. However it is easy to incorporate the non-Bayesian approach into the Bayesian framework by ignoring the a priori information about $\mathbf{u}$ and assuming that $\mathbf{u}$ follows a uniform distribution over $\boldsymbol{\mathcal{U}}$.

The a priori information about $\mathbf{u}$ and the information obtained from the observation $\mathbf{r}$ are combined by applying Bayes' rule to compute the a posteriori distribution of the unknown parameter vector $\mathbf{u}$

$$p(\mathbf{u}|\mathbf{r}) = \frac{p(\mathbf{r}|\mathbf{u})\,p(\mathbf{u})}{p(\mathbf{r})}. \tag{3.2}$$

Note that $p(\mathbf{r})$ is given by

$$p(\mathbf{r}) = \mathbb{E}_{\mathbf{u}}\left[p(\mathbf{r}|\mathbf{u})\right], \tag{3.3}$$

which is in general too complex to compute.

3.3 Estimation

In this part we will discuss the problem of Bayesian parameter estimation. We speak of *estimation* when the components of the parameter vector $\mathbf{u}$ are continuous-valued. Furthermore, the components of $\mathbf{u}$ are assumed to be real-valued. This is not a restriction since complex-valued parameters can be expressed as their real and imaginary parts which are both real-valued.

There exist several criteria to obtain an estimate of the parameter vector $\mathbf{u}$. One option is to search for the estimate of $\mathbf{u}$ that results in the highest a posteriori probability. This estimate is called the maximum a posteriori (MAP) estimate and is given by

$$\hat{\mathbf{u}}_{\mathrm{MAP}}(\mathbf{r}) = \arg\max_{\mathbf{u}} p(\mathbf{u}|\mathbf{r}). \tag{3.4}$$

This expression can be simplified by applying Bayes' rule (3.2), which yields

$$\hat{\mathbf{u}}_{\mathrm{MAP}}(\mathbf{r}) = \arg\max_{\mathbf{u}} p(\mathbf{r}|\mathbf{u})\, p(\mathbf{u}), \tag{3.5}$$

where we neglected the denominator $p(\mathbf{r})$ in (3.2) because it does not depend on $\mathbf{u}$. Note that the expression (3.5) does not require the computation of $p(\mathbf{r})$ in contrast with (3.4).

When the parameter vector $\mathbf{u}$ has a uniform a priori distribution, there is no a priori information available about $\mathbf{u}$ and the MAP estimate of $\mathbf{u}$ from (3.5) reduces to the maximum likelihood (ML) estimate of $\mathbf{u}$, which is given by

$$\hat{\mathbf{u}}_{\mathrm{ML}}(\mathbf{r}) = \arg\max_{\mathbf{u}} p(\mathbf{r}|\mathbf{u}). \tag{3.6}$$

Since the ML estimator does not exploit any a priori information about $\mathbf{u}$, it can also be considered as a non-Bayesian estimator.

To compare the performances of different estimators we need some kind of performance measure. The performance of an estimator of a continuous-valued parameter vector, is usually illustrated by means of its mean squared error (MSE) which is defined as

$$\mathrm{MSE}(\mathbf{u}) = \mathbb{E}_{\mathbf{r}|\mathbf{u}}\left[\left|\mathbf{u} - \hat{\mathbf{u}}(\mathbf{r})\right|^2\right]. \tag{3.7}$$

The MSE as a performance measure has the advantage that there are already lots of results available for all sorts of estimators in the literature. Besides

that, there exist a number of theoretical lower bounds for the MSE. The most well known is called the Cramer-Rao Lower bound (CRB). The CRB is a lower bound on the MSE of an unbiased estimate[1] of a deterministic parameter [37]. It results from the following inequality

$$\mathbb{E}_{\mathbf{r}|\mathbf{u}}\left[\left(\mathbf{u}-\hat{\mathbf{u}}\left(\mathbf{r}\right)\right)\left(\mathbf{u}-\hat{\mathbf{u}}\left(\mathbf{r}\right)\right)^{T}\right] \geq \mathbf{J}_{\mathbf{u}}^{-1}, \tag{3.8}$$

where the matrix $\mathbf{J_u}$ is the Fischer Information Matrix. The elements of $\mathbf{J_u}$ are defined as

$$\begin{aligned}\left(\mathbf{J_u}\right)_{m,n} &= \mathbb{E}_{\mathbf{r}|\mathbf{u}}\left[\frac{\partial \ln p\left(\mathbf{r}|\mathbf{u}\right)}{\partial u_m}\frac{\partial \ln p\left(\mathbf{r}|\mathbf{u}\right)}{\partial u_n}\right] \\ &= -\mathbb{E}_{\mathbf{r}|\mathbf{u}}\left[\frac{\partial^2 \ln p\left(\mathbf{r}|\mathbf{u}\right)}{\partial u_m \partial u_n}\right],\end{aligned} \tag{3.9}$$

or using matrix notations:

$$\mathbf{J_u} = \mathbb{E}_{\mathbf{r}|\mathbf{u}}\left[\left(\frac{\partial \ln p\left(\mathbf{r}|\mathbf{u}\right)}{\partial \mathbf{u}}\right)\left(\frac{\partial \ln p\left(\mathbf{r}|\mathbf{u}\right)}{\partial \mathbf{u}}\right)^{T}\right]. \tag{3.10}$$

From (3.8) it follows that

$$\mathbb{E}_{\mathbf{r}|\mathbf{u}}\left[\left(u_k-\hat{u}_k\left(\mathbf{r}\right)\right)^2\right] \geq \left(\mathbf{J}_{\mathbf{u}}^{-1}\right)_{k,k}, \tag{3.11}$$

so the lower bound for the MSE is given by

$$\mathrm{MSE}\left(\mathbf{u}\right) \geq \mathrm{tr}\left(\mathbf{J}_{\mathbf{u}}^{-1}\right) \tag{3.12}$$

where $\mathrm{tr}\left(\mathbf{X}\right)$ returns the trace of the square matrix $\mathbf{X}$.

The comparison of the MSE with the CRB gives an idea about the performance of an unbiased estimator: the CRB is a lower bound on the achievable variance for an unbiased estimator and any estimator which has an MSE which equals the CRB, is called an efficient estimator. An unbiased estimator which has a lower variance than all other unbiased estimators is called the minimum variance unbiased (MVU) estimator. As a consequence, when an efficient estimator exists, it is necessarily also the MVU estimator. Note however that if the MVU estimator exists, it is not necessarily efficient: having the lowest variance does not mean that the equality holds in (3.12).

3.4 Detection

This part treats the Bayesian detection problem. We speak of a detection problem when the components of the unknown parameter vector $\mathbf{u}$ are discrete-valued random variables. In contrast with the estimation problem where we

[1] We call $\hat{\mathbf{u}}\left(\mathbf{r}\right)$ an unbiased estimate of $\mathbf{u}$ when the expected value $\mathbb{E}_{\mathbf{r}|\mathbf{u}}\left[\hat{\mathbf{u}}\left(\mathbf{r}\right)\right]$ is equal to $\mathbf{u}$.

seek estimates which are closely located to the actual parameter, the detection reduces to a hypothesis testing. Based on the observation $\mathbf{r}$ we have to make a decision about the parameter vector $\mathbf{u}$.

The most well known detection algorithm is MAP detection. The MAP detector maximizes the a posteriori probability $p(\mathbf{u}|\mathbf{r})$ with respect to $\mathbf{u}$:

$$\begin{aligned} \hat{\mathbf{u}}_{\mathrm{MAP}}(\mathbf{r}) &= \arg\max_{\mathbf{u}\in\mathcal{U}} p(\mathbf{u}|\mathbf{r}). \\ &= \arg\max_{\mathbf{u}\in\mathcal{U}} p(\mathbf{r}|\mathbf{u})\, p(\mathbf{u}). \end{aligned} \tag{3.13}$$

The second step is obtained after applying Bayes' rule and neglecting $p(\mathbf{r})$ because it is independent of $\mathbf{u}$. This detector minimizes the probability for an erroneous detection, i.e $\Pr[\hat{\mathbf{u}} \neq \mathbf{u}]$ [1].

When there is no a priori information available about $\mathbf{u}$, the MAP detector yields the same results as the ML detector which is defined as

$$\hat{\mathbf{u}}_{ML}(\mathbf{r}) = \arg\max_{\mathbf{u}\in\mathcal{U}} p(\mathbf{r}|\mathbf{u}). \tag{3.14}$$

In a communication system, the transmitted data bits are an example of unknown discrete-valued random variables which have to be detected. The probability for an erroneous detection of a bit is called the *bit error rate* (BER). It is an important measure for the performance of a detection algorithm. In the literature, numerous examples of MAP and ML bit detectors can be found. As already mentioned those detectors have the favorable property that they minimize the BER. The performance of other detectors is usually compared with the ML or MAP detector and expressed in terms of the BER. Another class of detectors are the MAP and ML symbol detectors which in their turn result in the minimal symbol error rate (SER).

Sometimes the a posteriori probability of a single information bit or a symbol are too complex to obtain. In that case the total sequence of symbols is detected at once. The MAP and ML sequence detectors minimize the probability that the detected data sequence differs from from the transmitted data sequence.

3.5 The Expectation-Maximization Algorithm

As we have already mentioned in section 3.1, the presence of nuisance parameters gives rise to difficulties in obtaining the likelihood function. This makes parameter estimation often intractable in the presence of nuisance parameters. To solve this problem the iterative EM algorithm was proposed in [38]. It avoids the direct computation of (3.1). In this subsection we try to give an understandable reasoning behind the EM algorithm. This part is largely inspired by the part about the EM algorithm for the MAP estimator in [39, Section 2.2].

3.5.1 Problem Statement

We want to obtain an estimate of a parameter vector $\mathbf{u} \in \boldsymbol{\mathcal{U}}$ based on the observation $\mathbf{r}$. Besides $\mathbf{u}$, the observation $\mathbf{r}$ also depends on the unobserved parameter vector $\mathbf{v} \in \mathcal{V}$. The observation corresponds to the following statistical model

$$p\left(\mathbf{r}|\mathbf{u},\mathbf{v}\right). \tag{3.15}$$

Since the parameter vector $\mathbf{v}$ is unknown, we have to average the conditional probability density function over $\mathbf{v}$:

$$\begin{aligned} p\left(\mathbf{r}|\mathbf{u}\right) &= \int_{\mathbf{v}} p\left(\mathbf{r},\mathbf{v}|\mathbf{u}\right) d\mathbf{v} \\ &= \int_{\mathbf{v}} p\left(\mathbf{r}|\mathbf{u},\mathbf{v}\right) p\left(\mathbf{v}|\mathbf{u}\right) d\mathbf{v} \end{aligned} \tag{3.16}$$

and then we can apply the estimation / detection algorithms from the previous sections starting from the likelihood function $p\left(\mathbf{r}|\mathbf{u}\right)$. The integral in (3.16) is in general difficult to compute, so we need a different solution which avoids this problem. So let us consider the ML estimator in the presence of nuisance parameters, which yields the following maximization problem:

$$\begin{aligned} \hat{\mathbf{u}} &= \arg\max_{\mathbf{u}\in\mathcal{U}} \ln p\left(\mathbf{r}|\mathbf{u}\right) \\ &= \arg\max_{\mathbf{u}\in\mathcal{U}} \ln \int_{\mathbf{v}} p\left(\mathbf{r},\mathbf{v}|\mathbf{u}\right) d\mathbf{v}. \end{aligned} \tag{3.17}$$

Compared to (3.6), we have inserted an extra logarithm but this does not affect the result since the logarithm is a monotonically increasing function. Now we want to simplify the function that has to be maximized and that's in fact what the EM algorithm does.

3.5.2 A Lower Bound on the Likelihood Function

The maximization of the log likelihood function $\ln p\left(\mathbf{r}|\mathbf{u}\right)$ is too complex so we are going to derive a function which is easier to maximize. We start with deriving a lower bound on $\ln p\left(\mathbf{r}|\mathbf{u}\right)$.

The derivation is based on the Kullback-Leibler divergence [40], which is a widely used concept in the field of information theory. The Kullback-Leibler divergence measures the difference between two probability functions and is defined as follows

$$D\left(q\left(\mathbf{v}\right) || \; p\left(\mathbf{v}\right)\right) = \int_{\mathbf{v}} q\left(\mathbf{v}\right) \ln \frac{q\left(\mathbf{v}\right)}{p\left(\mathbf{v}\right)} d\mathbf{v}, \tag{3.18}$$

where $p\left(\mathbf{v}\right)$ and $q\left(\mathbf{v}\right)$ are probability functions. A well known property about $D\left(q\left(\mathbf{v}\right)||p\left(\mathbf{v}\right)\right)$ is that it is non-negative [40]. This fact can be exploited to

obtain a lower bound on $\ln p(\mathbf{r}|\mathbf{u})$:

$$\begin{aligned} F(\mathbf{u}, q) &= \ln p(\mathbf{r}|\mathbf{u}) - \int_{\mathbf{v}} q(\mathbf{v}) \ln \frac{q(\mathbf{v})}{p(\mathbf{v}|\mathbf{r}, \mathbf{u})} d\mathbf{v} \\ &\leq \ln p(\mathbf{r}|\mathbf{u}), \end{aligned} \tag{3.19}$$

where $q(\mathbf{v})$ represents an arbitrary distribution. By applying Bayes' rule, $F(\mathbf{u}, q)$ can be rewritten as

$$F(\mathbf{u}, q) = R(\mathbf{u}, q) + H(q) \tag{3.20}$$

where $R(\mathbf{u}, q)$ and $H(q)$ are defined as

$$R(\mathbf{u}, q) = \int_{\mathbf{v}} q(\mathbf{v}) \ln p(\mathbf{v}, \mathbf{r}|\mathbf{u}) \, d\mathbf{v} \tag{3.21}$$

$$H(q) = -\int_{\mathbf{v}} q(\mathbf{v}) \ln q(\mathbf{v}) \, d\mathbf{v} \geq 0. \tag{3.22}$$

Based on this lower bound, a new estimator of $\mathbf{u}$ can be defined:

$$\hat{\mathbf{u}} = \arg\max_{\mathbf{u}} F(\mathbf{u}, q) = \arg\max_{\mathbf{u}} R(\mathbf{u}, q), \tag{3.23}$$

where the last step is valid because $H(q)$ is independent of $\mathbf{u}$. The maximum of $R(\mathbf{u}, q)$ is easier to obtain than the maximum of (3.17) because the integral in $R(\mathbf{u}, q)$ is outside the logarithmic function. Since probability functions usually consist of exponential functions, the logarithmic and the exponential functions cancel out each other in $R(\mathbf{u}, q)$, which is a desirable property. As a result, the maximum with respect to $\mathbf{u}$ can be written in a closed form expression in many cases. This decreases the computational complexity of the maximization problem considerably. The next step consists of making an appropriate choice for the probability function $q(\mathbf{v})$, which is handled by the EM algorithm.

3.5.3 The actual EM algorithm

The EM algorithm starts from an initial estimate $\hat{\mathbf{u}}[0]$, which is used to make a choice about $q(\mathbf{v})$: the function $q(\mathbf{v})$ is chosen in such a way that the lower bound $F(\mathbf{u}, q)$ is equal to the true log likelihood function for $\mathbf{u} = \hat{\mathbf{u}}[0]$. From (3.19), we see that

$$F(\hat{\mathbf{u}}[0], q_0) = \ln p(\mathbf{r}|\hat{\mathbf{u}}[0]) \tag{3.24}$$

for

$$q_0(\mathbf{v}) = p(\mathbf{v}|\mathbf{r}, \hat{\mathbf{u}}[0]). \tag{3.25}$$

The estimate of $\mathbf{u}$ can be updated by maximizing the obtained lower bound $F(\mathbf{u}, q_0)$

$$\hat{\mathbf{u}}[1] = \arg\max_{\mathbf{u}} F(\mathbf{u}, q_0) = \arg\max_{\mathbf{u}} R(\mathbf{u}, q_0). \tag{3.26}$$

The new estimate $\hat{\mathbf{u}}[1]$ exhibits a higher log likelihood than $\hat{\mathbf{u}}\,[0]$, because

$$\begin{aligned} \ln p\left(\mathbf{r}|\hat{\mathbf{u}}[1]\right) &\geq F\left(\hat{\mathbf{u}}\,[1]\,, q_0\right) \\ &\geq F\left(\hat{\mathbf{u}}\,[0]\,, q_0\right) \\ &= \ln p\left(\mathbf{r}|\hat{\mathbf{u}}\,[0]\right), \end{aligned} \tag{3.27}$$

where we have used (3.19), (3.26) and (3.24). In every iteration i that follows, this process is repeated by the EM algorithm: First a new lower bound is computed based on the latest obtained estimate $\hat{\mathbf{u}}\,[i-1]$. This lower bound is then maximized with respect to $\mathbf{u}$ and a new estimate $\hat{\mathbf{u}}\,[i]$ is obtained which has a higher log likelihood than $\hat{\mathbf{u}}\,[i-1]$.

The first step of each EM iteration is called the expectation step (E-step) and consists of the computation of the function $Q\left(\mathbf{u}|\hat{\mathbf{u}}\,[i-1]\right) = R\left(\mathbf{u}, q_{i-1}\right)$. The second step is called the maximization step (M-step), which maximizes the obtained function $Q\left(\mathbf{u}|\hat{\mathbf{u}}\,[i-1]\right)$:

$$\text{E-step:} \quad Q\left(\mathbf{u}|\hat{\mathbf{u}}\,[i-1]\right) = \int_{\mathbf{v}} p\left(\mathbf{v}|\mathbf{r}, \hat{\mathbf{u}}\,[i-1]\right) \ln p\left(\mathbf{v}, \mathbf{r}|\mathbf{u}\right) d\mathbf{v} \tag{3.28}$$

$$\text{M-step:} \quad \hat{\mathbf{u}}\,[i] = \arg\max_{\mathbf{u}} Q\left(\mathbf{u}|\hat{\mathbf{u}}[i-1]\right). \tag{3.29}$$

When $\mathbf{u}$ and $\mathbf{v}$ are independent, the E-step can be further simplified to

$$Q\left(\mathbf{u}|\hat{\mathbf{u}}\,[i-1]\right) = \int_{\mathbf{v}} p\left(\mathbf{v}|\mathbf{r}, \hat{\mathbf{u}}\,[i-1]\right) \ln p\left(\mathbf{r}|\mathbf{v}, \mathbf{u}\right) d\mathbf{v}, \tag{3.30}$$

where we have omitted the term independent of $\mathbf{u}$.

The E-step and M-step are repeated until convergence is reached. The obtained estimates $\hat{\mathbf{u}}\,[0]$, $\hat{\mathbf{u}}\,[1]$, ... exhibit a non decreasing log likelihood, so when the initial estimate $\hat{\mathbf{u}}\,[0]$ is close to the ML estimate of $\mathbf{u}$, the EM algorithm converges towards the ML solution.

However, when the initial estimate $\hat{\mathbf{u}}\,[0]$ is in the neighborhood of a local maximum, the EM algorithm will get stuck in this local maximum.

3.6 Chapter Summary

In this chapter we have briefly introduced some topics from the detection and estimation theory which will come in handy in the next chapters.

The estimation theory will be used to obtain estimates about the channel and synchronization parameters. We have reviewed some well known estimators (MAP and ML). The performance of a particular estimator is measured by its MSE and is compared to the corresponding CRB, which is a theoretical lower bound.

The detection theory and its algorithms are necessary to the detect the transmitted data symbols at the receiver. The MAP and ML detector were defined in this chapter.

In the last part of this chapter, the EM algorithm has been treated. It is a suboptimal iterative algorithm, which can be applied when the optimal estimators (ML) are too complex.

4 Data Detection

This chapter deals with data detection. We assume a perfectly synchronized receiver, which means that all synchronization parameters are assumed to be known. For the actual data detection in a KSP-OFDM system, there are several possibilities. Basically, the different data detection algorithms can be categorized in two categories:

- Time-domain data detection, see section 4.1
- Frequency-domain data detection, see section 4.2

In the last part of this chapter we compare the performance of the different proposed data detection algorithms by means of some numerical results.

4.1 Time-domain Detection

Figure 4.1 shows a part of the time domain signals at the transmitter and the receiver respectively. For the data detection of block $i = 0$, we take the $N + \nu$ samples of the received signal from the observation interval corresponding to

block $i = 0$ shown in figure 4.1. The vector $\mathbf{z}_0 = [z(0), \ldots, z(N+\nu-1)]^T$ collects the $N+\nu$ received signal samples and can be expressed as

$$\mathbf{z}_0 = \mathbf{H}_{\text{ch}}\mathbf{s}_0 + \mathbf{w}_0 \tag{4.1}$$

where the $(N+\nu) \times (N+\nu)$ channel matrix $\mathbf{H}_{\text{ch}}$ is defined as

$$(\mathbf{H}_{\text{ch}})_{l,m} = h\left(|l-m|_{N+\nu} T\right). \tag{4.2}$$

The notation $|x|_K$ denotes the modulo-K reduction of x, returning a result in the interval $[0,\ K[$. Note that $h(lT) = 0$ for $l < 0$ or $l \geq L$. The vector $\mathbf{w}_0$ collects the noise samples $w(k)$, for $k = 0,\ \ldots,\ N+\nu-1$, and $\mathbf{s}_0$ is defined in (2.8).

The contributions from the time-domain pilot symbols (dark grey areas on figure 4.1) are first subtracted from the received signal. The resulting signal vector $\mathbf{y}_0$ is given by

$$\begin{aligned} \mathbf{y}_0 &= \mathbf{z}_0 - \sqrt{\frac{N}{N+\nu}}\mathbf{H}_{\text{ch}} \begin{pmatrix} \mathbf{0}_{(N-\nu)\times 1} \\ \mathbf{a}_{\text{g}} \end{pmatrix} && (4.3) \\ &= \sqrt{\frac{N}{N+\nu}}\mathbf{H}\mathbf{F}^H\mathbf{a}_0 + \mathbf{w}_0 && (4.4) \end{aligned}$$

where the $(N+\nu) \times N$ channel matrix $\mathbf{H}$ is defined as

$$(\mathbf{H})_{l,m} = h\left((l-m)T\right), \tag{4.5}$$

and the notation $\mathbf{0}_{m\times n}$ denotes an $m \times n$ matrix with all entries equal to zero.

The likelihood function associated with $\mathbf{a}_0$ is given by

$$p(\mathbf{y}_0|\mathbf{a}_0 = \boldsymbol{\omega}) \propto \exp\left(-\frac{1}{N_0}\left\|\mathbf{y}_0 - \sqrt{\frac{N}{N+\nu}}\mathbf{H}\mathbf{F}^H\boldsymbol{\omega}\right\|^2\right). \tag{4.6}$$

We assume that the receiver knows $\mathbf{H}$ perfectly.

4.1.1 Optimal Detection

In this section, we consider the detection of the transmitted data based on the observation model (4.4).

4.1.1.1 Uncoded System

Since a frequency selective fading channel is not memoryless, bit by bit decision or symbol by symbol decision are not easily performed because they respectively comprise the computation of the likelihood function of every bit of $\mathbf{a}_0$ or every component of $\mathbf{a}_0$. Therefore the symbol vector $\mathbf{a}_0$ is detected

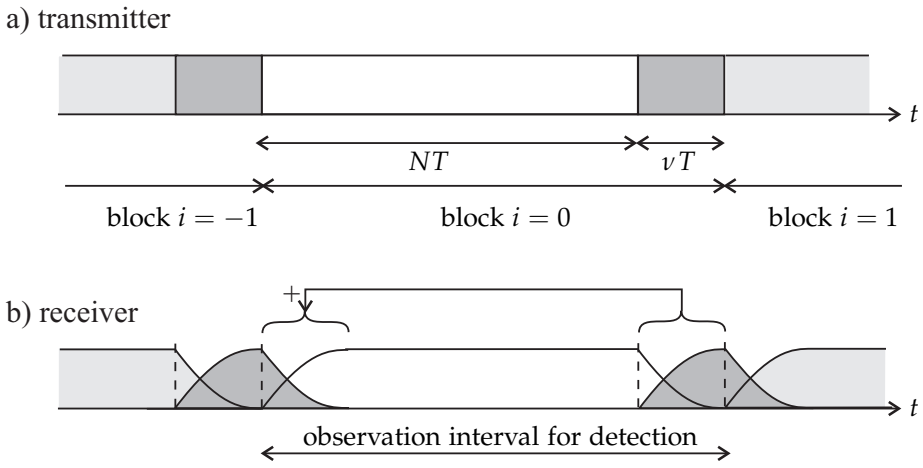

Figure 4.1: Time-domain signal of a KSP-OFDM block a) transmitted signal b) received signal and observation interval for detection.

at once. The theoretical optimal detector is called the ML sequence detector (MLSD) [2] and is defined as:

$$\begin{aligned} \hat{\mathbf{a}}_0 &= \arg \max_{\boldsymbol{\omega} \in \Omega_{M_\mathrm{s}}^N} p\left(\mathbf{y}_0 | \mathbf{a}_0 = \boldsymbol{\omega}\right) \\ &= \arg \min_{\boldsymbol{\omega} \in \Omega_{M_\mathrm{s}}^N} \left\| \mathbf{y}_0 - \sqrt{\frac{N}{N+\nu}} \mathbf{H}\mathbf{F}^H \boldsymbol{\omega} \right\|^2 . \end{aligned} \tag{4.7}$$

This detector is called optimal because it minimizes the probability that the detected symbol vector $\hat{\mathbf{a}}_0$ differs from the transmitted symbol vector $\mathbf{a}_0$.

The problem from (4.7) can only be solved by an exhaustive search over all M_s^N possible symbol vectors $\mathbf{a}_0 \in \Omega_{M_\mathrm{s}}^N$, which means that the complexity grows exponential with the number of carriers. The ML sequence detector is not suitable in a practical situation, but its performance serves as a benchmark for practical suboptimal detection algorithms.

4.1.1.2 Coded System

For a general error correcting code, the a posteriori probability of every information bit is not easily obtained, so the MAP information bit detector, which minimizes the BER, can not be applied. The optimal decoder is similar to the optimal detector for an uncoded system. An exhaustive search is performed over all 2^{N_b} possible legitimate symbol vectors $\mathbf{a}_0 \in \zeta_{\mathcal{D}}$. The computational complexity grows exponentially with the number of information bits.

For a turbo code, the standard detection algorithm is an iterative algorithm which approximates the MAP information bit detector [31, 41, 42]. The algorithm is an application of the sum-product algorithm to a factor graph. For the observation model from (4.4), the factor graph shows short cycles. As a result the iterative algorithm may not converge or converge to a wrong result.

The detection algorithm for a low-density parity-check (LDPC) code is also an application of the sum-product algorithm [41, 42]. The algorithm tries to iteratively estimate the a posteriori probabilities of the coded bits. After the computation of the MAP estimate of every code bit, the encoding process is reversed to determine the corresponding information bits. For a systematic LDPC code, this last step is trivial, since the N_b information bits are a part of the code word. When the algorithm converges, the performance of the MAP code bit detector is reached. However, the factor graph corresponding to the observation model from (4.4) exhibits short cycles. Convergence is not guaranteed and even if the algorithm converges, the result still might be wrong.

4.1.2 Zero-Forcing Detection

As convergence of the algorithms for data detection in the previous section is not assured for the observation model (4.4), we change the observation model by applying the observation vector $\mathbf{y}_0$ from (4.4) first to an equalizer.

The zero-forcing (ZF) detector belongs to the class of the linear receivers. We speak of a linear receiver when the detection algorithm consists of the following two steps (for an uncoded system):

- The considered observation vector $\mathbf{y}_0$ is multiplied by a $N \times (N+\nu)$ matrix $\mathbf{G}$, i.e. the equalizer

$$\mathbf{y}_0^{'} = \mathbf{G}\mathbf{y}_0. \tag{4.8}$$

- The detection problem is then reduced to finding the symbol vector which is closest to $\mathbf{y}_0^{'}$

$$\hat{\mathbf{a}}_0 = \arg \min_{\boldsymbol{\omega} \in \Omega_{M_\mathrm{s}}^N} \left\| \mathbf{y}_0^{'} - \boldsymbol{\omega} \right\|^2. \tag{4.9}$$

There exist several criteria to select the optimal matrix $\mathbf{G}$. The ZF detector wants to remove all the interference between the different data symbols that is present in $\mathbf{y}_0$, so an appropriate choice for the matrix $\mathbf{G}$ is given by

$$\mathbf{G}_\mathrm{ZF} = \sqrt{\frac{N+\nu}{N}} \left(\mathbf{F}\mathbf{H}^H\mathbf{H}\mathbf{F}^H \right)^{-1} \mathbf{F}\mathbf{H}^H. \tag{4.10}$$

It can be easily verified that $\mathbf{G}_\mathrm{ZF}$ indeed removes all the ISI:

$$\begin{aligned} \mathbf{y}_0^{'} &= \mathbf{G}_\mathrm{ZF}\mathbf{y}_0 \\ &= \mathbf{a}_0 + \mathbf{G}_\mathrm{ZF}\mathbf{w}_0, \end{aligned} \tag{4.11}$$

where we have replaced $\mathbf{y}_0$ by (4.4) in the second step. We see that the noise vector $\mathbf{w}_0$ is also multiplied by $\mathbf{G}_{\mathrm{ZF}}$, resulting in a noise vector with a covariance matrix $\mathbf{R}_{\mathrm{ZF}}$ equal to

$$\begin{aligned}\mathbf{R}_{\mathrm{ZF}} &= \mathbb{E}\left[\mathbf{G}_{\mathrm{ZF}}\mathbf{w}\mathbf{w}^H\mathbf{G}_{\mathrm{ZF}}^H\right] \\ &= \frac{N+\nu}{N}N_0\left(\mathbf{F}\mathbf{H}^H\mathbf{H}\mathbf{F}^H\right)^{-1}. \end{aligned} \tag{4.12}$$

The variance of the components of this new noise vector is usually increased as compared to the variance of the original noise samples $\mathbf{w}$, we call this effect *noise enhancement*. As a result, the performance of the ZF detector is worse than the optimal ML detector from 4.1.1, however the main advantage is the lower computational complexity of the ZF detector. Since $\mathbf{G}_{\mathrm{ZF}}$ cannot be precomputed because it needs the inversion of a matrix which depends on the channel, its computational complexity might still be too high [24].

For the coded case, we see from (4.11) that the likelihood of the symbol vector $\mathbf{a}_0$ is given by:

$$p\left(\mathbf{y}_0^{'}|\mathbf{a}_0\right) \propto \exp\left\{-\left(\mathbf{y}_0^{'}-\mathbf{a}_0\right)^H\mathbf{R}_{\mathrm{ZF}}^{-1}\left(\mathbf{y}_0^{'}-\mathbf{a}_0\right)\right\}. \tag{4.13}$$

The noise contribution in (4.11) is not white, so it is not easy to obtain the likelihood of the different symbols $a_0(n)$, $n = 0, \ldots, N-1$, from (4.13). Nevertheless, we assume that the noise contribution is white and we approximate its covariance matrix $\mathbf{R}_{\mathrm{ZF}}$ by a diagonal matrix with the diagonal elements equal to the diagonal elements of $\mathbf{R}_{\mathrm{ZF}}$. With this approximation, the n-th output of the ZF filter $y_0^{'}(n)$ consists of the n-th symbol $a_0(n)$ plus an AWGN sample with variance equal to $(\mathbf{R}_{\mathrm{ZF}})_{n,n}$. The different data symbols are now decoupled. The likelihood of $a_0(n)$is now given by

$$p\left(y_0^{'}(n)\,|a_0(n)\right) \propto \exp\left(-\frac{\left|y_0^{'}(n)-a_0(n)\right|^2}{(\mathbf{R}_{\mathrm{ZF}})_{n,n}}\right). \tag{4.14}$$

The obtained likelihood functions can now be used by the channel decoder to perform MAP bit detection.

4.2 Frequency-domain Detection

The data detection for CP-OFDM has a very low computational complexity: The CP is dropped from the received signal and the remaining N samples are transformed to the frequency domain by applying the FFT. Then per carrier single-tap equalization is performed and the data detection reduces to

symbol-by-symbol detection. The useful part of the N received signal samples in a CP-OFDM system is given by [25]

$$\mathbf{y}_{\mathrm{CP}} = \sqrt{\frac{N}{N+\nu}}\mathbf{H}_{\mathrm{CP}}\mathbf{F}^{H}\mathbf{a}_0, \tag{4.15}$$

where $\mathbf{H}_{\mathrm{CP}}$ is the $N \times N$ channel matrix:

$$(\mathbf{H}_{\mathrm{CP}})_{l,m} = h\left(|l-m|_N T\right). \tag{4.16}$$

This matrix is a circulant matrix. It is well known that $N \times N$ circulant matrices are diagonalized by pre- and post-multiplication with N-point FFT and IFFT matrices [7, 25]: $\mathbf{F}\mathbf{H}_{\mathrm{CP}}\mathbf{F}^{H} = \mathrm{diag}\left(H(0), \ldots, H(N-1)\right)$. The elements $H(n)$, $n = 0, \ldots, N-1$ are the N outputs of the N-point FFT of the channel impulse response $\mathbf{h}$, and are given by

$$H(n) = \sum_{l=0}^{L-1} h(l)\, e^{-j2\pi\frac{nl}{N}}. \tag{4.17}$$

So applying an FFT to $\mathbf{y}_{\mathrm{CP}}$ transforms the samples of $\mathbf{y}_{\mathrm{CP}}$ into samples which only depend on one data symbol. Equalization is easy and involves only a simple division.

We want to achieve something similar for our KSP-OFDM system. After removing the contributions from the guard interval pilot samples from the received signal, we obtain a system which is in fact a ZP-OFDM system [24]. The received signal is given by (4.4). To convert $\mathbf{H}$ into a circulant matrix, we have to add the last ν samples of $\mathbf{y}_0$ to the first ν samples [24, 25]. This operation can be formulated as follows

$$\mathbf{y}_0^{'} = \mathbf{\Omega}\mathbf{y}_0, \tag{4.18}$$

where $\mathbf{\Omega}$ is defined as:

$$\mathbf{\Omega} = \begin{pmatrix} \mathbf{I}_N & \begin{matrix} \mathbf{I}_\nu \\ \mathbf{0}_{(N-\nu)\times\nu} \end{matrix} \end{pmatrix}. \tag{4.19}$$

The resulting samples $\mathbf{y}_0^{'}$ are given by

$$\mathbf{y}_0^{'} = \sqrt{\frac{N}{N+\nu}}\mathbf{H}_{\mathrm{CP}}\mathbf{F}^{H}\mathbf{a}_0 + \mathbf{w}_0^{'}. \tag{4.20}$$

We see that the useful part in $\mathbf{y}_0^{'}$ is similar to (4.15). The noise samples $w_0^{'}(n)$, $n = 0, \ldots, N-1$, are slightly colored and have a covariance matrix equal to

$$\mathbf{R}_{\mathbf{w}'} = N_0\mathbf{\Omega}\mathbf{\Omega}^{H}. \tag{4.21}$$

The received signal samples $\mathbf{y}_0^{'}$ are transformed to the frequency domain, which yields:

$$\begin{aligned} \tilde{\mathbf{y}}_0 &= \mathbf{F}\mathbf{y}_0^{'} \\ &= \sqrt{\frac{N}{N+\nu}}\tilde{\mathbf{H}}\mathbf{a}_0 + \tilde{\mathbf{w}}_0, \end{aligned} \tag{4.22}$$

where $\tilde{\mathbf{H}}$ and $\tilde{\mathbf{w}}$ are the channel matrix and noise vector respectively, which are defined as

$$\begin{aligned} \tilde{\mathbf{H}} &= \mathbf{F}\mathbf{H}_{CP}\mathbf{F}^H \\ &= \operatorname{diag}\left(H\left(0\right), \ldots,\ H\left(N-1\right)\right) && (4.23) \\ \tilde{\mathbf{w}}_0 &= \mathbf{F}\mathbf{w}_0^{'} && (4.24) \end{aligned}$$

The components of $\tilde{\mathbf{w}}_0$ are zero mean and have a covariance matrix $\mathbf{R}_{\tilde{\mathbf{w}}}$ which is defined as

$$\left(\mathbf{R}_{\tilde{\mathbf{w}}}\right)_{n_1,n_2} = N_0\left(\delta_{n_2-n_1} + \frac{1}{N}\sum_{k=0}^{\nu-1} e^{j2\pi\frac{(n_2-n_1)k}{N}}\right). \tag{4.25}$$

The optimal detector based on the observations $\tilde{\mathbf{y}}_0$ is defined as

$$\hat{\mathbf{a}}_0 = \arg\min_{\boldsymbol{\omega}\in\Omega_{M_s}^N}\left(\tilde{\mathbf{y}}_0 - \sqrt{\frac{N}{N+\nu}}\tilde{\mathbf{H}}\boldsymbol{\omega}\right)^H\left(\mathbf{R}_{\tilde{\mathbf{w}}}\right)^{-1}\left(\tilde{\mathbf{y}}_0 - \sqrt{\frac{N}{N+\nu}}\tilde{\mathbf{H}}\boldsymbol{\omega}\right). \tag{4.26}$$

The optimal detector does not reduce to symbol-by-symbol detection because the noise samples $\tilde{\mathbf{w}}$ are correlated. As a consequence this minimization problem can only be solved by an exhaustive search over all M_s^N possible symbol vectors $\mathbf{a}_0 \in \Omega_{M_s}^N$, which comes with a high computational cost, as we have already mentioned in subsection 4.1.1.

However, the optimization problem of (4.26) can be simplified by noting that the non-diagonal elements of $\mathbf{R}_{\tilde{\mathbf{w}}}$ are much smaller than the diagonal elements, so we can approximate $\mathbf{R}_{\tilde{\mathbf{w}}}$ by a diagonal matrix:

$$\mathbf{R}_{\tilde{\mathbf{w}}} \approx N_0\frac{N+\nu}{N}\mathbf{I}_N. \tag{4.27}$$

When we substitute $\mathbf{R}_{\tilde{\mathbf{w}}}$ by its approximation in (4.26), the detection problem reduces to N symbol-by-symbol detection problems

$$\hat{a}_0\left(n\right) = \arg\min_{\omega\in\Omega_{M_s}}\left|\tilde{y}_0\left(n\right) - \sqrt{\frac{N}{N+\nu}}H\left(n\right)\omega\right|^2,\ n = 0,\ \ldots,\ N-1. \tag{4.28}$$

For a coded system, we can again assume that the noise samples $\tilde{w}\left(n\right)$, $n = 0,\ \ldots,\ N-1$, are white and approximate the covariance matrix $\mathbf{R}_{\tilde{\mathbf{w}}}$ by the

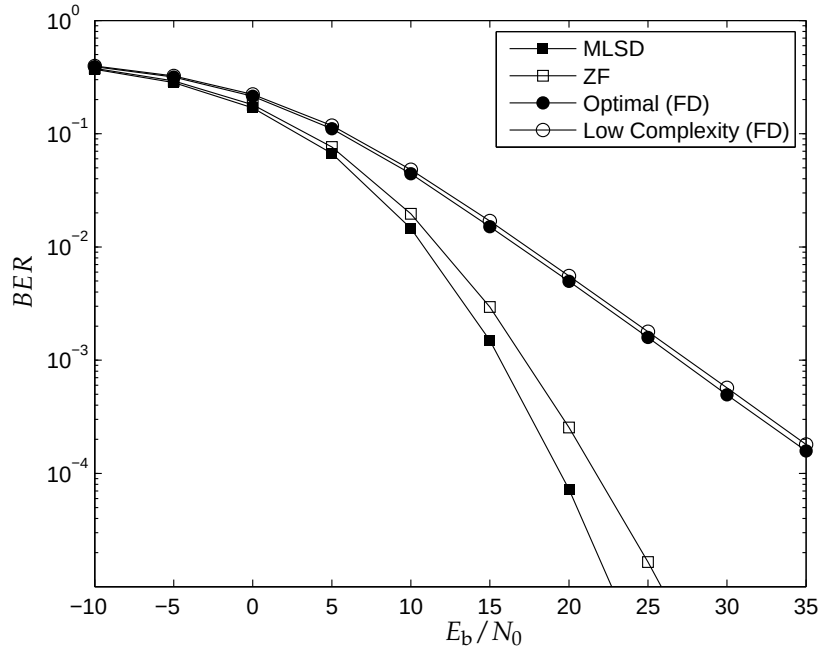

Figure 4.2: BER as a function of E_b/N_0 , $N = 8$, $\nu = 4$, $L = 4$.

diagonal matrix from (4.27). The likelihood of $a_0(n)$ based on the observation $\tilde{y}_0(n)$ is given by

$$p(\tilde{y}_0(n)|a_0(n)) \propto \exp\left(-\frac{\left|\tilde{y}_0(n) - \sqrt{\frac{N}{N+\nu}}H(n)\,a_0(n)\right|^2}{\frac{N+\nu}{N}N_0}\right).$$

Similar, as for the ZF detector, these likelihoods can now be used by the decoder to obtain the a posteriori probabilities of the code bits or information bits and MAP bit detection can be performed.

4.3 Performance Comparison and Discussion

4.3.1 Uncoded System

In this section we will illustrate the performance of the different detectors which were proposed in the previous sections.

To give an idea of the differences in performance between the various detection algorithms, we have simulated a simple KSP-OFDM system with

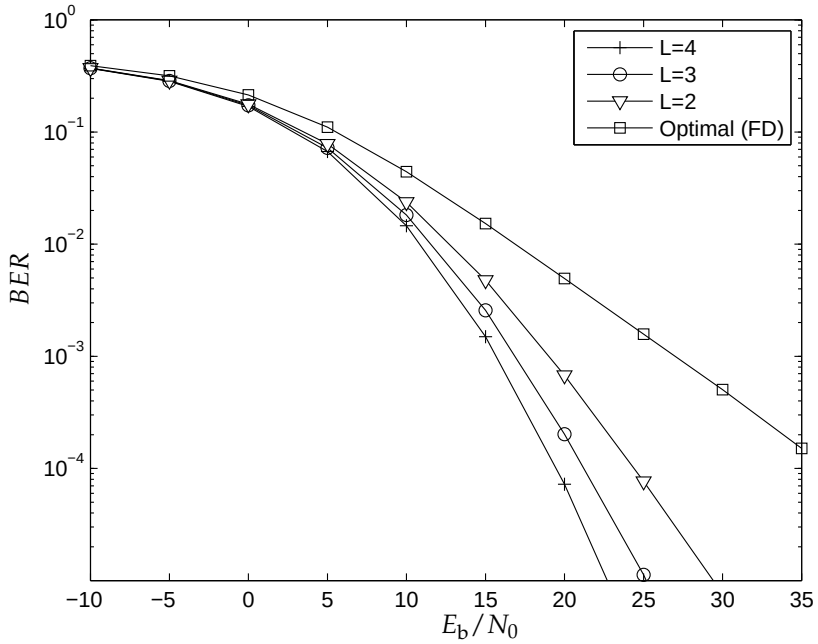

Figure 4.3: Influence of L on the BER of the MLSD, $N = 8$, $\nu = 4$.

$N = 8$ carriers and a guard interval consisting of $\nu = 4$ samples. The transmitted data symbols are randomly generated BPSK symbols. The channel is a frequency selective Rayleigh fading channel consisting of 4 symbol spaced taps ($\tau_l = lT$, $l = 0, \ldots, 3$, $\tilde{L} = 4$) with equal variance. The impulse response of the cascade of transmit pulse, channel impulse response and receiver filter is given by

$$h(t) = \sum_{l=0}^{\tilde{L}-1} \tilde{h}_l g(t - lT). \tag{4.29}$$

Because $g(t)$ is a Nyquist pulse, the samples of the channel impulse response are given by $h(kT) = \tilde{h}_k$, $k = 0, \ldots, 3$ and L is equal to the number of taps, i.e $L = \tilde{L} = 4$.

Figure 4.2 shows the BER versus the signal-to-noise ratio (SNR). As expected, the MLSD algorithm results in the lowest BER. The ZF algorithm gives rise to some degradation compared to the optimal detector, but still has a performance which is close to the BER of MLSD. Only for high values of E_b/N_0, the degradation can become significant. The FD detection algorithms result in a somewhat worse BER performance compared to the TD detection methods. The low complexity detection algorithm reaches the performance of

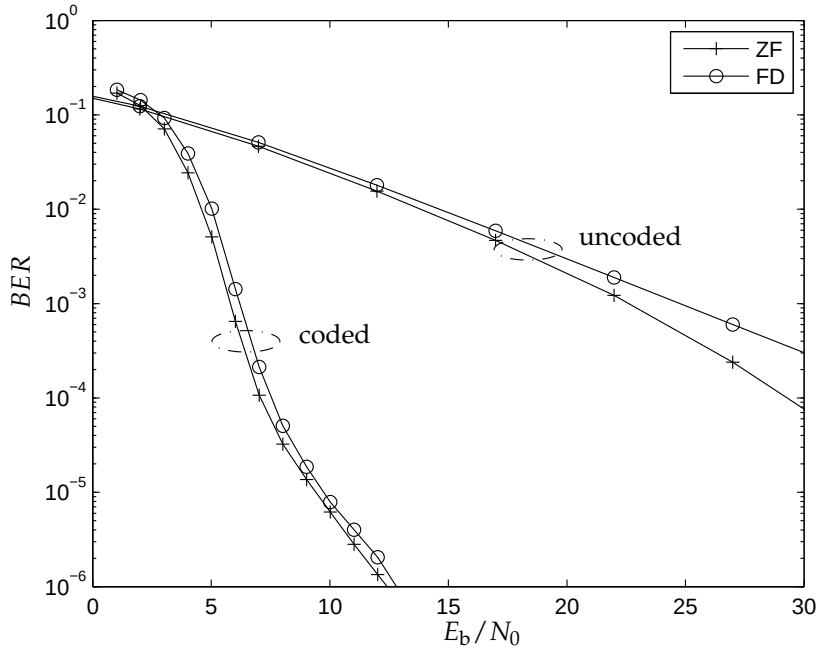

Figure 4.4: BER for coded and uncoded transmission as a function of E_b/N_0, $N = 1024$, $\nu = 100$, $L = 50$.

the optimal FD algorithm. So instead of using the optimal FD detector, the symbol-by-symbol FD detector is a valuable alternative with only a negligible loss in performance.

For high SNR, the BER has a constant slope, which is determined by the diversity of the considered data detection algorithm. The multipath diversity provided by the frequency selective fading channel is lost in the FD [24], while the MLSD achieves the maximum diversity which is equal to L in this example (see the appendix at the end of this chapter for more details, concerning the diversity order of the MLSD). The ZF detector introduces some noise enhancement and loses some of the available diversity.

Figure 4.3 illustrates the difference in diversity between the optimal TD detection algorithm and the optimal FD detection algorithm. We have simulated a channel with 2, 3 and 4 taps respectively, which means that L varies from 2 to 4. The BER of the optimal FD algorithm does not depend on the value of L and has a diversity equal to 1, while the BER of the MLSD decreases as L increases.

4.3.2 Coded System

To illustrate the different detectors in a coded system, we have simulated a more realistic KSP-OFDM system with $N = 1024$ carriers, which are all used as data carriers. The guard interval consists of $\nu = 100$ samples. The channel has a similar impulse response as the channel in the previous section, but now the number of paths L is equal to 50. The considered channel code is a turbo code that consists of the parallel concatenation of two identical 16-state rate-1/2 recursive systematic convolutional codes with generator sequences $(37)_8$ and $(21)_8$ (in octal notation) through a uniform interleaver of length $N_\mathrm{b} = 1019$ bits. The parity bits are punctured so that the length of 1 code word N_c is equal to 2048 bits. The considered symbol constellation is a 4-QAM constellation, which means that 1 code word is mapped on $N_\mathrm{d} = 1024$ symbols, or in other words, 1 OFDM block contains exactly 1 code word.

Due to the high computational complexity of the optimal detector, it is not possible to show some numerical results for this detector. Only the results for the ZF and the frequency detectors are shown. To illustrate the error correcting capabilities of the considered turbo code, we also show the results for an uncoded system. The error correcting code results in a large gain, regardless which detection method is applied. These results illustrate that the approximations made about the noise to obtain the likelihoods of the different data symbols in 4.1.2 4.2, are valid. In contrast with the uncoded case, there is not much difference between the ZF detector and the FD detector for the considered range of E_b/N_0 values. There is a small degradation of less than 0.5 dB though for the FD detector.

4.4 Chapter Summary

In this chapter we have discussed several data detection algorithms. First we have shown that the optimal detection algorithm, i.e. MLSD, has a very high computational complexity. To tackle this problem we have proposed some alternatives. The first candidate was the ZF detector, which is a linear receiver that operates in the time domain. Compared to the MLSD, the ZF combines a lowered computational complexity with some loss in performance, although it still might be too complex. So, we have continued our search for low complexity alternatives. By transforming the received signal to the frequency domain, the multipath interference is removed and symbol-by-symbol detection can be performed. This comes at a price, though: the available multipath diversity can not be exploited, while the MLSD has full diversity. There is a trade off between complexity and performance.

For a coded transmission, the results show that the ZF detector and FD detector are valuable alternatives for the computationally prohibitive optimal detector. The difference in performance between the ZF detector and the FD

detector is not as explicit as in the uncoded case. Although there is a small gap between the BER performances of both detectors, with the ZF detector yielding the better performance.

4.A Diversity of the ML Sequence Detector

In this appendix we are going to compute the diversity order of the ML sequence detector. In terms of the BER, the diversity order is defined as [3] the increase in the slope of the BER versus E_b/N_0 curve, compared to the BER curve of a system with no diversity where the BER is assymptotically inversely proportional to E_b/N_0. The BER is upper bounded by the union bound [2, 3] based on the pairwise error probability (PEP). So the diversity of the MLSD can be determined by computing the (worst case) PEP between two arbitrary symbol vectors. The probability that the likelihood of the symbol vector $\mathbf{a}_0^{'}$ is larger than the likelihood of the symbol vector $\mathbf{a}_0$, when the symbol vector $\mathbf{a}_0$ is originally transmitted and all the symbol vectors are equiprobable, is given by

$$p\left(\mathbf{a}_0 \rightarrow \mathbf{a}_0^{'} \middle| \mathbf{h}\right) = Q\left(\sqrt{\frac{N}{N+\nu}\frac{1}{2N_0}\left\|\mathbf{H}\mathbf{F}^H\left(\mathbf{a}_0 - \mathbf{a}_0^{'}\right)\right\|^2}\right) \tag{4.30}$$

where the function $Q(x)$ is defined as [2]

$$Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty \exp\left(-\frac{u^2}{2}\right) du. \tag{4.31}$$

This result still depends on the channel impulse response. In its current form, it is not that easy to average the PEP over the channel parameters. Therefore we are going to rewrite the argument of the Q-function and replace the Q-function by an appropriate upper bound. First we define $\bar{\mathbf{s}}_\mathrm{d}$ as the IFFT of the difference between the symbol vectors $\mathbf{a}_0$ and $\mathbf{a}_0^{'}$, i.e. $\bar{\mathbf{s}}_\mathrm{d} = \mathbf{F}^H\left(\mathbf{a}_0 - \mathbf{a}_0^{'}\right)$, and the $(N+\nu)\times L$ Toeplitz matrix $\mathbf{S}$ as

$$(\mathbf{S})_{l:l+N-1,l} = \frac{1}{\sqrt{E_\mathrm{s}}}\bar{\mathbf{s}}_\mathrm{d} \quad l = 0,\ \ldots,\ L-1. \tag{4.32}$$

The samples from the channel impulse $\mathbf{h}$ response are given by (2.25), which yields in vector format

$$\mathbf{h} = \mathbf{G}\tilde{\boldsymbol{h}} \tag{4.33}$$

where the vector $\tilde{\boldsymbol{h}}$ collects the complex path gains $\tilde{h}_l$, $l = 0,\ \ldots, \tilde{L}-1$ and the $L\times\tilde{L}$ matrix $\mathbf{G}$ is defined as

$$(\mathbf{G})_{k,l} = g\left(kT - \tau_l\right). \tag{4.34}$$

With these definitions the PEP can be expressed as

$$p\left(\mathbf{a}_0 \rightarrow \mathbf{a}_0^{'} \middle| \mathbf{h}\right) = Q\left(\sqrt{\frac{1}{2}\frac{N}{N+\nu}\frac{E_\mathrm{s}}{N_0}\tilde{\boldsymbol{h}}^H\mathbf{G}^H\mathbf{S}^H\mathbf{S}\mathbf{G}\tilde{\boldsymbol{h}}}\right). \tag{4.35}$$

The Q-function can be upper bounded by its Chernoff bound [2] :

$$Q(x) \leq \frac{1}{2} e^{-\frac{x^2}{2}}, \tag{4.36}$$

which results in the following upper bound for the PEP from (4.30):

$$p\left(\mathbf{a}_0 \rightarrow \mathbf{a}_0^{'} \middle| \mathbf{h}\right) \leq \frac{1}{2} \exp\left(-\frac{N}{N+\nu}\frac{E_s}{N_0}\frac{\tilde{\boldsymbol{h}}^H \mathbf{G}^H \mathbf{S}^H \mathbf{S} \mathbf{G} \tilde{\boldsymbol{h}}}{4}\right). \tag{4.37}$$

We assume that the path delays are constant so that we only have to average over the complex path gains. The components of the vector $\tilde{\boldsymbol{h}}$ are i.i.d. complex Gaussian random variables with zero mean and variance equal to $1/\tilde{L}$. To simplify things, we express the vector $\tilde{\boldsymbol{h}}$ as a unitary transformation of the vector $\tilde{\boldsymbol{h}}'$:

$$\tilde{\boldsymbol{h}} = \mathbf{V} \tilde{\boldsymbol{h}}' \tag{4.38}$$

where $\mathbf{V}$ is a unitary matrix. The components of the vector $\tilde{\boldsymbol{h}}'$ have the same statistical properties as the components of $\tilde{\boldsymbol{h}}$. Now we still have to make an appropriate choice for $\mathbf{V}$. The matrix $\mathbf{G}^H \mathbf{S}^H \mathbf{S} \mathbf{G}$ can be expressed as its eigendecomposition:

$$\mathbf{G}^H \mathbf{S}^H \mathbf{S} \mathbf{G} = \mathbf{U} \mathbf{Y} \mathbf{U}^{\boldsymbol{H}} \tag{4.39}$$

where $\mathbf{U}$ collects the eigenvectors and $\mathbf{Y}$ is a diagonal matrix with the eigenvalues v_l, $l = 0, \ldots, \tilde{L}-1$, on its diagonal. Because $\mathbf{G}^H \mathbf{S}^H \mathbf{S} \mathbf{G}$ is a Hermitian matrix, the matrix $\mathbf{U}$ is a unitary matrix so we choose $\mathbf{V}$ equal to $\mathbf{U}$. Substituting $\tilde{\boldsymbol{h}}$ by (4.38) in (4.37) results in

$$\begin{aligned} p\left(\mathbf{a}_0 \rightarrow \mathbf{a}_0^{'} \middle| \mathbf{h}\right) &\leq \frac{1}{2} \exp\left(-\frac{N}{N+\nu}\frac{E_s}{N_0}\frac{\tilde{\boldsymbol{h}}'^H \mathbf{U}^H \mathbf{G}^H \mathbf{S}^H \mathbf{S} \mathbf{G} \mathbf{U} \tilde{\boldsymbol{h}}'}{4}\right) \\ &= \frac{1}{2} \exp\left(-\frac{N}{N+\nu}\frac{E_s}{N_0}\frac{\tilde{\boldsymbol{h}}'^H \mathbf{Y} \tilde{\boldsymbol{h}}'}{4}\right) \end{aligned} \tag{4.40}$$

and averaging this result over the distribution of $\tilde{\boldsymbol{h}}'$ yields

$$\begin{aligned} p\left(\mathbf{a}_0 \rightarrow \mathbf{a}_0^{'}\right) &\leq \frac{\tilde{L}^{\tilde{L}}}{\pi^{\tilde{L}}} \int_{\tilde{\boldsymbol{h}}'} \frac{1}{2} \exp\left[-\tilde{\boldsymbol{h}}'^H \left(\frac{N}{N+\nu}\frac{E_s}{N_0}\frac{\mathbf{Y}}{4} + \tilde{L}\mathbf{I}_{\tilde{L}}\right) \tilde{\boldsymbol{h}}'\right] d\tilde{\boldsymbol{h}}' \\ &= \frac{1}{2}\tilde{L}^{\tilde{L}} \prod_{l=0}^{\tilde{L}-1} \int_{\tilde{h}_l'} \frac{1}{\pi}\frac{\sigma_l^2}{\sigma_l^2} \exp\left(-\frac{\left|\tilde{h}_l'\right|^2}{\sigma_l^2}\right) d\tilde{h}_l' \\ &= \frac{1}{2}\tilde{L}^{\tilde{L}} \prod_{l=0}^{\tilde{L}-1} \sigma_l^2 \\ &= \frac{1}{2} \prod_{l=0}^{\tilde{L}-1} \left(1 + \frac{N}{N+\nu}\frac{E_s}{N_0}\frac{v_l}{4\tilde{L}}\right)^{-1}, \end{aligned} \tag{4.41}$$

where $\sigma_l^2 = \left(\tilde{L} + \frac{N}{N+\nu}\frac{E_s}{N_0}\frac{v_l}{4}\right)^{-1}$.

The diversity is then given by the total number of non zero eigenvalues v_l, which corresponds to the rank of the matrix $\mathbf{G}^H\mathbf{S}^H\mathbf{S}\mathbf{G}$. The rank of the matrix $\mathbf{G}$ is equal to $\tilde{L}$ while the rank of the matrix $\mathbf{S}$ is always equal to L no matter which symbol vectors $\mathbf{a}_0$ and $\mathbf{a}_0^{'}$ we consider. Since the number of resolvable paths $\tilde{L}$ is always smaller than L, we can conclude that the matrix $\mathbf{G}^H\mathbf{S}^H\mathbf{S}\mathbf{G}$ has full rank and the diversity of the MLSD is equal to $\tilde{L}$.

5

Parameter Estimation: General Concepts

In this chapter we will give a short introduction to the general concepts about parameter estimation. Instead of averaging the probability function over the unknown parameters, which is in general a complex operation, the receiver makes an estimate of the unknown parameters and replaces the unknown parameters in the probability function by the obtained estimates, assuming that those estimates are equal to the true values of the unknown parameters.

First of all, data aided parameter estimation is mentioned. An important aspect of data aided estimation algorithms is the insertion of pilot symbols in the data stream which has to be transmitted. We show some examples of possible pilot transmission schemes.

When the possibilities of the pilot symbols are exhausted to obtain estimates of the unknown parameters, one can resort to decision aided estimation algorithms. This class of algorithms consists of two subcategories. The first category consists of the hard decision aided algorithms. In this case, after an initial estimation of the unknown parameters, the detected data symbols

are considered as known and are used as extra pilot symbols to improve the estimates of the unknown parameters. With these new 'pilot symbols', usually more accurate estimates of the unknown parameters can be obtained. A second class of decision aided estimation algorithms, consists of estimation algorithms that use some soft information about the unknown data symbols.

5.1 Pilot Aided Estimation

A first class of estimation algorithms makes use of known training symbols which are inserted in the data stream that has to be transmitted. Those algorithms are called pilot aided estimation algorithms. In a KSP-OFDM system, a part of the training symbols is transmitted during the guard interval between two adjacent OFDM blocks, i.e. the time domain pilots and the other part is transmitted on some of the data carriers, i.e. the frequency domain pilots. The receiver knows which pilot symbols have been transmitted by the transmitter and tries to exploit this knowledge to extract an estimate of the unknown parameters. This type of algorithm is usually easy to implement and exhibits good performance provided that the number of transmitted pilot symbols is high enough and that the placement of the pilots in the frequency domain is well thought-out.

The downside of this approach is the fact that the transmission of the pilot symbols costs bandwidth and consumes power. The main purpose of a communication system is to transmit unknown data symbols to the receiver, so we want to keep the part of the available resources that is spent on the transmission of pilot symbols as low as possible. Especially in low SNR environments providing good parameter estimates might be impossible without wasting to much resources at transmitting pilot symbols.

Since there are also data symbols available, which are not considered for the estimation process, a pilot aided estimation algorithm is not optimal in the sense that it does not make use of all the available information.

5.2 Pilot Symbol Arrangement

The placement of the time domain pilot symbols is straightforward and has been discussed already in chapter 2. After every OFDM block a total number of ν pilot symbols are transmitted.

In the frequency domain we have more possibilities to place the pilot symbols. Usually the number of pilot carriers is much smaller than the total number of carriers, so there are a lot of possibilities to choose which carriers are data carriers and which carriers are designated for the transmission of pilot symbols. Furthermore the pilot symbol arrangement in frequency domain may vary per transmitted OFDM block. In this work, we assume that a total

number of M pilot symbols are transmitted during 1 OFDM block + guard interval, which means that besides the ν pilot symbols in the guard interval, there are also $M-\nu$ pilot symbols which are transmitted on the pilot carriers. The set of carrier indices is divided in two subsets:

$$S_{\mathrm{p}} = \{\alpha_0, \ \ldots, \ \alpha_{M-\nu-1}\}, \tag{5.1}$$

which is the set of pilot carriers, and

$$S_{\mathrm{d}} = \{\beta_0, \ \ldots, \ \beta_{N+\nu-M-1}\}, \tag{5.2}$$

which contains the data carriers.

In this work we consider two possible pilot structures, e.g. the comb-type pilot structure [43, 44] and a random pilot structure. For the comb-type pilot structure, the pilot carrier positions are given by

$$S_{\mathrm{p}} = \left\{\alpha_0 + m\lambda_{\mathrm{pc}} |\ m = 0, \ \ldots, \ M-\nu-1\right\}, \tag{5.3}$$

where λ_{pc} is the distance between two pilot carriers expressed in number of carriers and defined as

$$\lambda_{\mathrm{pc}} = \left\lfloor \frac{N}{M-\nu} \right\rfloor, \tag{5.4}$$

where $\lfloor x \rfloor$ is the floor function x, yielding the largest integer which is smaller than x. The position of the first pilot carrier α_0 is chosen from $\{0, \ \ldots, \ \rho\}$, where ρ is given by

$$\rho = N - 1 - (M-\nu-1)\,\lambda_{\mathrm{pc}}. \tag{5.5}$$

In the random pilot structure, as the name already suggests, the pilot carrier positions are randomly selected. Figure 5.1 illustrates the two considered pilot structures.

We can rewrite the time domain samples of the i-th block $\mathbf{s}_i$ (see also 2.8) as the sum of a vector $\mathbf{s}_{\mathrm{p}}$, which collects the contributions from both the pilot symbols in time and frequency domain, and a vector $\mathbf{s}_{\mathrm{d}}^{(i)}$, which consists of the contributions from the unknown data symbols:

$$\mathbf{s}_i = \mathbf{s}_{\mathrm{p}} + \mathbf{s}_{\mathrm{d}}^{(i)}. \tag{5.6}$$

The contribution of the pilot symbols $\mathbf{s}_{\mathrm{p}}$ is given by:

$$\mathbf{s}_{\mathrm{p}} = \sqrt{\frac{N}{N+\nu}} \begin{pmatrix} \mathbf{F}_{\mathrm{p}}^{H} \mathbf{a}_{\mathrm{c}} \\ \mathbf{a}_{\mathrm{g}} \end{pmatrix}, \tag{5.7}$$

where $\mathbf{F}_{\mathrm{p}}$ is a $(M-\nu) \times N$ matrix which consists of a subset of rows of the FFT matrix $\mathbf{F}$ corresponding to the set S_{p} of pilot carriers. The vector $\mathbf{a}_{\mathrm{c}} = \left(a_{\mathrm{c}}(0), \ \ldots, \ a_{\mathrm{c}}(M-\nu-1)\right)^T$ collects the frequency domain pilot symbols, while the vector $\mathbf{a}_{\mathrm{g}}$ consists of the time domain pilot symbols and is

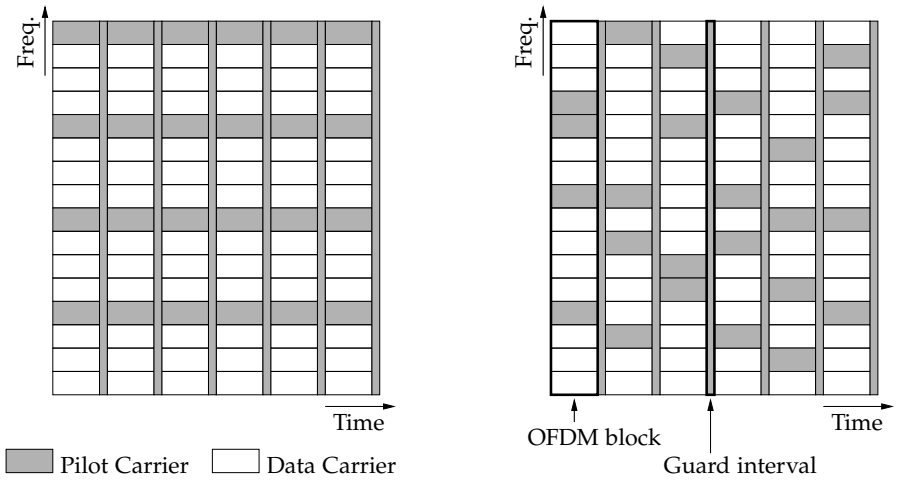

Figure 5.1: The comb-type pilot scheme (left) and the random pilot scheme (right).

defined in subsection 2.1.4. The contribution of the data symbols from the i-th OFDM block $\mathbf{s}_{\mathrm{d}}^{(i)}$ is defined as:

$$\mathbf{s}_{\mathrm{d}}^{(i)} = \sqrt{\frac{N}{N+\nu}} \begin{pmatrix} \mathbf{F}_{\mathrm{d}}^{H} \mathbf{a}_{\mathrm{d}}^{(i)} \\ \mathbf{0}_{\nu \times 1} \end{pmatrix}. \tag{5.8}$$

The matrix $\mathbf{F}_{\mathrm{d}}$ is a $(N - M + \nu) \times N$ matrix consisting of a subset of rows of the FFT matrix $\mathbf{F}$ corresponding to the set S_{d} of data carriers. The vector $\mathbf{a}_{\mathrm{d}}^{(i)} = \left(a_{\mathrm{d}}^{(i)}(0), \ldots, a_{\mathrm{d}}^{(i)}(N - M + \nu - 1)\right)$ contains the data symbols from the i-th OFDM block.

5.3 Decision Aided Estimation

In general, a decision aided estimation algorithm is an algorithm that exploits some information about the unknown data symbols from the detector or the decoder to perform the estimation process [45, 46]. It is expected that the estimator can benefit from the symbol information provided by the symbol detector or decoder.

A decision aided estimation algorithm is an iterative algorithm that usually starts from an initial estimate of the parameters obtained from a pilot aided initialization algorithm. This initial estimate is then used by the symbol detector / decoder to obtain some information about the unknown data symbols. The decision aided algorithm iterates between improving the parameter

estimation exploiting the pilot symbols and the data symbol information provided by the detector / decoder, and updating the data symbol information by applying the improved parameter estimates. This procedure is repeated several times until convergence has reached.

Now the question remains which information about the unknown data symbols is fed back to the parameter estimation unit. There exist several possibilities. The most straightforward decision aided estimation algorithm uses the hard decisions on the data symbols which are provided by the detector / decoder and assumes that they are correct. So the data symbols are assumed to be known by the estimation unit and are considered as extra pilot symbols. The extension of the pilot aided estimation algorithms that are applied to obtain the initial parameter estimates, to estimation algorithms that also consider the unknown data symbols, is straightforward.

In a second class of decision aided estimation algorithms the detector / decoder does not feed the estimation unit with hard decisions on the data symbols, but provides a measure about the symbol knowledge after detection / decoding [46]. This measure can be for example the a posteriori probabilities of the data symbols. This information about the unknown data symbols is called soft information. The EM algorithm which estimates the unknown parameters and considers the unknown data symbols as nuisance parameters, is a well known decision aided estimation algorithm which makes use of soft information.

5.3.1 Practical Implementation for Coded Systems: Embedded Estimation

Every iteration of the estimation process, the marginal a posteriori probabilities of the data symbols need to be computed to obtain the necessary hard decisions or soft information. When the computation of these marginal a posteriori probabilities is performed according to the sum product algorithm, as is the case with turbo codes and LDPC codes, then this operation itself is already iterative. So in theory, for every iteration of the estimator, the decoder needs to perform enough decoding iterations until convergence for the a posteriori probabilities is reached. This might result in a high computational complexity and makes it almost impossible to implement [47, 48].

To keep the number of computations tractable, we resort to the concept of embedded estimation [31, 39, 47, 48]. After every update of the estimates of the unknown parameters, the decoder performs only a single iteration, but the state information of the decoder is not reset but kept for the next iteration. So after every application of the estimation process, the detector is not reinitialized but starts from the state information from the previous step. The iterations from the estimator and the iterations from the sum product algorithm are intertwined, or in other words, the estimation process is embedded

in the decoding process. By doing so, the computational load is considerably reduced.

Note that performing only one decoding iteration at each estimation iteration might result in very coarse approximations of the actual marginal symbol a posteriori probabilities, particularly in the first estimation iterations. Furthermore, the decoder is not reinitialized every iteration, which means that the computed a posteriori probabilities depend on all the parameter estimates from previous iterations. This means that when the EM algorithm is embedded in the decoding process, it is no longer an EM algorithm, because the EM algorithm requires that the a posteriori probabilities used in the E-step, only depend on the parameter estimates from the M-step from the previous estimation iteration [31].

5.4 Conclusion

In this chapter, we have discussed some general approaches to perform parameter estimation. The most conventional approach is to obtain estimates based on the transmission of pilot symbols. In a KSP-OFDM system, the pilot symbols consist of time domain pilot symbols which are transmitted during the guard interval and frequency domain pilot which are transmitted on the pilot carriers. We have mentioned two possible pilot symbol structures: the comb-type pilot structure and the random pilot structure. The major drawbacks of data aided estimation are the inefficient use of energy and bandwidth.

To improve the quality of the obtained estimates, more advanced algorithms exist which exploit information about the unknown data symbols provided by the detector / decoder. They form the class of decision aided estimation algorithms. The decision aided algorithms typically iterate between a parameter estimation process and a detection / decoding process. For the more advanced coding schemes (turbo codes, LDPC codes, ...), the decoding process itself is already an iterative algorithm. To reduce the computational complexity of the decision aided estimation algorithms, the concept of embedded estimation is introduced. The estimation iterations and the decoding iterations are intertwined: after every estimation iteration, only one decoding operation is performed, while saving the state information of the decoder for the next iteration.

6

Channel Estimation

In this chapter we investigate several algorithms to estimate the channel impulse response. After the problem statement, we start with the CRB, which gives us a benchmark for the achievable performance in terms of the MSE. Secondly, we review the state of the art by discussing some existing channel estimation methods which can be applied for a KSP-OFDM system.

In the remaining part of this chapter we introduce our proposed channel estimators, starting with a pilot aided channel estimator which operates in the frequency domain. The quality of the estimate is then further improved by applying an iterative decision aided estimator. First hard decision aided estimation is introduced both in the time domain and in the frequency domain. Then a soft decision aided estimator based on the EM algorithm is proposed.

Finally the different estimators are compared with the CRB and the state-of the-art by means of some numerical results for the MSE and the BER.

6.1 Problem Statement

We consider a KSP-OFDM system with N carriers and a guard interval of length ν. Besides the time domain pilot symbols, there are also $M - \nu$ pilot carriers available, so per OFDM block + guard interval, there are M pilot symbols transmitted.

The OFDM blocks are transmitted over a frequency selective channel with an impulse response $\mathbf{h}$, consisting of L taps. We assume that the timing offset and frequency offset are perfectly known and compensated. The $N + \nu$ received signal samples corresponding to one block are given by (4.1):

$$\mathbf{z} = \mathbf{H}_{\text{ch}}\mathbf{s} + \mathbf{w}, \tag{6.1}$$

where we have dropped the block index for notational convenience. For the data detection, the receiver needs to know the matrix $\mathbf{H}_{\text{ch}}$. Since the components of $\mathbf{H}_{\text{ch}}$ are all samples from the channel impulse response $\mathbf{h}$, it suffices to estimate $\mathbf{h}$.

First of all, the received signal vector $\mathbf{z}$ must be rewritten in a more convenient form:

$$\begin{aligned} \mathbf{z} &= \mathbf{H}_{\text{ch}}\left(\mathbf{s}_{\text{p}} + \mathbf{s}_{\text{d}}\right) + \mathbf{w} \\ &= \mathbf{B}\mathbf{h} + \mathbf{H}\mathbf{F}_{\text{d}}^{H}\mathbf{a}_{\text{d}} + \mathbf{w}, \end{aligned} \tag{6.2}$$

where the time domain pilot vector $\mathbf{s}_{\text{p}}$ and the time domain data vector $\mathbf{s}_{\text{d}}$ are given by (5.7) and (5.8) respectively. The $(N + \nu) \times L$ pilot matrix $\mathbf{B}$ is defined as:

$$(\mathbf{B})_{k,l} = s_{\text{p}}\left(|k - l|_{N+\nu}\right). \tag{6.3}$$

Now, the problem consists of obtaining an estimate of $\mathbf{h}$, based on the received signal $\mathbf{z}$ (6.2), which still depends on the unknown data symbols $\mathbf{a}_{\text{d}}$.

In the literature, several channel estimation techniques have been proposed [27,49–52]. It turns out that channel estimation in KSP-OFDM is harder than in CP-OFDM and ZP-OFDM, because of the combination of frequency-domain pilots and time-domain pilots. ML channel estimation in KSP-OFDM is very complex and suboptimal estimation techniques must be used. The algorithms from [27,49,50] assume that the unknown data symbols are Gaussian distributed. In [27], a suboptimal ML-based channel estimation algorithm is proposed. However, [27] assumes that the autocorrelation matrix of the disturbance (containing contributions from the noise, the data symbols and the channel) is known. Hence, before this channel estimator can be used, first the autocorrelation matrix must be estimated from the received signal. Further, even if the autocorrelation matrix is perfectly known, the resulting MSE shows an error floor at high SNR, indicating that the presence of the unknown data symbols disturbs the channel estimation. In [50] channel estimation for KSP transmission over stationary frequency-selective channels is considered,

which is a special case of the situation considered in [27]. In [51], the signals transmitted on the pilot carriers are selected such that the last IFFT outputs correspond to the pilot symbols from the KSP sequence. Before the OFDM block, the same samples are transmitted, so we can consider this situation as a special case of CP-OFDM. An equalizer is then trained based on the received KSP sequences of several consecutive blocks. The transmitted signals on the pilot carriers also contain a contribution of the unknown data symbols, so only the KSP sequence in the time domain can be used to perform synchronization. For TDS-OFDM, several channel estimators are proposed, e.g. [53–55]. Since there are no pilot carriers available in a TDS-OFDM system, it is hard to perfectly separate the pilot symbol contributions from the data symbol contributions. As a result, the existing algorithms all estimate the channel impulse response in an iterative way, where an initial estimate of the channel impulse response is obtained by exploiting the presence of the pseudo noise sequence in the guard interval.

In the next sections we will introduce some possible solutions to tackle the presence of the unknown data symbols. First of all, we will start by deriving a CRB, which serves as a lower bound for the MSE of the proposed channel estimators.

6.2 Cramer-Rao Bound

This part is based on the paper from Steendam et al. [52]. The presence of the unknown data symbols in (6.2) makes it not straightforward to obtain the CRB of the estimation of $\mathbf{h}$. Therefore, we assume that the vector $\mathbf{F}_{\mathrm{d}}^{H}\mathbf{a}_{\mathrm{d}}$ is Gaussian distributed with zero mean. This approximation is good when the number of data carriers $N+\nu-M$ is large. Typically, the number of pilot carriers is small, so we can say that the approximation is especially reasonable for large block sizes. The obtained CRB is called the Gaussian CRB (GCRB). It has been shown (see [56]) that the GCRB is the worst case CRB and so the GCRB is an upper bound for the CRB which considers the true distribution of $\mathbf{F}_{\mathrm{d}}^{H}\mathbf{a}_{\mathrm{d}}$. We define the vector $\bar{\mathbf{w}}$ as the total noise vector:

$$\bar{\mathbf{w}} = \mathbf{H}\mathbf{F}_{\mathrm{d}}^{H}\mathbf{a}_{\mathrm{d}} + \mathbf{w} \tag{6.4}$$

We define the vector $\boldsymbol{\theta}$ as

$$\boldsymbol{\theta} = \begin{pmatrix} \Re\{\mathbf{h}\} \\ \Im\{\mathbf{h}\} \end{pmatrix}. \tag{6.5}$$

The Fisher information matrix for $\boldsymbol{\theta}$ is given by (3.10):

$$\mathbf{J}_{\theta} = \mathbb{E}_{\mathbf{z}|\boldsymbol{\theta}}\left[\left(\frac{\partial}{\partial\boldsymbol{\theta}}\ln p\left(\mathbf{z}|\mathbf{h}\right)\right)\left(\frac{\partial}{\partial\boldsymbol{\theta}}\ln p\left(\mathbf{z}|\mathbf{h}\right)\right)^{T}\right]. \tag{6.6}$$

We assume that $\mathbf{s}_\mathrm{d} = \mathbf{F}_\mathrm{d}^H \mathbf{a}_\mathrm{d}$ is zero-mean Gaussian distributed, so $\mathbf{z}$ given $\mathbf{h}$ is Gaussian distributed with autocorrelation matrix given by

$$\mathbf{R}_{\tilde{\mathbf{w}}} = E_\mathrm{s} \frac{N}{N+\nu} \mathbf{H} \mathbf{F}_\mathrm{d}^H \mathbf{F}_\mathrm{d} \mathbf{H}^H + N_0 \mathbf{I}_{N+\nu}. \tag{6.7}$$

Hence the log likelihood function of $\mathbf{h}$ is equal to

$$\ln p\left(\mathbf{z}|\mathbf{h}\right) = C - \frac{1}{2} \ln \det \mathbf{R}_{\tilde{\mathbf{w}}} - \left(\mathbf{z} - \mathbf{B}\mathbf{h}\right)^H \mathbf{R}_{\tilde{\mathbf{w}}}^{-1} \left(\mathbf{z} - \mathbf{B}\mathbf{h}\right). \tag{6.8}$$

The autocorrelation matrix $\mathbf{R}_{\tilde{\mathbf{w}}}$ depends on the components of the parameter vector $\boldsymbol{\theta}$, so we need the derivatives of $\det \mathbf{R}_{\tilde{\mathbf{w}}}$ and $\mathbf{R}_{\tilde{\mathbf{w}}}^{-1}$ with respect to $\Re\{\mathbf{h}\}$ and $\Im\{\mathbf{h}\}$ to obtain the Fisher information matrix. These derivatives are in general difficult to obtain, so the computation of the GCRB is in general very complex. To simplify things, the contribution of the data symbols $\mathbf{H}\mathbf{F}_\mathrm{d}^H \mathbf{a}_\mathrm{d}$ can be approximated by $\tilde{\mathbf{F}}\tilde{\mathbf{H}}\mathbf{a}_\mathrm{d}$, where the $(N+\nu) \times (N+\nu-M)$ matrix $\tilde{\mathbf{F}}$ is given by:

$$\left(\tilde{\mathbf{F}}\right)_{k,n} = \frac{1}{\sqrt{N}} e^{j2\pi \frac{k\beta_n}{N}},\ \beta_n \in S_\mathrm{d} \tag{6.9}$$

where S_d is the subset of data carriers (5.2), $\tilde{\mathbf{H}}$ is a diagonal matrix with diagonal elements $H\left(\beta_n\right)$:

$$H\left(\beta_n\right) = \sum_{l=0}^{L-1} h\left(l\right) e^{-j2\pi \frac{l\beta_n}{N}}. \tag{6.10}$$

Basically, this approximation comes down to neglecting the transient of the contribution from $\mathbf{a}_d$ at the edges of the received block $\mathbf{z}$. Especially for large blocks when the duration of an OFDM block is much larger than the guard interval, i.e. $N \gg \nu$, this approximation is valid.

Note that $\tilde{\mathbf{H}}\mathbf{a}_\mathrm{d}$ contains only $N+\nu-M < N+\nu$ components, so it is possible to find an invertible linear transform $\mathbf{T}$ that transforms $\mathbf{z}$ into an $(N+\nu) \times 1$ vector $\mathbf{z}' = \left[\mathbf{z}_1^T\ \mathbf{z}_2^T\right]^T$, where $\mathbf{z}_1$ is a $(N+\nu-M) \times 1$ vector which depends on the transmitted data symbols and $\mathbf{z}_2$ is a $M \times 1$ vector which is independent of $\mathbf{a}_\mathrm{d}$. Furthermore, an invertible linear transform of the observation vector $\mathbf{z}$ does not affect the GCRB. The transform can be found by performing a QR decomposition of the matrix $\tilde{\mathbf{F}}$. The QR decomposition of a $(N+\nu) \times (N+\nu-M)$ matrix factors the considered matrix as a product of a unitary $(N+\nu) \times (N+\nu)$ matrix and an upper triangular $(N+\nu) \times (N+\nu-M)$ matrix of which the M bottom rows consist entirely of zeros. The matrix $\tilde{\mathbf{F}}$ can be written as

$$\tilde{\mathbf{F}} = \mathbf{Q}\mathbf{V} = \mathbf{Q} \begin{pmatrix} \mathbf{U} \\ \mathbf{0}_{M \times (N+\nu-M)} \end{pmatrix}, \tag{6.11}$$

where $\mathbf{Q}$ is a unitary $(N+\nu) \times (N+\nu)$ matrix and $\mathbf{U}$ is a $(N+\nu-M) \times (N+\nu-M)$ triangular matrix. The matrix of the desired linear transform is

now given by $\mathbf{T} = \mathbf{Q}^{-1} = \mathbf{Q}^H$. Applying this transform to the observation vector $\mathbf{z}$ yields

$$\mathbf{z}' = \mathbf{T}\mathbf{z} = \begin{pmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{pmatrix} \mathbf{h} + \begin{pmatrix} \mathbf{U} \\ \mathbf{0} \end{pmatrix} \tilde{\mathbf{H}}\mathbf{a}_d + \begin{pmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \end{pmatrix}. \tag{6.12}$$

The noise samples $\left(\mathbf{w}_1^T,\ \mathbf{w}_2^T\right)^T = \mathbf{T}\mathbf{w}$ have the same mean and variance as the original noise samples $\mathbf{w}$, because $\mathbf{T}$ is a unitary matrix. The transformation of the total noise vector $\bar{\mathbf{w}}$ is denoted as $\mathbf{w}'$ and has an autocorrelation matrix $\mathbf{R}_{\mathbf{w}'}$ which is given by

$$\mathbf{R}_{\mathbf{w}'} = \mathbf{T}\mathbf{R}_{\bar{\mathbf{w}}}\mathbf{T}^H = \begin{pmatrix} \mathbf{R}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_2 \end{pmatrix} \tag{6.13}$$

where $\mathbf{R}_1$ and $\mathbf{R_2}$ are respectively defined as

$$\mathbf{R}_1 = E_\mathrm{s}\frac{N}{N+\nu}\mathbf{U}\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H\mathbf{U}^H + N_0\mathbf{I}_{N+\nu-M} \tag{6.14}$$

$$\mathbf{R}_2 = N_0\mathbf{I}_M. \tag{6.15}$$

Since the autocorrelation matrix $\mathbf{R}_{\mathbf{w}'}$ is a block diagonal matrix, the observation vectors $\mathbf{z}_1$ and $\mathbf{z}_2$ given $\mathbf{h}$ are statistically independent. As a result, the Fischer information matrix is given by $\mathbf{J}_{\boldsymbol{\theta}} = \mathbf{J}_1 + \mathbf{J}_2$, where (with $i = 1,\ 2$)

$$\mathbf{J}_i = \mathbb{E}_{\mathbf{z}_i|\boldsymbol{\theta}}\left[\left(\frac{\partial}{\partial\boldsymbol{\theta}}\ln p\left(\mathbf{z}_i|\mathbf{h}\right)\right)\left(\frac{\partial}{\partial\boldsymbol{\theta}}\ln p\left(\mathbf{z}_i|\mathbf{h}\right)\right)^T\right], \tag{6.16}$$

$$\ln p\left(\mathbf{z}_i|\mathbf{h}\right) = C - \frac{1}{2}\ln\det\mathbf{R}_i - \left(\mathbf{r}_i - \mathbf{B}_i\mathbf{h}\right)^H\mathbf{R}_i^{-1}\left(\mathbf{r}_i - \mathbf{B}_i\mathbf{h}\right). \tag{6.17}$$

The determination of $\mathbf{J}_1$ is based on the observation $\mathbf{z}_1$ only which is given by

$$\mathbf{z}_1 = \mathbf{B}_1\mathbf{h} + \mathbf{U}\tilde{\mathbf{H}}\mathbf{a}_\mathrm{d} + \mathbf{w}_1. \tag{6.18}$$

One notices immediately that the autocorrelation matrix $\mathbf{R}_1$ of the contribution of the data symbols and the noise vector depends on $\mathbf{h}$. As a result, we need the derivatives of $\det\mathbf{R}_1$ and $\left(\mathbf{R}_1\right)^{-1}$ with respect to $\mathbf{h}$. We start with expressing $\mathbf{R}_1$ in a more convenient form, using the following approximation: when the number of pilot carriers is much smaller than the total number of carriers, i.e. $M - \nu \ll N$, the matrix product $\tilde{\mathbf{F}}\tilde{\mathbf{F}}^H$ can be approximated by the identity matrix $\mathbf{I}_{N+\nu}$. Under this assumption, $\mathbf{R_1}$ can be rewritten as

$$\mathbf{R}_1 = \mathbf{T}_1\tilde{\mathbf{F}}\boldsymbol{\Delta}\tilde{\mathbf{F}}^H\mathbf{T}_1^H, \tag{6.19}$$

where $\mathbf{T_1}$ is given by the first $N + \nu - M$ of the linear transformation matrix $\mathbf{T}$, and $\boldsymbol{\Delta}$ is a diagonal matrix with elements λ_n defined as

$$\lambda_n = N_0 + \frac{N}{N+\nu}E_\mathrm{s}\left|H\left(\beta_n\right)\right|^2,\ \beta_n \in S_\mathrm{d}. \tag{6.20}$$

The matrix $\mathbf{T}_1\tilde{\mathbf{F}}$ is a square matrix, so $\mathbf{R}_1$ is the product of square matrices. The determinant of the product of three square matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ is given by $\det \mathbf{ABC} = \det \mathbf{A} \cdot \det \mathbf{B} \cdot \det \mathbf{C}$. Using this property, $\ln \det \mathbf{R}_1$ can be expressed as

$$\ln \det \mathbf{R}_1 = \ln \det \left(\mathbf{T}_1 \tilde{\mathbf{F}} \tilde{\mathbf{F}}^H \mathbf{T}_1^H \right) + \sum_{\beta_n \in S_d} \ln \lambda_n. \tag{6.21}$$

The matrix $\tilde{\mathbf{F}}$ has a rank equal to $N + \nu - M$, so the product $\mathbf{T}_1\tilde{\mathbf{F}}$ is a full-rank matrix. Using this last result, the inverse of $\mathbf{R}_1$ is computed as

$$(\mathbf{R}_1)^{-1} = \left(\tilde{\mathbf{F}}^H \mathbf{T}_1^H \right)^{-1} \mathbf{\Delta}^{-1} \left(\mathbf{T}_1 \tilde{\mathbf{F}} \right)^{-1}. \tag{6.22}$$

The derivatives of $\ln \det \mathbf{R}_1$ and $(\mathbf{R}_1)^{-1}$ are easily obtained using (6.21) and (6.22). For the interested reader, more details about the computation of these derivatives and the final expressions for the Fisher information matrix can be found in appendix 6.A. The resulting Fisher information matrix $\mathbf{J}_1$ can be expressed as a block matrix containing 4 $L \times L$ parts:

$$\mathbf{J}_1 = \begin{pmatrix} \mathbf{J}_{1,a} & \mathbf{J}_{1,b} \\ \mathbf{J}_{1,b}^T & \mathbf{J}_{1,c} \end{pmatrix}, \tag{6.23}$$

where the parts $\mathbf{J}_{1,a}$, $\mathbf{J}_{1,b}$ and $\mathbf{J}_{1,c}$ are given by

$$\begin{aligned} (\mathbf{J}_{1,a})_{l,l'} &= 2\Re \left\{ \left(\mathbf{B}_1^H \mathbf{R}_1^{-1} \mathbf{B}_1 \right)_{l,l'} \right\} \\ &\quad + 4 \left(\Re \{\eta_l\} \Re \{\eta_{l'}\} + \sum_{\beta_n \in S_\mathrm{d}} \frac{\Re \{\gamma_{l,n}\} \Re \left\{ \gamma_{l',n} \right\}}{\lambda_n^2} \right) \end{aligned} \tag{6.24}$$

$$\begin{aligned} (\mathbf{J}_{1,b})_{l,l'} &= -2\Im \left\{ \left(\mathbf{B}_1^H \mathbf{R}_1^{-1} \mathbf{B}_1 \right)_{l,l'} \right\} \\ &\quad - 4 \left(\Re \{\eta_l\} \Im \{\eta_{l'}\} + \sum_{\beta_n \in S_\mathrm{d}} \frac{\Re \{\gamma_{l,n}\} \Im \left\{ \gamma_{l',n} \right\}}{\lambda_n^2} \right) \end{aligned} \tag{6.25}$$

$$\begin{aligned} (\mathbf{J}_{1,c})_{l,l'} &= 2\Re \left\{ \left(\mathbf{B}_1^H \mathbf{R}_1^{-1} \mathbf{B}_1 \right)_{l,l'} \right\} \\ &\quad + 4 \left(\Im \{\eta_l\} \Im \{\eta_{l'}\} + \sum_{\beta_n \in S_\mathrm{d}} \frac{\Im \{\gamma_{l,n}\} \Im \left\{ \gamma_{l',n} \right\}}{\lambda_n^2} \right). \end{aligned} \tag{6.26}$$

The parameters $\gamma_{l,n}$ and η_l are defined as

$$\gamma_{l,n} = \frac{N}{N+\nu} E_\mathrm{s} H^* (\beta_n) \, e^{-j2\pi \frac{l\beta_n}{N}} \tag{6.27}$$

$$\eta_l = -\frac{1}{2} \sum_{\beta_n \in S_\mathrm{d}} \frac{\gamma_{l,n}}{\lambda_n}. \tag{6.28}$$

The computation of $\mathbf{J}_2$ is based only on the observation vector $\mathbf{z}_2$, which is given by

$$\mathbf{z}_2 = \mathbf{B}_2\mathbf{h} + \mathbf{w}_2. \tag{6.29}$$

Obtaining $\mathbf{J}_2$ is much more straightforward as the determination of $\mathbf{J}_1$; the details can be found in appendix 6.B, which yields for the Fisher Information matrix $\mathbf{J}_2$:

$$\mathbf{J_2} = \frac{2}{N_0}\begin{pmatrix} \Re\left\{\mathbf{B}_2^H\mathbf{B}_2\right\} & -\Im\left\{\mathbf{B}_2^H\mathbf{B}_2\right\} \\ \Im\left\{\mathbf{B}_2^H\mathbf{B}_2\right\} & \Re\left\{\mathbf{B}_2^H\mathbf{B}_2\right\} \end{pmatrix}. \tag{6.30}$$

The total Fisher information matrix $\mathbf{J}_{\boldsymbol{\theta}}$, which is based on both the observation vectors $\mathbf{z}_1$ and $\mathbf{z}_2$ is given by

$$\begin{aligned} \mathbf{J}_{\boldsymbol{\theta}} &= \mathbf{J}_1 + \mathbf{J}_2 \\ &= \begin{pmatrix} \mathbf{J}_{\boldsymbol{\theta},a} & \mathbf{J}_{\boldsymbol{\theta},b} \\ \mathbf{J}_{\boldsymbol{\theta},b}^T & \mathbf{J}_{\boldsymbol{\theta},c} \end{pmatrix} \end{aligned} \tag{6.31}$$

where the different parts $\mathbf{J}_{\boldsymbol{\theta},a}$, $\mathbf{J}_{\boldsymbol{\theta},b}$ and $\mathbf{J}_{\boldsymbol{\theta},c}$ are defined as

$$\begin{aligned} (\mathbf{J}_{\boldsymbol{\theta},a})_{l,l'} &= 2\Re\left\{\left(\mathbf{B}^H\mathbf{R}_{\tilde{\mathbf{w}}}^{-1}\mathbf{B}\right)_{l,l'}\right\} \\ &\quad +4\left(\Re\left\{\eta_l\right\}\Re\left\{\eta_{l'}\right\} + \sum_{\beta_n\in S_{\mathrm{d}}}\frac{\Re\left\{\gamma_{l,n}\right\}\Re\left\{\gamma_{l',n}\right\}}{\lambda_n^2}\right) \end{aligned} \tag{6.32}$$

$$\begin{aligned} (\mathbf{J}_{\boldsymbol{\theta},b})_{l,l'} &= -2\Im\left\{\left(\mathbf{B}^H\mathbf{R}_{\tilde{\mathbf{w}}}^{-1}\mathbf{B}\right)_{l,l'}\right\} \\ &\quad -4\left(\Re\left\{\eta_l\right\}\Im\left\{\eta_{l'}\right\} + \sum_{\beta_n\in S_{\mathrm{d}}}\frac{\Re\left\{\gamma_{l,n}\right\}\Im\left\{\gamma_{l',n}\right\}}{\lambda_n^2}\right) \end{aligned} \tag{6.33}$$

$$\begin{aligned} (\mathbf{J}_{\boldsymbol{\theta},c})_{l,l'} &= 2\Re\left\{\left(\mathbf{B}^H\mathbf{R}_{\tilde{\mathbf{w}}}^{-1}\mathbf{B}\right)_{l,l'}\right\} \\ &\quad +4\left(\Im\left\{\eta_l\right\}\Im\left\{\eta_{l'}\right\} + \sum_{\beta_n\in S_{\mathrm{d}}}\frac{\Im\left\{\gamma_{l,n}\right\}\Im\left\{\gamma_{l',n}\right\}}{\lambda_n^2}\right), \end{aligned} \tag{6.34}$$

where we have used the fact that

$$\begin{aligned} \mathbf{B}_1^H\mathbf{R}_1^{-1}\mathbf{B}_1 + \frac{1}{N_0}\mathbf{B}_2^H\mathbf{B}_2 &= \begin{pmatrix} \mathbf{B}_1^H & \mathbf{B}_2^H \end{pmatrix}\mathbf{R}_{\mathbf{w}'}^{-1}\begin{pmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{pmatrix} \\ &= (\mathbf{TB})^H\left(\mathbf{TR}_{\tilde{\mathbf{w}}}\mathbf{T}^H\right)^{-1}\mathbf{TB} \\ &= \mathbf{B}^H\mathbf{R}_{\tilde{\mathbf{w}}}^{-1}\mathbf{B}. \end{aligned} \tag{6.35}$$

The channel impulse response vector is a linear function of the parameter vector $\boldsymbol{\theta}$: $\mathbf{h} = \Re\{\mathbf{h}\} + j\Im\{\mathbf{h}\} = \begin{pmatrix} \mathbf{I} & j\mathbf{I} \end{pmatrix}\boldsymbol{\theta} = \mathbf{C}\boldsymbol{\theta}$. As a result, the GCRB on

the estimate of $\mathbf{h}$ can be expressed as

$$\mathbb{E}\left[\left(\mathbf{h}-\hat{\mathbf{h}}\right)\left(\mathbf{h}-\hat{\mathbf{h}}\right)^H\right] = \mathbf{C}\mathbb{E}\left[\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}\right)\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}\right)^H\right]\mathbf{C}^H \tag{6.36}$$

$$\geq \mathbf{C}\mathbf{J}_{\boldsymbol{\theta}}^{-1}\mathbf{C}^H \tag{6.37}$$

To gather some insight about the obtained results, we are going to look at the behavior of the GCRB for low and high values of E_s/N_0. From the definitions of λ_n, $\gamma_{l,n}$ and η_l, we can see that for (very) low values of E_s/N_0, the second and third terms in (6.108)-(6.110) are proportional to $(E_s/N_0)^2$, whereas the first terms in (6.108)-(6.110) are proportional to E_s/N_0. Furthermore, the autocorrelation matrix $\mathbf{R}_{\tilde{\mathbf{w}}}$ can be approximated as $\mathbf{R}_{\tilde{\mathbf{w}}} = N_0\mathbf{I}_{N+\nu}$. As a result the Fischer Information matrix $\mathbf{J}_{\boldsymbol{\theta}}$ reduces to

$$\mathbf{J}_{\boldsymbol{\theta},\text{low}} = \frac{2}{N_0}\begin{pmatrix} \Re\left\{\mathbf{B}^H\mathbf{B}\right\} & -\Im\left\{\mathbf{B}^H\mathbf{B}\right\} \\ \Im\left\{\mathbf{B}^H\mathbf{B}\right\} & \Re\left\{\mathbf{B}^H\mathbf{B}\right\} \end{pmatrix}. \tag{6.38}$$

It is easily shown that the inverse of $\mathbf{J}_{\boldsymbol{\theta},\text{low}}$ is given by

$$\mathbf{J}_{\boldsymbol{\theta},\text{low}}^{-1} = \frac{N_0}{2}\begin{pmatrix} \Re\left\{\left(\mathbf{B}^H\mathbf{B}\right)^{-1}\right\} & -\Im\left\{\left(\mathbf{B}^H\mathbf{B}\right)^{-1}\right\} \\ \Im\left\{\left(\mathbf{B}^H\mathbf{B}\right)^{-1}\right\} & \Re\left\{\left(\mathbf{B}^H\mathbf{B}\right)^{-1}\right\} \end{pmatrix}. \tag{6.39}$$

The GCRB for low E_s/N_0 is then given by

$$GCRB_{\text{low}} = N_0\text{tr}\left(\left(\mathbf{B}^H\mathbf{B}\right)^{-1}\right). \tag{6.40}$$

The same result can be obtained by immediately neglecting the contribution of the unknown data symbols in the observation vector $\mathbf{z}$ (6.2). This GCRB still depends on the specific values of the transmitted pilot symbols. The averaging of $GCRB_{\text{low}}$ is carried out in appendix 6.C and yields

$$\mathbb{E}_{\mathbf{a}_c,\mathbf{a}_g}\left[GCRB_{\text{low}}\right] \geq \frac{L}{M}\frac{N+\nu}{N}\frac{N_0}{E_s}. \tag{6.41}$$

Hence, $GCRB_{\text{low}}$ is inversely proportional to the number of pilot symbols M.

For high values of E_s/N_0, the second and third terms (6.108)-(6.110) are independent of E_s/N_0. When we express the first term as

$$\mathbf{B}^H\mathbf{R}_{\tilde{\mathbf{w}}}^{-1}\mathbf{B} = \mathbf{B}_1^H\mathbf{R}_1^{-1}\mathbf{B}_1 + \frac{1}{N_0}\mathbf{B}_2^H\mathbf{B}_2, \tag{6.42}$$

we see that $\mathbf{B}_1^H\mathbf{R}_1^{-1}\mathbf{B}_1$ also becomes independent of E_s/N_0 and the second term is proportional to E_s/N_0. Hence the Fisher information matrix $\mathbf{J}_\theta$ is dominated by $\frac{1}{N_0}\mathbf{B}_2^H\mathbf{B}_2$ for high values of E_s/N_0. As a result, the high SNR limit of the GCRB is given by

$$GCRB_{\text{high}} = N_0\text{tr}\left(\left(\mathbf{B}_2^H\mathbf{B}_2\right)^{-1}\right). \tag{6.43}$$

Note that this $GCRB_{\text{high}}$ is equal to the $GCRB$ corresponding to the observation $\mathbf{z}_2$ only. We can conclude that for high E_s/N_0, the observation $\mathbf{z}_1$, which depends on the unknown data symbols, does not contain much additional information about the channel impulse response compared to the observation $\mathbf{z}_2$.

In the special case that an OFDM block consists entirely of pilot symbols, the GCRB is exactly equal to $GCRB_{\text{low}} = N_0 \text{tr}\left(\left(\mathbf{B}^H\mathbf{B}\right)^{-1}\right)$ and averaged over the pilot symbols, this all pilot GCRB is lower bounded by

$$\mathbb{E}_{\mathbf{a}_c,\mathbf{a}_g}\left[GCRB_{\text{allpilot}}\right] \geq \frac{L}{N}\frac{N_0}{E_s}. \tag{6.44}$$

This lower bound is obtained by substituting M by $N+\nu$ in (6.41).

6.3 Low SNR ML Based Estimator

The first practical channel estimator is proposed in [23]. The algorithm is an ML based estimator and is a simplified version of the estimator proposed in [27] . The algorithm from [27] is slightly modified to take into account all pilot symbols, and not only the guard interval pilots.

For low E_s/N_0, the presence of the data symbols in the observation $\mathbf{z}$ (6.2) can be neglected:

$$\mathbf{z} \approx \mathbf{B}\mathbf{h} + \mathbf{w}. \tag{6.45}$$

Based on this observation, the ML estimate is given by

$$\hat{\mathbf{h}} = \left(\mathbf{B}^H\mathbf{B}\right)^{-1}\mathbf{B}^H\mathbf{z}. \tag{6.46}$$

The MSE of this estimator is given by

$$\mathbb{E}\left[\left|\mathbf{h}-\hat{\mathbf{h}}\right|^2\right] = \text{tr}\left(\left(\mathbf{B}^H\mathbf{B}\right)^{-1}\mathbf{B}^H\mathbf{R}_{\bar{\mathbf{w}}}\mathbf{B}\left(\mathbf{B}^H\mathbf{B}\right)^{-1}\right). \tag{6.47}$$

For low E_s/N_0, this MSE is approximately equal to $N_0\text{tr}\left(\left(\mathbf{B}^H\mathbf{B}\right)^{-1}\right)$ because $\mathbf{R}_{\bar{\mathbf{w}}}$ reduces to $N_0\mathbf{I}_{N+\nu}$. Hence at low E_s/N_0, the MSE is inversely proportional to E_s/N_0. This estimator shows an error floor for high E_s/N_0 due to the presence of the unknown data symbols in the observation vector $\mathbf{z}$ (6.2): the MSE becomes equal to:

$$MSE = E_s\frac{N}{N+\nu}\text{tr}\left(\left(\mathbf{B}^H\mathbf{B}\right)^{-1}\mathbf{B}^H\mathbf{H}\mathbf{F}_d^H\mathbf{F}_d\mathbf{H}^H\mathbf{B}\left(\mathbf{B}^H\mathbf{B}\right)^{-1}\right). \tag{6.48}$$

Since $\mathbf{B}$ is known by the receiver, the matrix $\left(\mathbf{B}^H\mathbf{B}\right)^{-1}\mathbf{B}^H$ can be precomputed and the estimate (6.46) can be obtained with a low computational complexity.

6.4 Subset Estimator

In section 6.2, it is shown that after applying the transform $\mathbf{T}$, the observation vector $\mathbf{z}_2$ is independent of the unknown data symbols and linear in the channel impulse response vector $\mathbf{h}$ (6.12):

$$\mathbf{z}_2 = \mathbf{B}_2\mathbf{h} + \mathbf{w}_2. \tag{6.49}$$

Based on the observation $\mathbf{z}_2$, the authors from [52] derive the ML estimate of $\mathbf{h}$

$$\hat{\mathbf{h}} = \left(\mathbf{B}_2^H\mathbf{B}_2\right)^{-1}\mathbf{B}_2^H\mathbf{z}_2. \tag{6.50}$$

The MSE of this estimate is given by

$$\mathbb{E}\left[\left|\mathbf{h} - \hat{\mathbf{h}}\right|^2\right] = N_0\mathrm{tr}\left(\left(\mathbf{B}_2^H\mathbf{B}_2\right)^{-1}\right). \tag{6.51}$$

This MSE is equal to the GCRB based on the observation $\mathbf{z}_2$. Unfortunately, one should recall that the observation model for $\mathbf{z}_2$ is only valid when the assumption $\mathbf{H}\mathbf{F}_\mathrm{d} = \tilde{\mathbf{F}}\tilde{\mathbf{H}}$ holds. This is only the case for infinite block sizes. For finite block sizes the observation $\mathbf{z}_2$ is affected by a residual contribution from the unknown data symbols. So, for finite block sizes the MSE is given by [52]

$$\mathbb{E}\left[\left|\mathbf{h} - \hat{\mathbf{h}}\right|^2\right] = \mathrm{tr}\left(\mathbf{D}\mathbf{R}_{\bar{\mathbf{w}}}\mathbf{D}^H\right), \tag{6.52}$$

with the matrix $\mathbf{D}$ defined as $\mathbf{D} = \left(\mathbf{B}_2^H\mathbf{B}_2\right)^{-1}\mathbf{B}_2^H\mathbf{T}_2$, where the matrix $\mathbf{T}_2$ consists of the last M rows of $\mathbf{T}$.

For low E_s/N_0, the autocorrelation matrix $\mathbf{R}_{\bar{\mathbf{w}}}$ can be approximated by $N_0\mathbf{I}_{N+\nu}$ and the MSE of the subset estimator reaches (6.51), while for high E_s/N_0, the contribution of the data symbols becomes dominant and the MSE shows an error floor equal to $E_\mathrm{s}\frac{N}{N+\nu}\mathrm{tr}\left(\mathbf{D}\mathbf{H}\mathbf{F}_\mathrm{d}^H\mathbf{F}_\mathrm{d}\mathbf{H}^H\mathbf{D}^H\right)$.

The main advantage of this estimator is its low computational complexity, because the necessary transform $\mathbf{T}$ and the pilot matrix $\mathbf{B}_2$ are known by the receiver. As a result the matrix $\mathbf{D}$ can be precomputed and the estimate is then given by

$$\hat{\mathbf{h}} = \mathbf{D}\mathbf{z}, \tag{6.53}$$

where the observation $\mathbf{z}$ is given by (6.2).

6.5 Special Case 1: The All Pilots ML TD Estimator

When all the carriers are used as pilot carriers, i.e. $S_\mathrm{p} = \{0, \ldots, N-1\}$, the observation is exactly given by

$$\mathbf{z} = \mathbf{B}\mathbf{h} + \mathbf{w}. \tag{6.54}$$

In that case, both the estimators from section 6.3 and 6.4 result in the same estimator, which is the ML estimator based on the observation $\mathbf{z}$:

$$\hat{\mathbf{h}} = \left(\mathbf{B}^H\mathbf{B}\right)^{-1}\mathbf{B}^H\mathbf{z}. \tag{6.55}$$

Its MSE is given by

$$\mathbb{E}\left[\left|\mathbf{h}-\hat{\mathbf{h}}\right|^2\right] = N_0\mathrm{tr}\left(\left(\mathbf{B}^H\mathbf{B}\right)^{-1}\right). \tag{6.56}$$

This MSE is equal to the Cramer Rao bound and the MSE averaged over the pilot symbols is lower bounded by (6.44)

$$\mathbb{E}\left[\left|\mathbf{h}-\hat{\mathbf{h}}\right|^2\right] \geq \frac{L}{N}\frac{N_0}{E_\mathrm{s}} = MSE_{\mathrm{allpilots,LB,TD}}. \tag{6.57}$$

Of course this estimator does not have any practical value, it only serves as a benchmark for other more practical estimation algorithms.

6.6 FD Data Aided Estimator

The major drawback of both estimators from the previous sections is the fact that their MSE shows an error floor at high E_s/N_0 because the unknown data symbols are also present in the observation $\mathbf{z}$ (6.2). In this section, we derive a new channel estimator following a similar reasoning as in section 6.4. We consider an invertible transform independent of the parameter to be estimated that results in a part of the observation to be data-free. However in contrast with section 6.4, we do not make any approximations. The obvious way to separate the data carriers from the pilot carriers, is by transforming the observation to the FD. First we multiply the observation vector $\mathbf{z}$ with the matrix $\mathbf{\Omega}$ (see section 4.2, equation (4.19)) which corresponds to adding the last ν samples of $\mathbf{z}$ to the first ν samples of $\mathbf{z}$, and the resulting first N samples are then applied to an FFT:

$$\tilde{\mathbf{z}} = \mathbf{F}\mathbf{\Omega}\mathbf{z}. \tag{6.58}$$

The $M-\nu$ FFT outputs corresponding to the pilot carriers are given by

$$\begin{aligned}\tilde{\mathbf{z}}_2 &= \left(\tilde{z}\left(\alpha_0\right), \ldots, \tilde{z}\left(\alpha_{M-\nu-1}\right)\right)^T,\ \alpha_n \in S_\mathrm{p} \\ &= \tilde{\mathbf{B}}_2\mathbf{h} + \tilde{\mathbf{w}}_2.\end{aligned} \tag{6.59}$$

The noise vector $\tilde{\mathbf{w}}_2$ is zero-mean Gaussian distributed with an autocorrelation matrix equal to

$$\left(\tilde{\mathbf{R}}_2\right)_{n,m} = N_0\left(\delta_{n-m} + \frac{1}{N}\sum_{l=0}^{\nu-1} e^{-j2\pi\frac{(\alpha_n-\alpha_m)l}{N}}\right), \tag{6.60}$$

while the pilot matrix $\tilde{\mathbf{B}}_2$ is a $(M-\nu) \times L$ matrix which can be expressed as the sum of a part that depends on the pilot symbols on the pilot carriers, denoted as $\tilde{\mathbf{B}}_{\mathrm{c}}$ and a part that depends on the TD pilot symbols in the guard interval, denoted as $\tilde{\mathbf{B}}_{\mathrm{g}}$:

$$\tilde{\mathbf{B}}_2 = \tilde{\mathbf{B}}_{2,\mathrm{c}} + \tilde{\mathbf{B}}_{2,\mathrm{g}}. \tag{6.61}$$

The contribution of the pilot symbols in the pilot carriers $\tilde{\mathbf{B}}_{2,\mathrm{c}}$ is defined as

$$\left(\tilde{\mathbf{B}}_{2,\mathrm{c}}\right)_{n,l} = \sqrt{\frac{N}{N+\nu}} a_{\mathrm{c}}(n)\, e^{-j2\pi\frac{\alpha_n l}{N}},\ \alpha_n \in S_{\mathrm{p}},\ l = 0,\ \ldots,\ L-1; \tag{6.62}$$

where $a_{\mathrm{c}}(n)$ is the pilot symbol transmitted on the n-th pilot carrier (see section 5.2). The matrix $\tilde{\mathbf{B}}_{\mathrm{g}}$ is given by

$$\tilde{\mathbf{B}}_{2,\mathrm{g}} = \sqrt{\frac{N}{N+\nu}} \mathbf{F}_{\nu,\mathrm{p}} \mathbf{B}_{\mathrm{g},\nu} \tag{6.63}$$

where $\left(\mathbf{F}_{\nu,\mathrm{p}}\right)_{n,k} = \frac{1}{\sqrt{N}} e^{-j2\pi\frac{\alpha_n k}{N}}$ and $\left(\mathbf{B}_{\mathrm{g},\nu}\right)_{k,l} = a_{\mathrm{g}}\left(|k-l|_\nu\right)$, $\alpha_n \in S_{\mathrm{p}}$, $k = 0,\ \ldots,\ \nu-1$, $l = 0,\ \ldots,\ L-1$.

The ML estimate of $\mathbf{h}$ based on the observation $\tilde{\mathbf{z}}_2$ is easily found to be

$$\hat{\mathbf{h}} = \left(\tilde{\mathbf{B}}_2^H \tilde{\mathbf{R}}_2^{-1} \tilde{\mathbf{B}}_2\right)^{-1} \tilde{\mathbf{B}}_2^H \tilde{\mathbf{R}}_2^{-1} \tilde{\mathbf{z}}_2, \tag{6.64}$$

and its MSE is easily found by substituting $\tilde{\mathbf{z}}_2$ by its definition (6.59):

$$\mathbb{E}\left[\left|\mathbf{h} - \hat{\mathbf{h}}\right|^2\right] = \mathrm{tr}\left(\left(\tilde{\mathbf{B}}_2^H \tilde{\mathbf{R}}_2^{-1} \tilde{\mathbf{B}}_2\right)^{-1}\right). \tag{6.65}$$

The MSE is proportional to N_0/E_{s}, so there is no error floor for $N_0/E_{\mathrm{s}} \to 0$. This MSE still depends on the actual values of the pilot symbols $\mathbf{a}_{\mathrm{c}}$ and $\mathbf{a}_{\mathrm{g}}$. The averaging of (6.65) over all possible pilot sequences is not straightforward, so we derive a lower bound instead. In appendix 6.D we show that

$$\mathbb{E}\left[\left(\tilde{\mathbf{B}}_2^H \tilde{\mathbf{R}}_2^{-1} \tilde{\mathbf{B}}_2\right)^{-1}\right] \geq \frac{N+\nu}{N} \frac{1}{M-\nu} \frac{N_0}{E_{\mathrm{s}}} \mathbf{I}_L, \tag{6.66}$$

so the MSE is lower bounded by

$$\mathbb{E}\left[\left|\mathbf{h} - \hat{\mathbf{h}}\right|^2\right] \geq \frac{N+\nu}{N} \frac{L}{M-\nu} \frac{N_0}{E_{\mathrm{s}}}, \tag{6.67}$$

i.e. the MSE lower bound is inversely proportional to the number of pilot carriers.

The matrices $\tilde{\mathbf{B}}_2$ and $\tilde{\mathbf{R}}_2$ depend only on the known pilot symbols and the known pilot carrier positions. Hence, $\tilde{\mathbf{B}}_2$ and $\tilde{\mathbf{R}}_2$ are known at the receiver and the matrix $\left(\tilde{\mathbf{B}}_2^H \tilde{\mathbf{R}}_2^{-1} \tilde{\mathbf{B}}_2\right)^{-1} \tilde{\mathbf{B}}_2^H \tilde{\mathbf{R}}_2^{-1}$ can be precomputed. Therefore, the estimate (6.64) can be obtained with low complexity.

6.7 Special Case 2: The All Pilots FD Estimator

Similar as in section 6.5, we can consider the all pilot case, i.e. all the data symbols are known by the receiver. This case serves again as a benchmark for the achievable performance of a channel estimation algorithm that operates in the FD.

The estimate of $\mathbf{h}$ is given by (6.64) but where S_{p} contains all the N carriers. The MSE averaged over the pilot symbols is lower bounded by

$$\mathbb{E}\left[\left|\mathbf{h}-\hat{\mathbf{h}}\right|^2\right] \geq \frac{N+\nu}{N}\frac{L}{N}\frac{N_0}{E_{\mathrm{s}}} = MSE_{\mathrm{allpilots,LB,FD}}. \tag{6.68}$$

It is interesting to compare this lower bound with the lower bound on the MSE of the TD all pilots estimator from section 6.5:

$$MSE_{\mathrm{allpilots,LB,TD}} = \frac{L}{N}\frac{N_0}{E_{\mathrm{s}}}. \tag{6.69}$$

There is an extra factor $\frac{N+\nu}{N}$ in the lower bound $MSE_{\mathrm{allpilots,LB,FD}}$, suggesting that the achievable performance in FD is slightly worse than the achievable performance in the TD. This is caused by the fact that before the FFT is applied, the last ν samples of the observation vector are added to the first ν samples. This operation causes noise enhancement. We can see this from the definition of the autocorrelation matrix of the FD noise samples $\tilde{\mathbf{R}}_2$ (6.60). The diagonal elements are equal to $\frac{N+\nu}{N}N_0$, while the variance of the TD noise samples is given by N_0.

6.8 Iterative Hard Decision Aided Channel Estimation

To improve the performance of the already proposed data aided channel estimation algorithms, we need to resort to iterative decision aided algorithms. The most simple type of a decision aided algorithm updates every iteration the channel impulse response estimate using the obtained hard decisions on the data symbols from the previous iteration. Then after every update of the channel impulse response estimate, new hard decisions on the unknown data symbols are obtained using the most recent channel impulse response estimate. Here we propose two strategies than can be applied.

6.8.1 TD Hard Decision Aided Channel Estimation

A first one operates on the TD observation vector $\mathbf{z}$ (6.2). The signal model from (6.2) can be expressed as

$$\mathbf{z} = \mathbf{B}\mathbf{h} + \mathbf{A}\mathbf{h} + \mathbf{w}, \tag{6.70}$$

where we have rewritten the contribution of the data symbols $\mathbf{H}\mathbf{F}_{\mathrm{d}}^{H}\mathbf{a}_{\mathrm{d}}$ as $\mathbf{A}\mathbf{h}$, where $\mathbf{A}$ is a $(N+\nu)\times L$ matrix which solely depends on the unknown data symbols:

$$(\mathbf{A})_{k,l} = s_{\mathrm{d}}\left(|k-l|_{N+\nu}\right), \tag{6.71}$$

with $k = 0, \ldots, N+\nu-1$, $l = 0, \ldots, L-1$ and the components of the vector $\mathbf{s}_{\mathrm{d}}$ are defined by (5.8).

When the matrix $\mathbf{A}$ is known, the ML estimate of the channel impulse response given the observation $\mathbf{z}$ can easily found to be

$$\hat{\mathbf{h}} = \left(\mathbf{C}^{H}\mathbf{C}\right)^{-1}\mathbf{C}^{H}\mathbf{z}, \tag{6.72}$$

where the matrix $\mathbf{C}$ is defined as

$$\mathbf{C} = \mathbf{A} + \mathbf{B}. \tag{6.73}$$

Of course the matrix $\mathbf{A}$ is not a priori known by the receiver. Now suppose that we have obtained a hard decision on every data symbol, we then can use an estimate $\hat{\mathbf{A}}$ of the matrix $\mathbf{A}$, by replacing the unknown data symbols in the vector $\mathbf{s}_{\mathrm{d}}$ (5.8) by their hard decisions. So the expression (6.72) for the estimate of the channel impulse response remains the same but we replace $\mathbf{C}$ by an estimate $\hat{\mathbf{C}}$.

Now we still have to obtain the hard decisions on the data symbols. Since, we are working in the TD, it seems reasonable to apply the ZF detector from section (4.1.2).

So every iteration the algorithm performs two steps: i) a new channel impulse response estimate is obtained by applying (6.72) using the obtained hard decisions from the previous iteration, and ii) using this new channel impulse response estimate to apply the ZF detector to make new hard decisions on the unknown data symbols. For the initialization of the algorithm, the data aided FD channel impulse response estimator (6.64) generates the initial estimate of the channel impulse response, which is used to apply the ZF detector for the first time.

6.8.2 FD Hard Decision Aided Channel Estimation

A second approach starts from FD observation $\tilde{\mathbf{z}}$ (6.58), which can be expressed as

$$\tilde{\mathbf{z}} = \tilde{\mathbf{C}}\mathbf{h} + \tilde{\mathbf{w}}, \tag{6.74}$$

where the matrix $\tilde{\mathbf{C}}$ is the FD version of the matrix $\mathbf{C}$ from the previous subsection 6.8.1: $\tilde{\mathbf{C}} = \mathbf{F}\boldsymbol{\Omega}\mathbf{C}$. The components of $\tilde{\mathbf{C}}$ are given by

$$\left(\tilde{\mathbf{C}}\right)_{\alpha_n,l} = \sqrt{\frac{N}{N+\nu}}a_{\mathrm{c}}(n)\,e^{-j2\pi\frac{\alpha_n l}{N}} + \left(\tilde{\mathbf{B}}_{\mathrm{g}}\right)_{\alpha_n,l}, \tag{6.75}$$

$$\left(\tilde{\mathbf{C}}\right)_{\beta_n,l} = \sqrt{\frac{N}{N+\nu}}a_{\mathrm{d}}(n)\,e^{-j2\pi\frac{\beta_n l}{N}} + \left(\tilde{\mathbf{B}}_{\mathrm{g}}\right)_{\beta_n,l}, \tag{6.76}$$

with $\alpha_n \in S_\mathrm{p}$, $\beta_n \in S_\mathrm{d}$ and $l = 0, \ldots, L-1$. The $N \times L$ matrix $\tilde{\mathbf{B}}_\mathrm{g}$ collects the contributions from the guard interval pilot symbols and is defined as

$$\tilde{\mathbf{B}}_\mathrm{g} = \sqrt{\frac{N}{N+\nu}}\mathbf{F}_\nu \mathbf{B}_{\mathrm{g},\nu}$$

where $(\mathbf{F}_\nu)_{n,k} = \frac{1}{\sqrt{N}} e^{-j2\pi\frac{nk}{N}}$ and $\left(\mathbf{B}_{\mathrm{g},\nu}\right)_{k,l} = a_\mathrm{g}\left(|k-l|_\nu\right)$, $n = 0, \ldots, N$, $k = 0, \ldots, \nu-1$, $l = 0, \ldots, L-1$. The components of the noise vector $\tilde{\mathbf{w}}$ are Gaussian distributed with zero mean and an autocorrelation matrix equal to $\mathbf{R}_{\tilde{\mathbf{w}}}$ (4.25).

When the data symbols $\mathbf{a}_\mathrm{d}$ would be known by the receiver, the ML estimate of $\mathbf{h}$ based on the observation $\tilde{\mathbf{z}}$ would be given by

$$\hat{\mathbf{h}} = \left(\tilde{\mathbf{C}}^H \mathbf{R}_{\tilde{\mathbf{w}}}^{-1} \tilde{\mathbf{C}}\right)^{-1} \tilde{\mathbf{C}}^H \tilde{\mathbf{z}}. \tag{6.77}$$

Since the data symbols are not a priori known by the receiver, we can replace them by their hard decisions in the matrix $\tilde{\mathbf{C}}$. To obtain the necessary hard decisions the FD symbol by symbol detector from (4.28) is applied.

Every iteration the channel impulse response estimate is updated by applying (6.77) using the latest obtained hard decisions on the data symbols. To get new hard decisions in the unknown data symbols, the receiver makes use of the most recent estimate of the channel impulse response. The initialization of this algorithm is carried out by the data aided frequency domain channel impulse response estimator from section 6.6.

6.9 EM based Channel Estimation

To deal with the unknown data symbols, we can apply the EM algorithm. In [57], EM-based channel estimation algorithms operating in the frequency domain are proposed for CP-OFDM. However, when this algorithm is applied to KSP-OFDM, the pilot symbols from the guard interval cannot optimally be used for the channel estimation, as in the algorithm from [57], the samples from the guard interval are thrown away. The extension of this algorithm to take also into account the guard interval samples to estimate the channel in the frequency domain is not straightforward. Therefore, we propose an EM based channel estimator for KSP-OFDM operating in the time domain, using both the samples from the data part and the samples from the guard interval.

We consider the channel impulse response as the parameter vector that needs to be estimated, while the unknown data symbols are nuisance parameters. Every iteration, the algorithm performs the E-step and the M-step. In the E-step of the i-th iteration, the log likelihood $\ln p\left(\mathbf{z}|\mathbf{h}, \mathbf{a}_\mathrm{d}\right)$ is averaged over the unknown data symbols, given the observation $\mathbf{z}$ and the most recent estimate of $\mathbf{h}$:

$$Q\left(\mathbf{h}|\hat{\mathbf{h}}\left[i-1\right]\right) = \mathbb{E}_{\mathbf{a}_\mathrm{d}}\left[\ln p\left(\mathbf{z}|\mathbf{h}, \mathbf{a}_\mathrm{d}\right)|\mathbf{z}, \hat{\mathbf{h}}\left[i-1\right]\right]. \tag{6.78}$$

The vector of received samples $\mathbf{z}$ (6.70), given the channel impulse response $\mathbf{h}$ and the data symbols $\mathbf{a}_\mathrm{d}$, is Gaussian distributed with mean $\mathbf{Ch}$ (see (6.73) for the definition of $\mathbf{C}$) and autocorrelation matrix $N_0\mathbf{I}_{N+\nu}$, so the log likelihood $\ln p\left(\mathbf{z}|\mathbf{h},\mathbf{a}_\mathrm{d}\right)$is given by

$$\ln p\left(\mathbf{z}|\mathbf{h},\mathbf{a}_\mathrm{d}\right) = C - \frac{1}{N_0}\left(\mathbf{z}-\mathbf{Ch}\right)^H\left(\mathbf{z}-\mathbf{Ch}\right), \tag{6.79}$$

where C is a constant which is independent of the channel impulse response $\mathbf{h}$ and the data symbols $\mathbf{a}_\mathrm{d}$. The averaging of the log likelihood $\ln p\left(\mathbf{z}|\mathbf{h},\mathbf{a}_\mathrm{d}\right)$ over the unknown data symbols yields

$$\begin{aligned} Q\left(\mathbf{h}|\hat{\mathbf{h}}\left[i-1\right]\right) = \mathbb{E}_{\mathbf{a}_\mathrm{d}}\left[\ln p\left(\mathbf{z}|\mathbf{h},\mathbf{a}_\mathrm{d}\right)|\mathbf{z},\hat{\mathbf{h}}\left[i-1\right]\right] = \\ C - \frac{1}{N_0}\left(\mathbf{z}^H\mathbf{z} - \mathbf{z}^H\bar{\mathbf{C}}\mathbf{h} - \mathbf{h}^H\bar{\mathbf{C}}^H\mathbf{z} + \mathbf{h}^H\bar{\mathbf{R}}_\mathbf{C}\mathbf{h}\right), \end{aligned} \tag{6.80}$$

where the matrices $\bar{\mathbf{C}}$ and $\bar{\mathbf{R}}_\mathbf{C}$ are defined as

$$\bar{\mathbf{C}} = \mathbb{E}_{\mathbf{a}_\mathrm{d}}\left[\mathbf{C}|\mathbf{z},\hat{\mathbf{h}}\left[i-1\right]\right] \tag{6.81}$$

$$\bar{\mathbf{R}}_\mathbf{C} = \mathbb{E}_{\mathbf{a}_\mathrm{d}}\left[\mathbf{C}^H\mathbf{C}|\mathbf{z},\hat{\mathbf{h}}\left[i-1\right]\right]. \tag{6.82}$$

Taking into account (6.73), the former equations can be rewritten as

$$\bar{\mathbf{C}} = \mathbf{B} + \mathbb{E}_{\mathbf{a}_\mathrm{d}}\left[\mathbf{A}|\mathbf{z},\hat{\mathbf{h}}\left[i-1\right]\right] \tag{6.83}$$

$$\bar{\mathbf{R}}_\mathbf{C} = \mathbf{B}^H\mathbf{B} + \mathbb{E}_{\mathbf{a}_\mathrm{d}}\left[\mathbf{B}^H\mathbf{A} + \mathbf{A}^H\mathbf{B} + \mathbf{A}^H\mathbf{A}|\mathbf{z},\hat{\mathbf{h}}\left[i-1\right]\right]. \tag{6.84}$$

To obtain $\bar{\mathbf{C}}$ and $\bar{\mathbf{R}}_\mathbf{C}$, we need to compute

$$\mathbb{E}\left[a_\mathrm{d}\left(n\right)|\mathbf{z},\hat{\mathbf{h}}\left[i-1\right]\right] = \sum_{a_\mathrm{d}(n)} a_\mathrm{d}\left(n\right) p\left(a_\mathrm{d}\left(n\right)|\mathbf{z},\hat{\mathbf{h}}\left[i-1\right]\right) \tag{6.85}$$

and

$$\begin{aligned} &\mathbb{E}\left[a_\mathrm{d}\left(n\right)a_\mathrm{d}^*\left(n'\right)|\mathbf{z},\hat{\mathbf{h}}\left[i-1\right]\right] \\ &\quad= \begin{cases} \sum\limits_{a_\mathrm{d}(n)} \left|a_\mathrm{d}\left(n\right)\right|^2 p\left(a_\mathrm{d}\left(n\right)|\mathbf{z},\hat{\mathbf{h}}\left[i-1\right]\right) & n = n' \\ \mathbb{E}\left[a_\mathrm{d}\left(n\right)|\mathbf{z},\hat{\mathbf{h}}\left[i-1\right]\right]\left(\mathbb{E}\left[a_\mathrm{d}\left(n'\right)|\mathbf{z},\hat{\mathbf{h}}\left[i-1\right]\right]\right)^* & n \neq n' \end{cases} \end{aligned} \tag{6.86}$$

For an uncoded system, the a posteriori distribution of the data symbols $\mathbf{a}_\mathrm{d}$ given the observation $\mathbf{z}$ and the last obtained estimate of the channel impulse response $\hat{\mathbf{h}}[i-1]$, is given by

$$\begin{aligned} p\left(\mathbf{a}_\mathrm{d}|\mathbf{z},\hat{\mathbf{h}}\left[i-1\right]\right) &= \frac{p\left(\mathbf{a}_\mathrm{d},\mathbf{z}|\hat{\mathbf{h}}\left[i-1\right],\mathbf{a}_\mathrm{d}\right)}{p\left(\mathbf{z}|\hat{\mathbf{h}}\left[i-1\right]\right)} \\ &\propto p\left(\mathbf{z}|\mathbf{a}_\mathrm{d},\hat{\mathbf{h}}\left[i-1\right]\right) p\left(\mathbf{a}_\mathrm{d}\right). \end{aligned} \tag{6.87}$$

We assume that all data sequences are equiprobable, so the a posteriori distribution of the data symbols is proportional to

$$p\left(\mathbf{a}_{\mathrm{d}}|\mathbf{z},\hat{\mathbf{h}}\left[i-1\right]\right) \propto \exp\left(-\frac{1}{N_0}\left|\left(\mathbf{z}-\mathbf{B}\hat{\mathbf{h}}\left[i-1\right]-\sqrt{\frac{N}{N+\nu}}\hat{\mathbf{H}}\left[i-1\right]\mathbf{F}_{\mathrm{d}}^{H}\mathbf{a}_{\mathrm{d}}\right)\right|^{2}\right). \quad (6.88)$$

Looking at the a posteriori distribution (6.88) of the data symbols $\mathbf{a}_{\mathrm{d}}$, we see that the data symbols $\mathbf{a}_{\mathrm{d}}$ are not statistically independent because of the presence of the matrix $\mathbf{F}_{\mathrm{d}}\hat{\mathbf{H}}^{H}\left[i-1\right]\hat{\mathbf{H}}\left[i-1\right]\mathbf{F}_{\mathrm{d}}^{H}$, which is not a diagonal matrix. To obtain the a posteriori distribution of one data symbol $a_{\mathrm{d}}\left(n\right)$ we have to average the joint a posteriori distribution of all the data symbols $\mathbf{a}_{\mathrm{d}}$ over all the data symbols $a_{\mathrm{d}}\left(n'\right)$ with $n' \neq n$. This is a computationally hard operation so we are going to look for an alternative.

The a posteriori distribution of the data symbols $\mathbf{a}_{\mathrm{d}}$ given $\mathbf{z}$ and $\hat{\mathbf{h}}\left[i-1\right]$ can be rewritten as

$$p\left(\mathbf{a}_{\mathrm{d}}|\mathbf{z},\hat{\mathbf{h}}\left[i-1\right]\right) \propto \exp\left(-\left(\mathbf{a}_{\mathrm{d}}-\mathbf{m}_{\mathbf{a}}\right)^{H}\mathbf{R}_{\mathbf{a}}^{-1}\left(\mathbf{a}_{\mathrm{d}}-\mathbf{m}_{\mathbf{a}}\right)\right) \quad (6.89)$$

where the matrix $\mathbf{R}_{\mathbf{a}}$ and the vector $\mathbf{m}_{\mathbf{a}}$ are given by

$$\mathbf{R}_{\mathbf{a}} = \frac{N+\nu}{N}N_0\left(\mathbf{F}_{\mathrm{d}}\hat{\mathbf{H}}\left[i-1\right]^{H}\hat{\mathbf{H}}\left[i-1\right]\mathbf{F}_{\mathrm{d}}^{H}\right)^{-1} \quad (6.90)$$

$$\mathbf{m}_{\mathbf{a}} = \frac{1}{N_0}\sqrt{\frac{N}{N+\nu}}\mathbf{R}_{\mathbf{a}}\mathbf{F}_{\mathrm{d}}\hat{\mathbf{H}}\left[i-1\right]^{H}\left(\mathbf{z}-\mathbf{B}\hat{\mathbf{h}}\left[i-1\right]\right). \quad (6.91)$$

Note that the vector $\mathbf{m}_{\mathbf{a}}$ corresponds to the output of the ZF filter (see subsection 4.1.2) when the signal $\mathbf{z}-\mathbf{B}\hat{\mathbf{h}}\left[i-1\right]$ is sent through the ZF filter. Further, the matrix $\hat{\mathbf{H}}\left[i-1\right]^{H}\hat{\mathbf{H}}\left[i-1\right]$ is a Toeplitz matrix. For large N we can approximate $\hat{\mathbf{H}}\left[i-1\right]^{H}\hat{\mathbf{H}}\left[i-1\right]$ by a circulant matrix. The eigenvectors of a circulant $N \times N$ matrix are given by the columns of the $N \times N$ FFT matrix $\mathbf{F}$. The matrix $\mathbf{F}_{\mathrm{d}}$ consists of a subset of those eigenvectors so $\left(\mathbf{R}_{\mathbf{a}}\right)^{-1}$ can be approximated by a diagonal matrix in (6.89). This means that the data symbols $\mathbf{a}_{\mathrm{d}}$, given $\mathbf{z}$ and $\hat{\mathbf{h}}\left[i-1\right]$ can be considered as (approximately) statistically independent. So, this yields for the a posteriori distribution of the data symbols $\mathbf{a}_{\mathrm{d}}$ given $\mathbf{z}$ and $\hat{\mathbf{h}}\left[i-1\right]$:

$$p\left(\mathbf{a}_{\mathrm{d}}|\mathbf{z},\hat{\mathbf{h}}\left[i-1\right]\right) \approx \prod_{n=0}^{N-M+\nu-1} p\left(a_{\mathrm{d}}\left(n\right)|\mathbf{z},\hat{\mathbf{h}}\left[i-1\right]\right) \quad (6.92)$$

with

$$p\left(a_{\mathrm{d}}\left(n\right)|\mathbf{z},\hat{\mathbf{h}}\left[i-1\right]\right) \propto \exp\left(-\left(\mathbf{R}_{\mathbf{a}}^{-1}\right)_{n,n}\left|a_{\mathrm{d}}\left(n\right)-m_{\mathbf{a}}\left(n\right)\right|^{2}\right). \quad (6.93)$$

Now (6.93) can be used in (6.88) and (6.86) to compute the soft information $\mathbb{E}\left[a_{\mathrm{d}}(n)\,|\mathbf{z},\hat{\mathbf{h}}[i-1]\right]$ and $\mathbb{E}\left[a_{\mathrm{d}}(n)\,a_d^*(n')\,|\mathbf{z},\hat{\mathbf{h}}[i-1]\right]$ respectively.

In a coded system, the decoder provides every iteration the necessary a posteriori probability of every data symbol $a_{\mathrm{d}}(n)$.

The M-step of the i-th iteration of the EM algorithm comprises the maximization of $Q\left(\mathbf{h}|\hat{\mathbf{h}}[i-1]\right)$ (6.80). The new estimate $\hat{\mathbf{h}}[i]$ is given by

$$\hat{\mathbf{h}}[i] = \left(\bar{\mathbf{R}}_{\mathbf{C}}\right)^{-1}\bar{\mathbf{C}}^H\mathbf{z}. \tag{6.94}$$

The algorithm terminates once the estimate has reached convergence.

Note that the same expressions for the a posteriori probabilities of the data symbols could have been obtained as in (6.93) by assuming that the ZF detector is used to obtain the a posteriori probabilities of the data symbols. This means that the here proposed EM-based algorithm and the TD hard decision aided algorithm from subsection 6.8.1 can be respectively seen as the soft decision version and the hard decision version of the same TD decision aided algorithm.

6.10 Numerical results and Discussion

In this section, we illustrate the obtained CRBs, MSE lower bounds and the performance of the different proposed channel estimation algorithms in terms of the MSE and the BER.

Unless otherwise mentioned, we consider the comb-type pilot structure. The pilot symbols and the data symbols are both 4-QAM modulated. The channel impulse response is given by (4.29). We consider Rayleigh fading [3], so the samples of $\mathbf{h}$ are Gaussian distributed with zero mean and variance equal to $1/L$. We consider two different systems. Table 6.1 shows the parameters for the two considered systems. For the second system, we have also simulated a coded system. The considered channel code is a turbo code that consists of the parallel concatenation of two identical 16-state rate-1/2 recursive systematic convolutional codes with generator sequences $(37)_8$ and $(21)_8$ (in octal notation) through a uniform interleaver of length $N_{\mathrm{b}} = 919$ bits. The parity bits are punctured so that the length N_{c} of one code word is equal to 1848 bits. The modulation of this code word on 4-QAM symbols results in 924 data symbols, so that the data symbols corresponding to the code word can be transmitted on the data carriers of exactly one OFDM symbol.

6.10.1 CRB

In this subsection, we evaluate the GCRB (6.37) and its low SNR limit $GCRB_{\mathrm{low}}$ (6.40) and high SNR limit $GCRB_{\mathrm{high}}$ (6.43) for both system 1 and system 2.

Parameter	System 1	System 2
N	1024	1024
M	40	200
ν	7	100
L	8	50

Table 6.1: Simulation parameters.

In figures 6.1 and 6.2 we show the results for the normalized CRB, which is defined as:

$$NCRB = CRB \cdot SNR, \tag{6.95}$$

where SNR is defined as

$$SNR = \frac{N}{N+\nu}\frac{E_s}{N_0}. \tag{6.96}$$

Besides the NCRB, the low SNR limit $NGCRB_{\text{low}}$ and its lower bound L/M and the high SNR limit $NGCRB_{\text{low}}$ are also shown. The curve L/M is the normalized version of the lower bound (6.41) on the low SNR limit $GCRB_{\text{low}}$ (6.40). First of all, we see for both systems that the NGCRB is almost constant over the considered range of E_s/N_0. This means that the GCRB is approximately inversely proportional to E_s/N_0. The low and high SNR limits $NGCRB_{\text{low}}$ and $NGCRB_{\text{high}}$ do not change and are independent of E_s/N_0, so the corresponding GCRB's are inversely proportional to E_s/N_0.

For system 1, one can see that the GCRB coincides with its low SNR limit $GCRB_{\text{low}}$ for (very) low E_s/N_0 and with its high SNR limit $GCRB_{\text{high}}$ for (very) high E_s/N_0, respectively for $E_s/N_0 < -50$ dB and for $E_s/N_0 > 50$ dB. Only for moderate values of E_s/N_0, the GCRB distinguishes itself from its limits. The low SNR limit and the high SNR limit are very close to each other and are approximately equal to the lower bound on $GCRB_{\text{low}}$ (6.41), which is equal to $(L/M)\,SNR^{-1}$. Especially for $GCRB_{\text{low}}$, the lower bound (6.41) is very tight.

For the second system, the results are similar as those for system 1. However, the difference between the low SNR and high SNR limit is more pronounced. The GCRB still coincides with $GCRB_{\text{low}}$ for low values E_s/N_0 (< -30 dB) and with $GCRB_{\text{high}}$ for high values of E_s/N_0 (> 30 dB). Compared for the results of system 1, the difference between the GCRB and its limits is smaller for moderate E_s/N_0, because there are less data carriers in system 2 and their contribution to the GCRB is less important. In the limit, when there are only pilot carriers, the GCRB and its lower bounds $GCRB_{\text{low}}$ and $GCRB_{\text{high}}$will perfectly coincide. Furthermore, the lower bound $L/M \cdot SNR^{-1}$ is less tight in situation 2 than in situation 1.

In practice, the GCRB itself is not that useful, since it is not that simple to

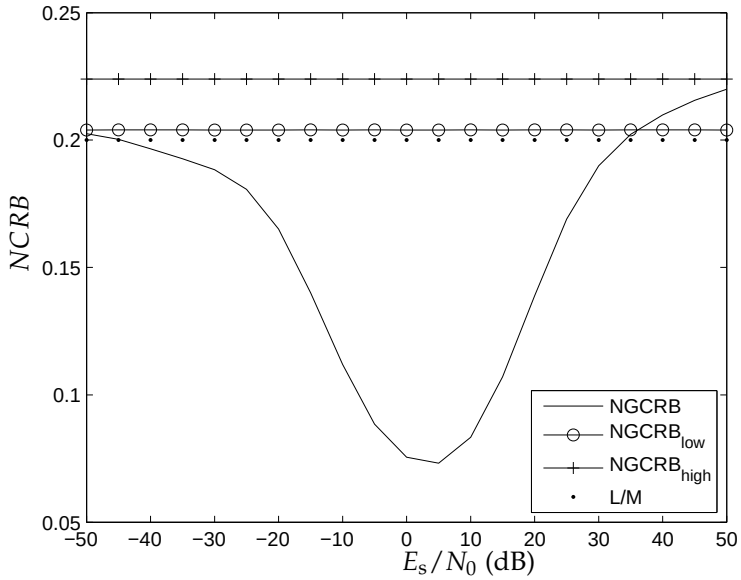

Figure 6.1: NGCRB as a function of E_s/N_0 , $N = 1024$, $\nu = 7$, $L = 8$, $M = 40$.

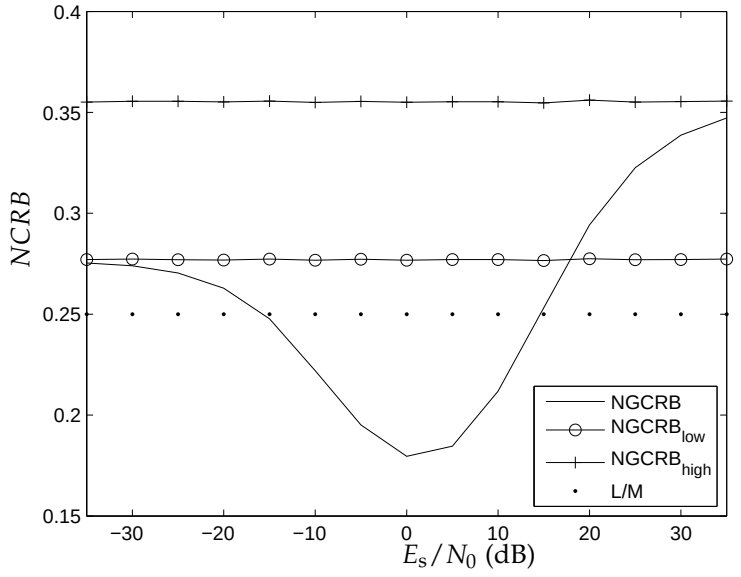

Figure 6.2: NGCRB as a function of E_s/N_0 , $N = 1024$, $\nu = 100$, $L = 50$, $M = 200$.

find a data aided estimation algorithm that can exploit the extra information about the channel impulse response in the noise vector $\bar{\mathbf{w}}$ (6.4). For a practical algorithm, one should compare the MSE with $GCRB_{\text{low}}$ (6.40) and its lower bound $L/M \cdot SNR^{-1}$ (6.41), and $GCRB_{\text{high}}$ (6.43).

6.10.2 MSE

6.10.2.1 Pilot Aided Estimators

First, we evaluate the performance of the pilot aided estimators, which can be used in the initialization phase of the iterative decision aided estimators.

In figures 6.3 and 6.4, the MSE of the low SNR ML based estimator, the subset estimator and the FD estimator are shown as function of E_s/N_0 for system 1 and system 2, respectively. In addition, the low SNR limit of the GCRB ($GCRB_{\text{low}}$), the high SNR limit of the GCRB ($GCRB_{\text{high}}$) and the lower bound on the MSE of the FD DA estimator $L/(M-\nu) \cdot SNR^{-1}$ are also shown. As can be observed, the FD estimator does not suffer from an error floor at high E_s/N_0, in contrast with the subset estimator and the low SNR ML based estimator.

For situation 1, the performances of the different pilot aided estimators are very close to each other and to the lower bounds for low E_s/N_0. Both the subset estimator and the low SNR ML based estimator show an error floor at high E_s/N_0. This error floor is caused by the interference of the unknown data symbols. In this case, the subset estimator yields the lowest error floor.

For situation 2, the MSE results are similar but there are some differences. First of all we see that the MSEs of the different pilot aided estimators remain close to the corresponding lower bounds. For low E_s/N_0, the low SNR ML based estimator results in the lowest MSE followed by the subset estimator. In fact for this case, the low SNR ML based estimator totally outperforms the subset estimator: although both estimators show an error floor at high E_s/N_0, the low SNR ML based estimator yields the lowest error floor. The performance of the low SNR ML based estimator is better in this situation because of several factors. There are less data carriers in situation 2 so the presence of the data carriers is less dominant. Secondly the length of the guard interval ν is also larger in situation 2, so the assumption that we made for the derivation of the subset estimator, i.e. $N \gg \nu$ is less realistic in situation 2 than in situation 1. The FD estimator has the highest MSE for low E_s/N_0, but does not show any error floor for high E_s/N_0. The lower bound $L/(M-\nu) \cdot SNR^{-1}$ is not as tight as in situation 1. For moderate to high E_s/N_0, it seems that the FD estimator yields the best performance in terms of the MSE and as a result the FD estimator is the most appropriate choice for the initialization phase of the iterative decision aided algorithms.

To further evaluate the MSEs, we consider the normalized MSE (NMSE),

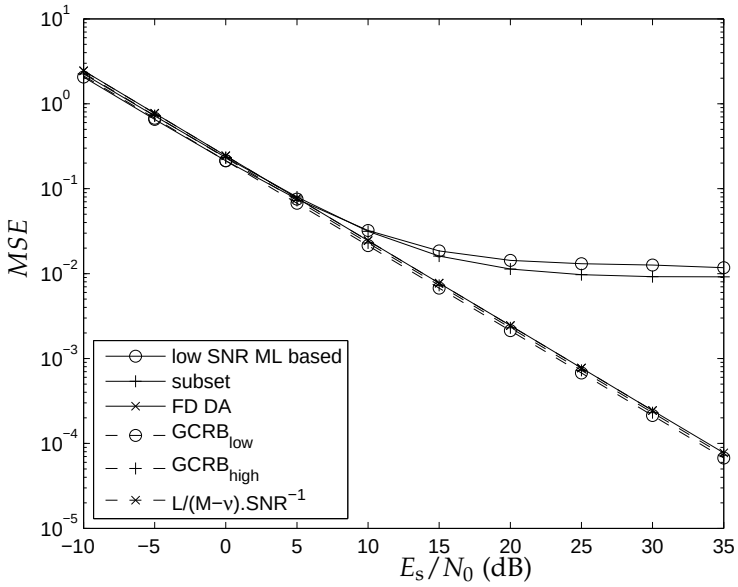

Figure 6.3: MSE as a function of E_s/N_0 , $N = 1024$, $\nu = 7$, $L = 8$, $M = 40$.

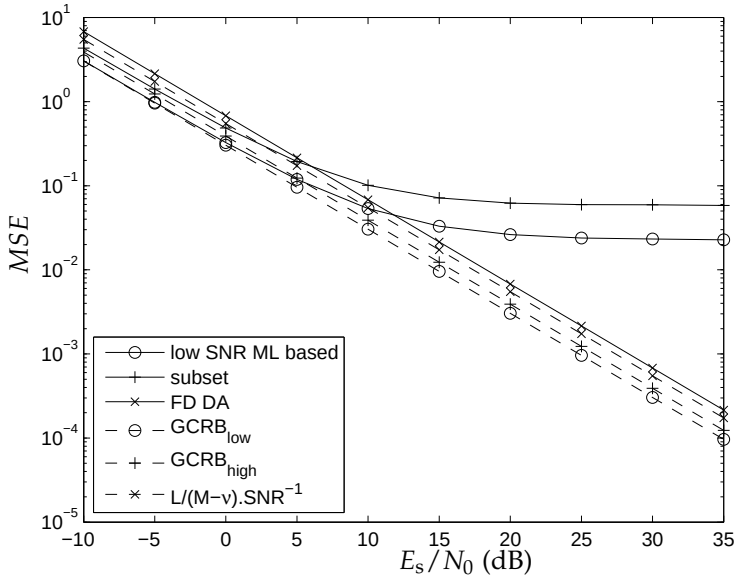

Figure 6.4: MSE as a function of E_s/N_0 , $N = 1024$, $\nu = 100$, $L = 50$, $M = 200$.

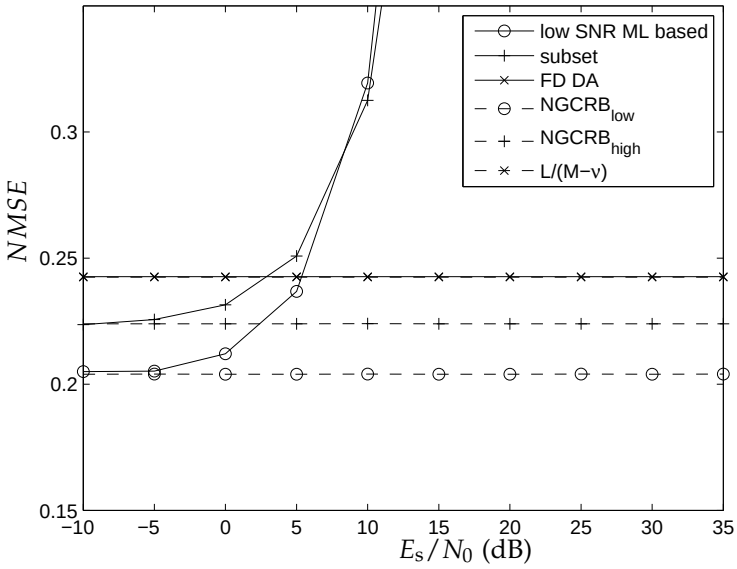

Figure 6.5: NMSE as a function of E_s/N_0 , $N = 1024$, $\nu = 7$, $L = 8$, $M = 40$.

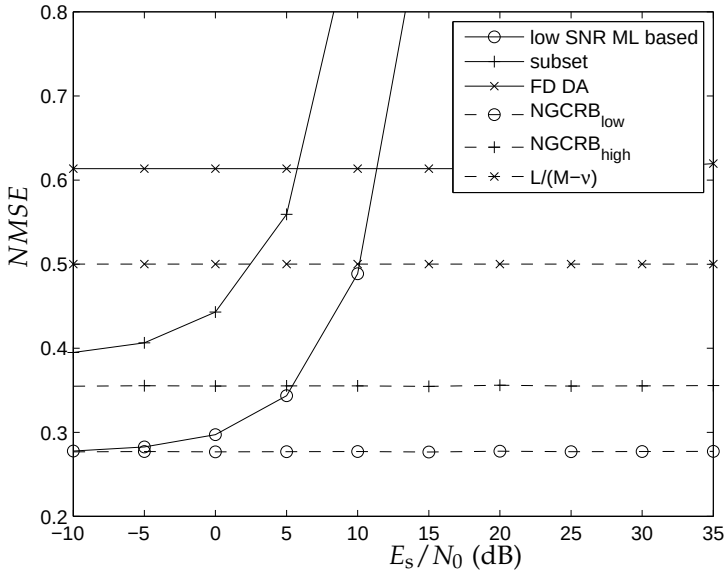

Figure 6.6: NMSE as a function of E_s/N_0 , $N = 1024$, $\nu = 100$, $L = 50$, $M = 200$.

defined as

$$NMSE = SNR \cdot MSE. \tag{6.97}$$

From figures 6.5 and 6.6 it follows that at low E_s/N_0, the low SNR ML based estimator and the subset estimator slightly outperform the FD estimator. This follows from the fact that the low SNR ML based estimator uses the total contribution of the pilot symbols and the subset estimator uses a larger subset of observations than the frequency domain estimator (i.e. M observations for the subset estimator versus $M-\nu$ observations for the frequency domain estimator). The NMSE of the FD estimator is constant with E_s/N_0, whereas the NMSEs of the low SNR ML based estimator and the subset estimator strongly increase for high E_s/N_0. This can be explained as in the FD estimator, the observations are data-free and therefore the data symbols have no influence on the performance of the estimator, whereas in the low SNR ML based estimator and the subset estimator, the data symbols cause an increasing amount of interference, resulting in the error floors in figures 6.3 and 6.4. From figure 6.5, it also can be observed that for situation 1 the NMSE of the frequency-domain estimator is very close (the curves can not be distinguished from each other) to the theoretical lower bound $L/(M-\nu)$ corresponding to (6.67), as was shown in section 6.6. For situation 2, the difference between the theoretical lower bound $L/(M-\nu)$ and the NMSE of the FD estimator is visible, but they are still close to each other. These results verify the approximations made to obtain the lower bound (6.67).

Figure 6.7 shows the influence of the number of pilots on the MSE of the FD estimator for situation 1. We first assume that the pilot carriers are equally spaced. As expected (see (6.67)), the MSE is essentially equal to $L/(M-\nu) \cdot SNR^{-1}$ for a wide range of M, i.e. the MSE is inversely proportional to the number of pilot carriers. For large M, the pilot spacing (5.4) becomes $\lambda_{\text{pc}} = 2$ (for $N/4 = 256 < M-\nu < N/2 = 512$) and $\lambda_{\text{pc}} = 1$ (for $M-\nu > N/2 = 512$); in that case pilots are not evenly spread over the carriers but grouped in one part of the spectrum, such that (6.139) can no longer be approximated by $SNR \cdot (M-\nu)\mathbf{I}_L$. This causes the peaks in the figure.

The influence of random pilot carrier positions on the frequency domain estimator performance is also shown in figure 6.7. The MSE is shown for 50 randomly generated pilot carrier positions, along with the average over the simulations. For small M, we observe that the performance of the FD estimator strongly depends on the pilot positions, whereas for large M, the FD estimator becomes essentially independent of the pilot positions (as long as the pilot positions are not equally spaced for really large M). This indicates that for small M, the second approximation (i.e. the pilots are evenly distributed, see appendix 6.D for more details) to obtain the MSE lower bound (6.67) is no longer valid. Hence, for small M, fixed, equally spaced pilot positions are preferred. For large M, equally spaced pilot positions are not suitable because of the peak in the MSE. Therefore, at large M, random pilot positions

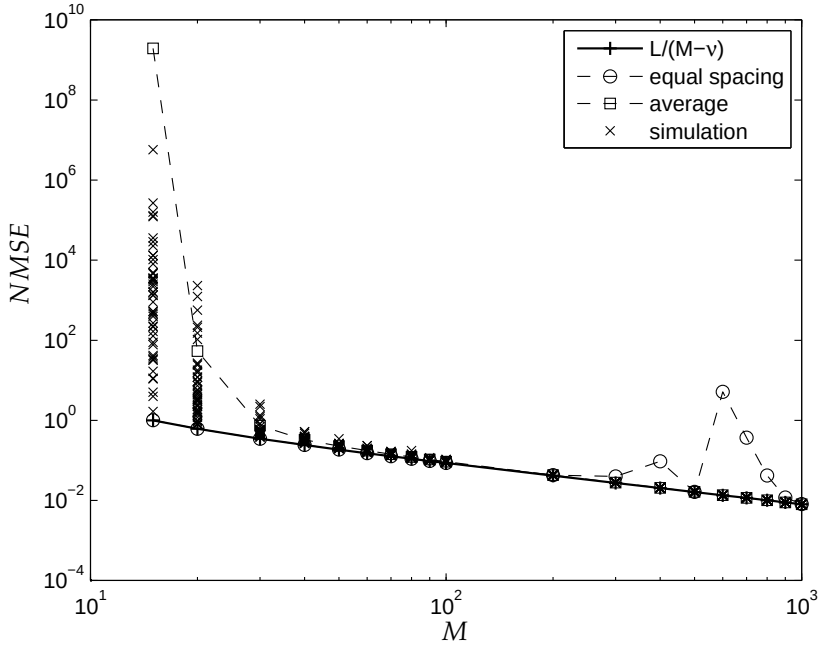

Figure 6.7: NMSE as a function of M , $N = 1024$, $\nu = 7$, $L = 8$.

are advised.

6.10.2.2 Decision Aided Estimators

First, we consider the hard decision aided estimator in the time domain (6.72) from subsection 6.8.1. In figures 6.8 and 6.9, the MSE of the TD hard decision aided estimator (6.72) is shown as function of E_s/N_0 for different number of iterations. The necessary initial estimate of the channel impulse response is provided by the FD DA estimator (6.64). In addition, the MSE (6.65) of the FD DA estimator and the lower bound on the MSE of the TD all pilots estimator (6.57) are shown. The MSE lower bound of the all pilots estimator is equal to $L/\left(N+\nu\right) \cdot SNR^{-1}$. At high E_s/N_0, the MSE of the TD hard decision aided estimator (6.72) converges to the MSE of the TD all pilots estimator, whereas at low E_s/N_0, the decisions on the data symbols are not reliable and disturb the channel estimation; at low E_s/N_0, the unreliable data decisions may even increase the MSE as compared to the FD DA case, as is the case for system 1. For both considered systems, the MSE converges after 2-4 iterations.

Additionally, in figure 6.9 the results are also shown for the coded system. In this case the turbo decoder performs one decoding iteration and provides

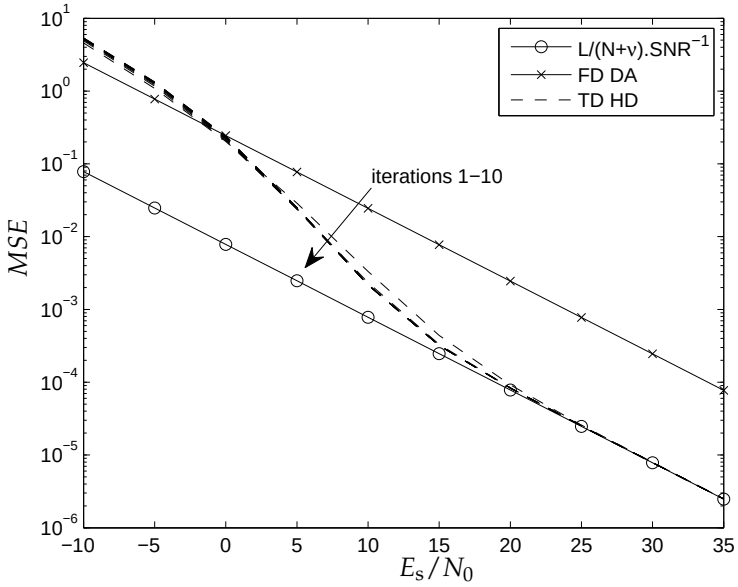

Figure 6.8: MSE as a function of E_s/N_0 , $N = 1024$, $\nu = 7$, $L = 8$, $M = 40$.

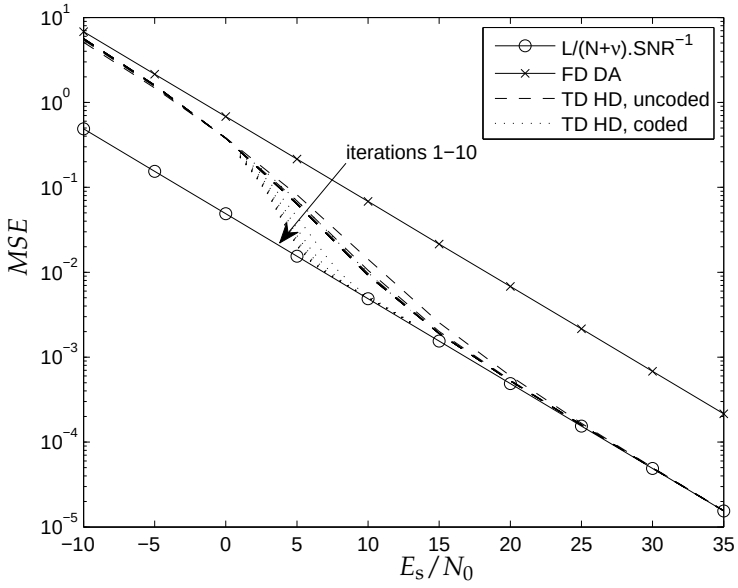

Figure 6.9: MSE as a function of E_s/N_0 , $N = 1024$, $\nu = 100$, $L = 50$, $M = 200$.

the a posteriori probabilities of the coded bits, which are then used to determine the necessary hard decisions on the data symbols. One can see the large improvement of coded transmission: the MSE of the TD hard decision estimator already reaches the MSE (6.57) of the TD all pilots estimator for considerably lower values of E_s/N_0.

Secondly, the performance of the hard decision aided estimator in the frequency domain (6.77) from subsection 6.8.2 is illustrated. Figures 6.10 and 6.11, show the results for the MSE of the FD hard decision aided estimator (6.77) as function of E_s/N_0 for different number of iterations. The FD hard decision aided estimator is initialized by the FD DA estimator (6.64). Additionally, the MSE (6.65) of the FD DA estimator and the lower bound on the MSE (6.57) of the TD all pilots estimator are shown. For situation 1, at high E_s/N_0, the MSE of the FD hard decision aided estimator essentially converges to the MSE of the TD all pilots estimator (when $N \gg \nu$, the MSE of the TD and the FD all pilots estimators become essentially equal), whereas at low E_s/N_0, the decisions on the data symbols are not always very reliable and disturb the channel estimation. At low E_s/N_0, erroneous data decisions even increase the MSE as compared to the FD DA case. If we compare the results for the TD hard decision aided estimator (6.72) from figure 6.8 with the results for the FD hard decision aided estimator (6.77) from figure 6.10, it is difficult to make a distinction between the two estimators for situation 1.

For situation 2, the results are slightly different. For high E_s/N_0 the MSE of the FD hard decision aided estimator (6.77) becomes very close to the MSE (6.57) of the TD all pilots estimator, but both curves do not coincide. In fact, later on in this section it is shown that the MSE of FD hard decision aided estimator converges to the MSE of the FD all pilots estimator (see section 6.7, (6.68)). At low E_s/N_0, the results are not as dramatic as in situation 1: for the considered range of E_s/N_0, the MSE of the FD hard decision aided estimator is never higher than the MSE of the FD DA estimator. Now there is a notable difference between the performance of the TD hard decision aided estimator (6.72) and the performance of the FD hard decision aided estimator (6.77) if we compare figures 6.9 and 6.11. For low E_s/N_0 (≤ -5 dB), the FD hard decision aided estimator does not give any improvement compared to the FD DA estimator (6.64), while the TD hard decision aided estimator clearly yields a visible, albeit small, improvement compared to the FD DA estimator (6.64). Furthermore, the TD hard decision aided estimator reaches the performance of the TD all pilots estimator for a lower value of E_s/N_0. The convergence rate is similar for both hard decision aided estimators, regardless of which situation (situation 1 or 2) we consider: the MSE converges after 2-4 iterations.

For the considered turbo coded system, it can be seen from figure 6.11, that the FD hard decision aided algorithm (6.77) converges to the performance of the FD all pilots estimator for an E_s/N_0 higher than 5 dB. The coded system results in substantial gain in performance compared to the uncoded system.

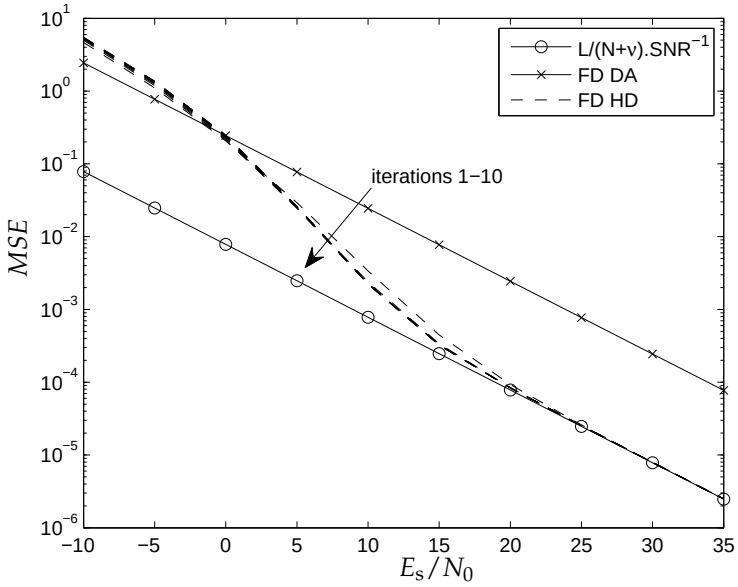

Figure 6.10: MSE as a function of E_s/N_0 , $N = 1024$, $\nu = 7$, $L = 8$, $M = 40$.

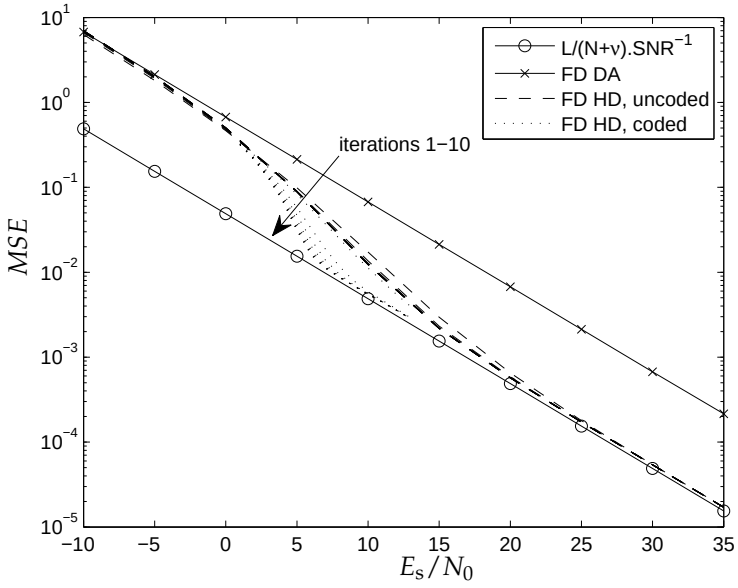

Figure 6.11: MSE as a function of E_s/N_0 , $N = 1024$, $\nu = 100$, $L = 50$, $M = 200$.

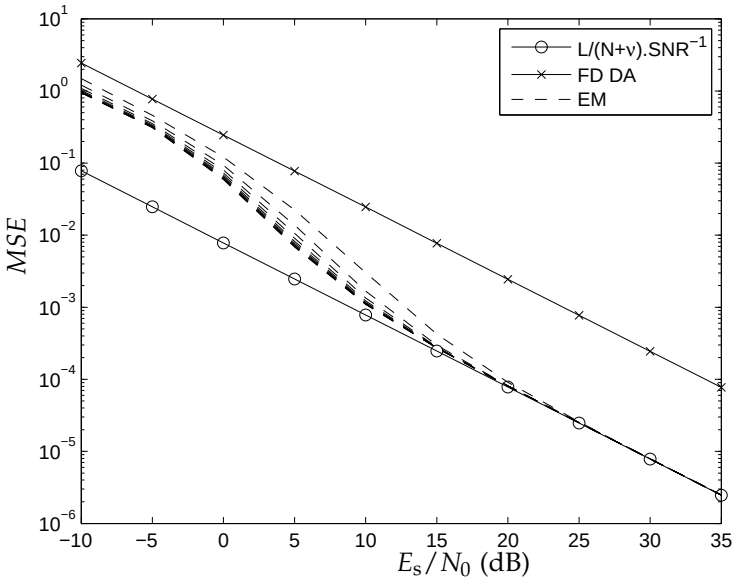

Figure 6.12: MSE as a function of E_s/N_0 , $N = 1024$, $\nu = 7$, $L = 8$, $M = 40$.

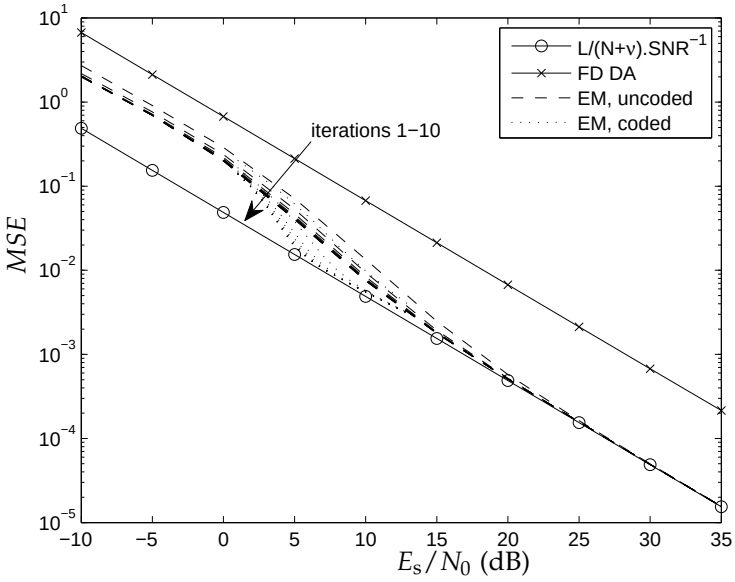

Figure 6.13: MSE as a function of E_s/N_0 , $N = 1024$, $\nu = 100$, $L = 50$, $M = 200$.

Finally, the EM based estimator (6.94) from section 6.9 is considered for channel estimation. The EM algorithm uses the FD DA estimator (6.64) to initialize the estimation, and then iteratively updates the channel estimates in a soft decision aided way. Figures 6.12 and 6.13 show the MSE of the EM algorithm as function of E_s/N_0 for different numbers of iterations. Also shown are the MSE (6.65) of the FD DA estimator and the lower bound on the MSE (6.57) of the TD all pilots estimator are shown. The MSE of the EM algorithm converges to the MSE of the TD all pilots estimator for high E_s/N_0, while for low E_s/N_0 this is not the case. However, there is still an improvement in performance if we compare with the MSE of the FD DA estimator. From figures 6.12 and 6.13, we observe that less than 10 iterations of the EM-algorithm are necessary to obtain convergence. In contrast with the TD and the FD hard decision aided estimators, the EM algorithm always results in a significant improvement in terms of the MSE for the considered range of E_s/N_0. Compared to the hard decision aided estimators (both TD and FD), the EM algorithm results in far better performance at low E_s/N_0. Furthermore, for both considered situations, the EM algorithm reaches the performance of the TD all pilots estimator for a larger range of E_s/N_0 ($>$ 15 dB) than the hard decision aided algorithms ($>$ 20 dB) . The convergence speed of the hard decision aided algorithms is somewhat faster than the convergence speed of the EM algorithm: the hard decision aided algorithms converge after 2-4 iterations while the EM algorithm needs at least 5 iterations (the EM algorithm is known to yield slow convergence).

To apply the EM algorithm in the coded system, the necessary soft information about the data symbols is computed by the receiver every iteration using the a posteriori probabilities of the coded bits, which are provided by the decoder. The MSE results for the EM algorithm applied to the coded system are similar to the results for the hard decision aided algorithms, as can be seen from figure 6.13. The performance of the TD all pilots estimator is already reached for E_s/N_0 equal to 5 dB. The differences in performance between the EM algorithm and the hard decision aided algorithms are not as pronounced as in the uncoded case.

Figure 6.14 compares the NMSE (6.97) (after 10 iterations, for system 2, uncoded) for the different proposed iterative decision aided algorithms with the theoretical lower bounds for the NMSE of the TD (6.57) and the FD all pilots estimator (6.68). Besides that, the lower bound (6.67) and the NMSE of the FD DA estimator (6.65) are also displayed at the figure. For low E_s/N_0, the hard decision aided estimators yield hardly any better performance than the FD DA estimator, this is caused by unreliable decisions on the data symbols. The EM algorithm, on the contrary, still yields in a noteworthy performance improvement. For high E_s/N_0, both the TD hard decision aided and the EM algorithm converge to the NMSE lower bound of the TD all pilots estimator, which equals $L/(N+\nu)$. This is clearly not the case for the FD hard decision

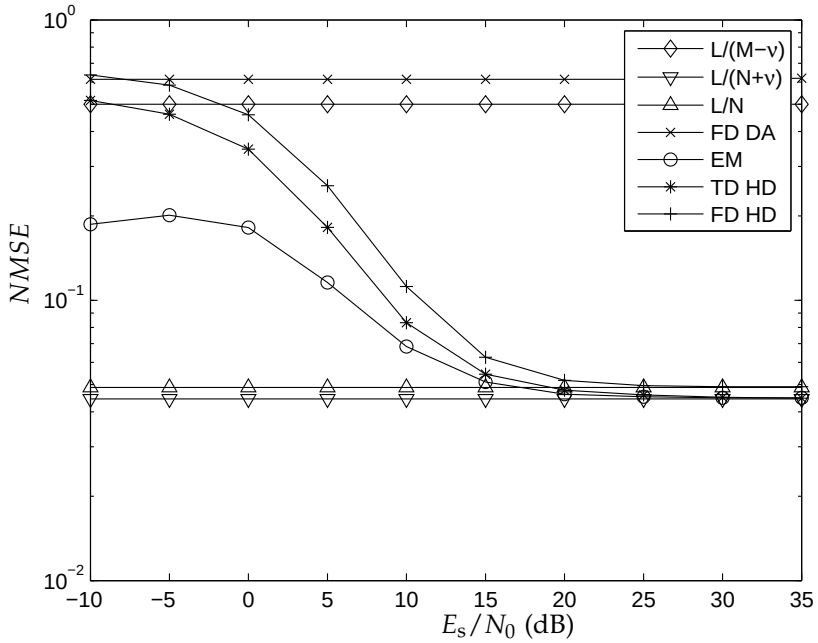

Figure 6.14: NMSE as a function of E_s/N_0 , $N = 1024$, $\nu = 100$, $L = 50$, $M = 200$.

aided estimator. The latter converges to the NMSE lower bound of the FD all pilots estimator, which is given by L/N. We see although the EM based estimator and the TD hard decision aided estimator converge to the same lower bound for high E_s/N_0, the EM algorithm results in a lower NMSE for low to moderate E_s/N_0. We can conclude that the EM algorithm results in the smallest MSE for the whole range of considered E_s/N_0, while the FD hard decision aided estimator has the highest MSE of the three considered iterative decision aided algorithms.

6.10.3 BER

Figures 6.15 and 6.16 show the BER performance when using the EM algorithm (6.94), the TD hard decision aided estimator (6.72), the FD hard decision aided estimator (6.77) and the FD DA estimator (6.64). The performances of a receiver with perfect channel knowledge which uses the ZF detector (4.10) and a receiver with perfect knowledge which uses the FD detector (4.28) are also added.

First we consider uncoded transmission. We see that all three of the it-

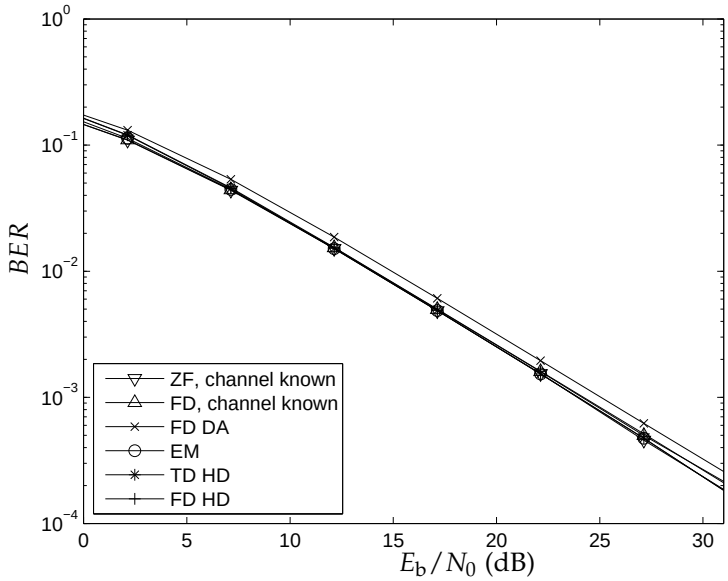

Figure 6.15: BER as a function of E_b/N_0 , $N = 1024$, $\nu = 7$, $L = 8$, $M = 40$.

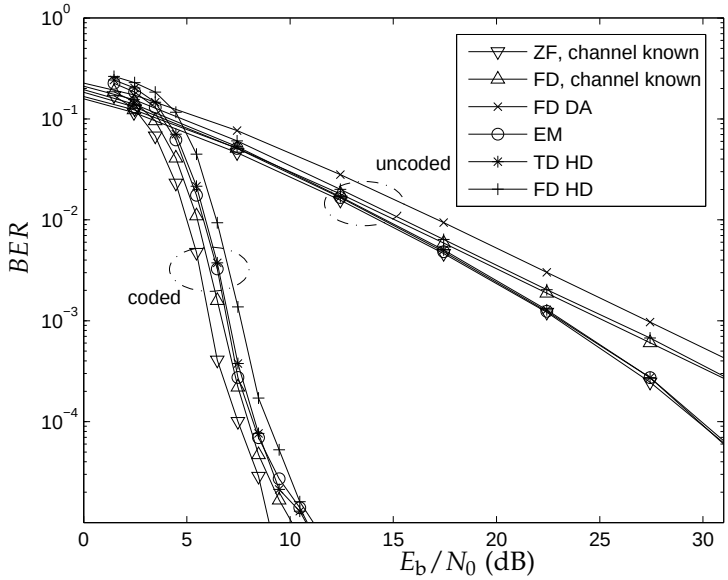

Figure 6.16: BER as a function of E_b/N_0 , $N = 1024$, $\nu = 100$, $L = 50$, $M = 200$.

erative decision aided algorithms converge to the respective curves for perfect channel knowledge. The application of the decision aided estimators to estimate the channel yields a considerable performance gain compared to a receiver which only considers the FD DA estimator. These results highlight again the advantage of considering the ZF detector instead of the FD detector: the receivers that use the ZF detector take advantage of the multipath diversity, especially at high E_b/N_0. The BER curves corresponding to the EM estimator and the TD hard decision aided estimator are almost indistinguishable from one another and from the BER of the receiver with perfect channel knowledge for the considered range of E_b/N_0. Only at very low E_b/N_0 they show a degradation compared to the BER for perfect channel knowledge. For this situation the decision aided algorithms tend to result in the same BER as the FD DA estimator.

The BER results for a turbo coded transmission in system 2 are shown in figure 6.16. The BER results are shown for 11 decoder iterations. First of all, the BER for a ZF receiver and an FD receiver both with perfect channel knowledge are shown. They serve as references for the results of the decision aided algorithms. The EM estimator and the TD hard decision aided estimator yield approximately the same performance. Compared to the uncoded situation, the BER for a receiver with one of both considered TD estimators (the TD hard decision aided algorithm and the EM algorithm), now shows a degradation compared to the TD reference curve, i.e. the ZF receiver (4.10). However, the gap remains small and is less than 1 dB. Similar conclusions can be drawn for the FD hard decision aided estimator, its performance in terms of the BER is somewhat worse than the performance of the FD detector with perfect channel state information. Nevertheless, the degradation is limited and smaller than 1 dB.

6.10.4 Optimization of the Number of Iterations

To obtain the results for the decision aided algorithms in a coded system, we have considered the concept of embedded estimation (see subsection 5.3.1). In order to to keep the computational complexity under control, every estimation operation is followed by only one iteration of the decoder. Typically a channel decoder (e.g. a turbo decoder or an LDPC decoder) needs to perform several iterations before the transmitted information bits can be recovered, while the channel estimate converges usualy much faster. So the computational complexity is further reduced by only updating the channel estimate during the first few iterations.

In this subsection we will illustrate this for the turbo coded transmission in system 2. From figures 6.9, 6.11 and 6.13 it can be seen that the estimates obtained with one of the proposed decision aided estimators (the TD hard decision aided (6.72), the FD hard decision aided (6.77) and the EM based es-

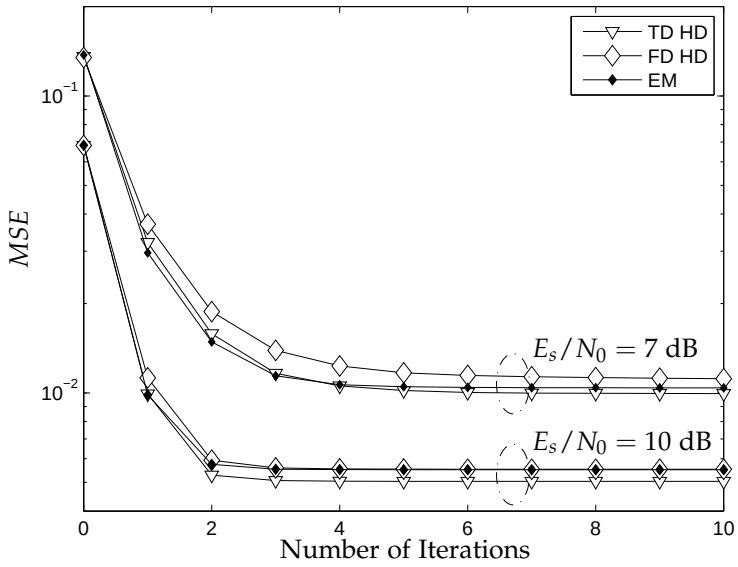

Figure 6.17: MSE as a function of the number of iterations , $N = 1024$, $\nu = 100$, $L = 50$, $M = 200$.

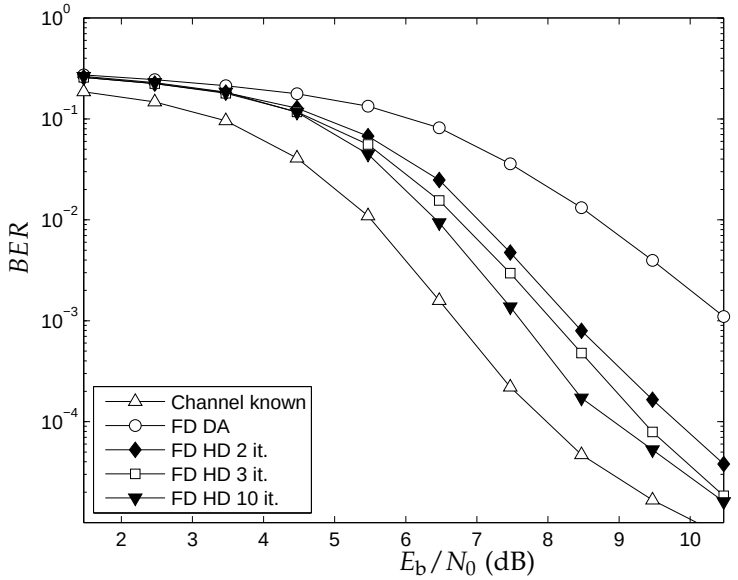

Figure 6.18: BER of the coded system as a function of E_b/N_0 , $N = 1024$, $\nu = 100$, $L = 50$, $M = 200$.

timator (6.94)) converge already after a few iterations. To optimize the computational complexity, we look at the MSE of the proposed decision aided channel estimators as a function of the number of performed updates. The results are shown in figure 6.17. First, We consider the situation for $E_s/N_0 = 10$ dB. The MSEs of the three proposed decision aided estimators show similar trends as a function of the number of iterations. Iteration '0' corresponds to MSE of the FD DA estimator, which provides the initial channel estimate. After two iterations, the channel estimates have almost converged. However, there is still a (very) small improvement in the third iteration. So after two or three iterations, the channel estimate does not need another update and in the remainig iterations only decoder operations need to be performed. To prove that this strategy is valid, we investigate the BER performance of the coded system with a reduced number of channel updates. For the channel estimation, the FD hard decision aided algorithm (6.77) is applied while for the data detection, the FD detector is considered. Figure 6.18 shows the BER as a function of E_b/N_0. To obtain these results we have applied 11 decoder iterations. First of all, we see the BER of a receiver with perfect channel knowledge, which serves as an absolute lower bound. Secondly, the BER curve for the coded system which applies the FD hard decision aided estimator and updates the channel estimate every iteration (which is the same as the one in figure (6.16) is also displayed (labeled as 'FD HD 10 it.' in the figure). Furthermore, the BER results are shown for a receiver which only applies the FD DA estimator to obtain a channel estimate. Finally, the BER results are shown for the situation where the channel estimate is kept constant after 2 and 3 iterations of the FD hard decision aided estimator, respectively. It can be seen, that applying the FD hard decision aided estimation algorithm (even if its only in a few iterations) yields a large improvement in terms of BER compared to a receiver which only performs the FD DA channel estimation. For $E_b/N_0 > 10$ dB, we see that performing only 2 iterations of the FD hard decision aided estimator yields a BER that is already close to the BER of a receiver which updates the channel estimate every decoder iteration, while applying 3 iterations of the FD hard decision aided estimator yields essentially the same performance as performing the channel estimation every iteration. However for E_b/N_0 between 4 and 9 dB, there is a degradation. This means that in this region, the FD DA channel estimation algorithm has not converged after 2-3 iterations. We have added the MSEs of the proposed decision aided algorithms as a function of the number of iterations for $E_s/N_0 = 7$ dB to figure 6.17 to illustrate this: the results show that for $E_s/N_0 = 7$ dB, convergence is only reached after 5 iterations.

6.11 Conclusion

In this chapter we have treated the problem of an unknown channel impulse response by the receiver. First we have derived the GCRB and a low and high SNR limit, to gain more insight on what could be achieved in terms of the MSE for a channel estimator. We have also reviewed two existing channel estimators from the literature. Both suffer from an error floor for moderate to high SNR because they both neglect the presence of the unknown data symbols. For the all pilots case they both result in the same estimator, i.e. the TD all pilot estimator. This estimator has no practical use but serves as a benchmark for the performance of the TD iterative decision aided algorithms.

We have proposed a DA estimator that operates on the FFT outputs of the received signal at the pilot carrier positions. The main advantage of this estimator is the fact that its MSE does not show any error floor for high SNR. The FFT perfectly separates the contributions of the pilot carriers and the data carriers, although the total contribution of the pilot symbols that are transmitted during the guard interval is spread over all the available carriers by applying the FFT. As a result the guard interval pilot contributions on the data carriers are lost, resulting in a MSE that is a little bit higher for low SNRs compared to the MSE of the existing estimators. For the all pilots case, the proposed FD DA estimator yields a second all pilots estimator, i.e. the FD all pilots estimator. Its performance serves as benchmark for the MSE of the iterative FD decision aided algorithm.

The quality of the obtained channel estimate based solely on the presence of pilot symbols, can be further improved by applying decision aided estimation algorithms. We have proposed three of them. A first one operates in the TD and uses every iteration the hard decisions, which are provided by the detector / decoder to update the channel estimate. A second algorithm is based on the EM algorithm and exploits the soft information about the data symbols that is available from the detector / decoder to obtain a better channel estimate. The third one operates in the FD and is in fact an extension of the proposed FD DA estimator: the hard decisions provided by the detector / decoder are considered as extra pilot symbols, so that all the carriers are considered as pilot carriers. An algorithm similar to the FD DA estimator is then applied but considering also the data carriers besides the pilot carriers.

In the last part of this chapter some numerical results are shown to illustrate the performance of the different proposed algorithms. We have considered two reference systems. First of all, the GCRB has been studied to verify the validity of the proposed low and high SNR limits and their lower bounds. Then the proposed estimators are compared with existing methods and theoretical lower bound in terms of the MSE. The FD DA estimator shows good results at high SNRs making it an appropriate candidate to serve as the initialization algorithm of the iterative decision aided estimators. For moderate to

high SNR, the EM based algorithm and the TD hard decision aided estimator both reach the performance of the TD all pilots estimator, while the MSE of the FD hard decision estimator converges to the MSE of FD all pilots estimator. For a coded system, this performance is already achieved for a considerably lower range of SNR values. In terms of the BER, the performance of a receiver with perfect channel state information is reached for uncoded transmission, while in a coded system, there is a small degradation.

The research described in this chapter has yielded 4 publications. The derivation of the FD DA estimator and a study of its performance have been treated in [33]. The TD hard decision aided estimator is proposed in [58], while the EM based algorithm is derived and studied in [59]. Furthermore we have written one journal paper ([34]) that combines the results for the FD DA estimator and the EM based algorithm.

6.A Appendix A

For the computation of the Fisher information matrix $\mathbf{J}_1$ based on the observation $\mathbf{z}_1 = \mathbf{B}_1\mathbf{h} + \mathbf{w}_1'$ (6.18) we need the derivatives of $\ln\det\mathbf{R}_1$ and $\mathbf{R}_1^{-1}$. It is easily shown that those derivatives are given by

$$\frac{\partial}{\partial\Re\{h(l)\}}\ln\det\mathbf{R}_1 = -4\Re\{\eta_l\} \tag{6.98}$$

$$\frac{\partial}{\partial\Im\{h(l)\}}\ln\det\mathbf{R}_1 = 4\Im\{\eta_l\} \tag{6.99}$$

$$\frac{\partial}{\partial\Re\{h(l)\}}\mathbf{R}_1^{-1} = \tilde{\mathbf{Q}}_{\Re,l} \tag{6.100}$$

$$\frac{\partial}{\partial\Im\{h(l)\}}\mathbf{R}_1^{-1} = \tilde{\mathbf{Q}}_{\Im,l} \tag{6.101}$$

where $\tilde{\mathbf{Q}}_{\Re,l}$ and $\tilde{\mathbf{Q}}_{\Im,l}$ are defined (for the definition of λ_n, $\gamma_{l,n}$ and η_l, see (6.20), (6.27) and (6.28)) as

$$\tilde{\mathbf{Q}}_{\Re,l} = \left(\tilde{\mathbf{F}}^H\mathbf{T}_1^H\right)^{-1}\Re\{\mathbf{X}_l\}\left(\mathbf{T}_1\tilde{\mathbf{F}}\right)^{-1} \tag{6.102}$$

$$\tilde{\mathbf{Q}}_{\Im,l} = \left(\tilde{\mathbf{F}}^H\mathbf{T}_1^H\right)^{-1}\Im\{\mathbf{X}_l\}\left(\mathbf{T}_1\tilde{\mathbf{F}}\right)^{-1} \tag{6.103}$$

$$\mathbf{X}_l = \operatorname{diag}\left(\frac{\gamma_{l,n}}{\lambda_n^2}\right). \tag{6.104}$$

The log likelihood function of $\mathbf{h}$ given the observations $\mathbf{z}_1$ is given by

$$\ln p(\mathbf{z}_1|\mathbf{h}) = C - \frac{1}{2}\ln\det\mathbf{R}_1 - (\mathbf{z}_1 - \mathbf{B}_1\mathbf{h})^H\mathbf{R}_1^{-1}(\mathbf{z}_1 - \mathbf{B}_1\mathbf{h}). \tag{6.105}$$

Taking into account equations (6.98)-(6.104), yields for the derivatives of the log likelihood function (6.105) with respect to $\Re\{h(l)\}$ and $\Im\{h(l)\}$:

$$\begin{aligned}\frac{\partial}{\partial\Re\{h(l)\}}\ln p(\mathbf{z}_1|\mathbf{h}) &= 2\Re\{\eta_l\} + 2\Re\left\{(\mathbf{z}_1 - \mathbf{B}_1\mathbf{h})^H\mathbf{R}_1^{-1}\mathbf{B}_1\mathbf{1}_l\right\} \\ &\quad +2(\mathbf{z}_1 - \mathbf{B}_1\mathbf{h})^H\tilde{\mathbf{Q}}_{\Re,l}(\mathbf{z}_1 - \mathbf{B}_1\mathbf{h}) \qquad (6.106)\\ \frac{\partial}{\partial\Im\{h(l)\}}\ln p(\mathbf{z}_1|\mathbf{h}) &= -2\Im\{\eta_l\} - 2\Im\left\{(\mathbf{z}_1 - \mathbf{B}_1\mathbf{h})^H\mathbf{R}_1^{-1}\mathbf{B}_1\mathbf{1}_l\right\} \\ &\quad -2(\mathbf{z}_1 - \mathbf{B}_1\mathbf{h})^H\tilde{\mathbf{Q}}_{\Im,l}(\mathbf{z}_1 - \mathbf{B}_1\mathbf{h}) \qquad (6.107)\end{aligned}$$

for $l = 0, \ldots, L-1$, $\mathbf{1}_l$ is a vector of length L with a one in the l-th position and zeros elsewhere. Hence, the elements of the different blocks of $\mathbf{J}_1$ are

easily obtained:

$$\begin{aligned}
(\mathbf{J}_{1,a})_{l,l'} &= \mathbb{E}\left[\frac{\partial}{\partial \Re\{h(l)\}} \ln p(\mathbf{z}_1|\mathbf{h}) \frac{\partial}{\partial \Re\{h(l')\}} \ln p(\mathbf{z}_1|\mathbf{h})\right] \quad (6.108)\\
&= 2\Re\left\{\left(\mathbf{B}_1^H \mathbf{R}_1^{-1} \mathbf{B}_1\right)_{l,l'}\right\} + 4\Re\{\eta_l\}\Re\{\eta_{l'}\}\\
&\quad +4\Re\{\eta_l\}\,\mathrm{tr}\left(\tilde{\mathbf{Q}}_{\Re,l'}\mathbf{R}_1\right) + 4\Re\{\eta_{l'}\}\,\mathrm{tr}\left(\tilde{\mathbf{Q}}_{\Re,l}\mathbf{R}_1\right)\\
&\quad +4\mathrm{tr}\left(\tilde{\mathbf{Q}}_{\Re,l}\mathbf{R}_1\right)\mathrm{tr}\left(\tilde{\mathbf{Q}}_{\Re,l'}\mathbf{R}_1\right)\\
&\quad +4\mathrm{tr}\left(\tilde{\mathbf{Q}}_{\Re,l}\mathbf{R}_1\tilde{\mathbf{Q}}_{\Re,l'}\mathbf{R}_1\right)
\end{aligned}$$

$$\begin{aligned}
(\mathbf{J}_{1,b})_{l,l'} &= \mathbb{E}\left[\frac{\partial}{\partial \Re\{h(l)\}} \ln p(\mathbf{z}_1|\mathbf{h}) \frac{\partial}{\partial \Im\{h(l')\}} \ln p(\mathbf{z}_1|\mathbf{h})\right] \quad (6.109)\\
&= -2\Im\left\{\left(\mathbf{B}_1^H \mathbf{R}_1^{-1} \mathbf{B}_1\right)_{l,l'}\right\} - 4\Re\{\eta_l\}\Im\{\eta_{l'}\}\\
&\quad -4\Re\{\eta_l\}\,\mathrm{tr}\left(\tilde{\mathbf{Q}}_{\Im,l'}\mathbf{R}_1\right) - 4\Im\{\eta_{l'}\}\,\mathrm{tr}\left(\tilde{\mathbf{Q}}_{\Re,l}\mathbf{R}_1\right)\\
&\quad -4\mathrm{tr}\left(\tilde{\mathbf{Q}}_{\Re,l}\mathbf{R}_1\right)\mathrm{tr}\left(\tilde{\mathbf{Q}}_{\Im,l'}\mathbf{R}_1\right)\\
&\quad -4\mathrm{tr}\left(\tilde{\mathbf{Q}}_{\Re,l}\mathbf{R}_1\tilde{\mathbf{Q}}_{\Im,l'}\mathbf{R}_1\right)
\end{aligned}$$

$$\begin{aligned}
(\mathbf{J}_{1,c})_{l,l'} &= \mathbb{E}\left[\frac{\partial}{\partial \Im\{h(l)\}} \ln p(\mathbf{z}_1|\mathbf{h}) \frac{\partial}{\partial \Im\{h(l')\}} \ln p(\mathbf{z}_1|\mathbf{h})\right] \quad (6.110)\\
&= 2\Re\left\{\left(\mathbf{B}_1^H \mathbf{R}_1^{-1} \mathbf{B}_1\right)_{l,l'}\right\} + 4\Im\{\eta_l\}\Im\{\eta_{l'}\}\\
&\quad +4\Im\{\eta_l\}\,\mathrm{tr}\left(\tilde{\mathbf{Q}}_{\Im,l'}\mathbf{R}_1\right) + 4\Im\{\eta_{l'}\}\,\mathrm{tr}\left(\tilde{\mathbf{Q}}_{\Im,l}\mathbf{R}_1\right)\\
&\quad +4\mathrm{tr}\left(\tilde{\mathbf{Q}}_{\Im,l}\mathbf{R}_1\right)\mathrm{tr}\left(\tilde{\mathbf{Q}}_{\Im,l'}\mathbf{R}_1\right)\\
&\quad +4\mathrm{tr}\left(\tilde{\mathbf{Q}}_{\Im,l}\mathbf{R}_1\tilde{\mathbf{Q}}_{\Im,l'}\mathbf{R}_1\right)
\end{aligned}$$

These expressions can be further simplified: taking into account the definitions of $\mathbf{R}_1$, $\tilde{\mathbf{Q}}_{\Re,l}$ and $\tilde{\mathbf{Q}}_{\Im,l}$ (see (6.19), (6.102) and (6.103)) and one of the properties of the trace of a product of matrices, i.e. $\mathrm{tr}(\mathbf{XY}) = \mathrm{tr}(\mathbf{YX})$, then it follows that

$$\begin{aligned}
\mathrm{tr}\left(\tilde{\mathbf{Q}}_{\Re,l}\mathbf{R}_1\right) &= \mathrm{tr}\left(\Re\{\mathbf{X}_l\}\boldsymbol{\Delta}\right)\\
&= -2\Re\{\eta_l\} \quad (6.111)
\end{aligned}$$

$$\begin{aligned}
\mathrm{tr}\left(\tilde{\mathbf{Q}}_{\Im,l}\mathbf{R}_1\right) &= \mathrm{tr}\left(\Im\{\mathbf{X}_l\}\boldsymbol{\Delta}\right)\\
&= -2\Im\{\eta_l\} \quad (6.112)
\end{aligned}$$

$$\begin{aligned}
\mathrm{tr}\left(\tilde{\mathbf{Q}}_{\Re,l}\mathbf{R}_1\tilde{\mathbf{Q}}_{\Re,l'}\mathbf{R}_1\right) &= \mathrm{tr}\left(\Re\{\mathbf{X}_l\}\boldsymbol{\Delta}\Re\{\mathbf{X}_{l'}\}\boldsymbol{\Delta}\right)\\
&= \sum_{\beta_n \in S_{\mathrm{d}}} \frac{\Re\{\gamma_{l,n}\}\Re\{\gamma_{l',n}\}}{\lambda_n^2} \quad (6.113)
\end{aligned}$$

$$\begin{aligned}\operatorname{tr}\left(\tilde{\mathbf{Q}}_{\Re,l}\mathbf{R}_1\tilde{\mathbf{Q}}_{\Im,l'}\mathbf{R}_1\right) &= \operatorname{tr}\left(\Re\left\{\mathbf{X}_l\right\}\boldsymbol{\Delta}\Im\left\{\mathbf{X}_{l'}\right\}\boldsymbol{\Delta}\right)\\ &= \sum_{\beta_n\in S_{\mathrm{d}}}\frac{\Re\left\{\gamma_{l,n}\right\}\Im\left\{\gamma_{l',n}\right\}}{\lambda_n^2} \qquad (6.114)\end{aligned}$$

$$\begin{aligned}\operatorname{tr}\left(\tilde{\mathbf{Q}}_{\Im,l}\mathbf{R}_1\tilde{\mathbf{Q}}_{\Im,l'}\mathbf{R}_1\right) &= \operatorname{tr}\left(\Im\left\{\mathbf{X}_l\right\}\boldsymbol{\Delta}\Im\left\{\mathbf{X}_{l'}\right\}\boldsymbol{\Delta}\right)\\ &= \sum_{\beta_n\in S_{\mathrm{d}}}\frac{\Im\left\{\gamma_{l,n}\right\}\Im\left\{\gamma_{l',n}\right\}}{\lambda_n^2} \qquad (6.115)\end{aligned}$$

The substitution of equations (6.111) to (6.115) in (6.108), (6.109) and (6.110) yields

$$\begin{aligned}\left(\mathbf{J}_{1,a}\right)_{l,l'} &= 2\Re\left\{\left(\mathbf{B}_1^H\mathbf{R}_1^{-1}\mathbf{B}_1\right)_{l,l'}\right\}\\ &\quad +4\left(\Re\left\{\eta_l\right\}\Re\left\{\eta_{l'}\right\}+\sum_{\beta_n\in S_{\mathrm{d}}}\frac{\Re\left\{\gamma_{l,n}\right\}\Re\left\{\gamma_{l',n}\right\}}{\lambda_n^2}\right) \qquad (6.116)\\ \left(\mathbf{J}_{1,b}\right)_{l,l'} &= -2\Im\left\{\left(\mathbf{B}_1^H\mathbf{R}_1^{-1}\mathbf{B}_1\right)_{l,l'}\right\}\\ &\quad -4\left(\Re\left\{\eta_l\right\}\Im\left\{\eta_{l'}\right\}+\sum_{\beta_n\in S_{\mathrm{d}}}\frac{\Re\left\{\gamma_{l,n}\right\}\Im\left\{\gamma_{l',n}\right\}}{\lambda_n^2}\right) \qquad (6.117)\\ \left(\mathbf{J}_{1,c}\right)_{l,l'} &= 2\Re\left\{\left(\mathbf{B}_1^H\mathbf{R}_1^{-1}\mathbf{B}_1\right)_{l,l'}\right\}\\ &\quad +4\left(\Im\left\{\eta_l\right\}\Im\left\{\eta_{l'}\right\}+\sum_{\beta_n\in S_{\mathrm{d}}}\frac{\Im\left\{\gamma_{l,n}\right\}\Im\left\{\gamma_{l',n}\right\}}{\lambda_n^2}\right). \qquad (6.118)\end{aligned}$$

6.B Appendix B

In this appendix, we are going to compute the Fischer information matrix $\mathbf{J}_2$ which is based on the observation vector $\mathbf{z}_2$ (6.29).

The observation $\mathbf{z}_2$ is independent of the unknown data symbols $\mathbf{a}_{\mathrm{d}}$ and is given by

$$\mathbf{z}_2 = \mathbf{B}_2\mathbf{h} + \mathbf{w}_2. \qquad (6.119)$$

The log likelihood function of $\mathbf{h}$ given the observations $\mathbf{z}_2$ is defined as

$$\ln p\left(\mathbf{z}_2|\mathbf{h}\right) = C - \frac{1}{N_0}\left(\mathbf{z}_2-\mathbf{B}_2\mathbf{h}\right)^H\left(\mathbf{z}_2-\mathbf{B}_2\mathbf{h}\right). \qquad (6.120)$$

The derivative of $\ln p\left(\mathbf{z}_2|\mathbf{h}\right)$ with respect to $\boldsymbol{\theta} = \left(\Re\left\{\mathbf{h}\right\}^T,\ \Im\left\{\mathbf{h}\right\}^T\right)^T$ can easily be obtained:

$$\frac{\partial}{\partial\boldsymbol{\theta}}\ln p\left(\mathbf{z}_2|\mathbf{h}\right) = \frac{2}{N_0}\begin{pmatrix}\Re\left\{\mathbf{B}_2^H\left(\mathbf{z}_2-\mathbf{B}_2\mathbf{h}\right)\right\}\\ \Im\left\{\mathbf{B}_2^H\left(\mathbf{z}_2-\mathbf{B}_2\mathbf{h}\right)\right\}\end{pmatrix}. \qquad (6.121)$$

Using this result the computation of the Fisher information matrix becomes straightforward and yields

$$\mathbf{J}_2 = \frac{2}{N_0} \begin{pmatrix} \Re\left\{\mathbf{B}_2^H \mathbf{B}_2\right\} & -\Im\left\{\mathbf{B}_2^H \mathbf{B}_2\right\} \\ \Im\left\{\mathbf{B}_2^H \mathbf{B}_2\right\} & \Re\left\{\mathbf{B}_2^H \mathbf{B}_2\right\} \end{pmatrix} \tag{6.122}$$

6.C Appendix C

The averaging of $N_0 \mathrm{tr}\left(\left(\mathbf{B}^H \mathbf{B}\right)^{-1}\right)$ (6.40) over the pilot symbols is not that easy, so instead we are going to compute a lower bound. Therefore, we apply Jensen's inequality for the inverse of a matrix [60]:

$$\mathbb{E}\left[\left(\mathbf{B}^H \mathbf{B}\right)^{-1}\right] \geq \left(\mathbb{E}\left[\mathbf{B}^H \mathbf{B}\right]\right)^{-1}. \tag{6.123}$$

The average $\mathbb{E}\left[\mathbf{B}^H \mathbf{B}\right]$ consists of two parts and is given by

$$\mathbb{E}\left[\mathbf{B}^H \mathbf{B}\right] = \mathbb{E}\left[\mathbf{B}_\mathrm{g}^H \mathbf{B}_\mathrm{g}\right] + \mathbb{E}\left[\mathbf{B}_\mathrm{c}^H \mathbf{B}_\mathrm{c}\right]. \tag{6.124}$$

This last equation is valid because the pilot symbols in the guard interval are independent of the pilot symbols on the pilot carriers.

Since the different guard interval pilot symbols are independent of each other, i.e. $\mathbb{E}\left[a_\mathrm{g}(k_1)\, a_\mathrm{g}^*(k_2)\right] = E_s \delta_{k_1 - k_2}$, the first contribution $\mathbb{E}\left[\mathbf{B}_\mathrm{g}^H \mathbf{B}_\mathrm{g}\right]$ reduces to a diagonal matrix with components given by

$$\begin{aligned} \mathbb{E}\left[\left(\mathbf{B}_\mathrm{g}^H \mathbf{B}_\mathrm{g}\right)_{l,l}\right] &= \frac{N}{N+\nu} \sum_{k=0}^{\nu-1} \mathbb{E}\left[\left|a_\mathrm{g}(k)\right|^2\right] \\ &= \frac{N}{N+\nu} \nu E_\mathrm{s}. \end{aligned} \tag{6.125}$$

The components of the second contribution $\mathbb{E}\left[\mathbf{B}_\mathrm{c}^H \mathbf{B}_\mathrm{c}\right]$ consist of sums of terms like $\mathbb{E}\left[s_\mathrm{p}(k_1)\, s_\mathrm{p}^*(k_2)\right]$, $k_1,\ k_2 = 0,\ \ldots,\ N-1$, which are given by:

$$\begin{aligned} \mathbb{E}\left[s_\mathrm{p}(k_1)\, s_\mathrm{p}^*(k_2)\right] &= \frac{1}{N+\nu} \sum_{\alpha_{n_1}, \alpha_{n_2} \in S_\mathrm{p}} \mathbb{E}\left[a_\mathrm{c}(n_1)\, a_\mathrm{c}^*(n_2)\right] e^{j2\pi \frac{\alpha_{n_1} k_1 - \alpha_{n_2} k_2}{N}} \\ &= \frac{E_\mathrm{s}}{N+\nu} \sum_{\alpha_n \in S_\mathrm{p}} e^{j2\pi \frac{\alpha_n (k_1 - k_2)}{N}}. \end{aligned} \tag{6.126}$$

For the comb-type pilot structure (see section 5.2), this yields:

$$\mathbb{E}\left[s_\mathrm{p}(k_1)\, s_\mathrm{p}^*(k_2)\right] = \frac{M-\nu}{N+\nu} E_\mathrm{s} \delta_{k_1 - k_2}. \tag{6.127}$$

For the random pilot structure, this is not entirely true, but we will assume that (6.127) is valid anyway. Using this last result, one can show that $\mathbb{E}\left[\mathbf{B}_{\mathrm{c}}^{H}\mathbf{B}_{\mathrm{c}}\right]$ is a diagonal matrix with components defined as:

$$\begin{aligned}\mathbb{E}\left[\left(\mathbf{B}_{\mathrm{c}}^{H}\mathbf{B}_{\mathrm{c}}\right)_{l,l}\right] &= \sum_{k=0}^{N-1}\mathbb{E}\left[\left|s_{\mathrm{p}}\left(k_1\right)\right|^2\right] \\ &= \frac{N}{N+\nu}\left(M-\nu\right)E_{\mathrm{s}}. \end{aligned} \tag{6.128}$$

Substituting (6.125) and (6.128) in (6.124) yields

$$\mathbb{E}\left[\mathbf{B}^{H}\mathbf{B}\right] = \frac{N}{N+\nu}ME_{\mathrm{s}}\mathbf{I}_{L}. \tag{6.129}$$

So the average of $GCRB_{\mathrm{low}}$ over the pilot symbols is lower bounded by

$$\mathbb{E}\left[GCRB_{\mathrm{low}}\right] \geq \frac{L}{M}\frac{N+\nu}{N}\frac{N_0}{E_{\mathrm{s}}}. \tag{6.130}$$

6.D Appendix D

The averaging of $\left(\tilde{\mathbf{B}}_2^{H}\tilde{\mathbf{R}}_2^{-1}\tilde{\mathbf{B}}_2\right)^{-1}$ (6.65) over the pilot symbols is really difficult, so as an alternative, we are going to compute a lower bound for the expectation $\mathbb{E}\left[\left(\tilde{\mathbf{B}}_2^{H}\tilde{\mathbf{R}}_2^{-1}\tilde{\mathbf{B}}_2\right)^{-1}\right]$. Since the inverse of a matrix is a matrix convex function, Jensen's inequality for matrices can be applied [60]:

$$\mathbb{E}\left[\left(\tilde{\mathbf{B}}_2^{H}\tilde{\mathbf{R}}_2^{-1}\tilde{\mathbf{B}}_2\right)^{-1}\right] \geq \left(\mathbb{E}\left[\tilde{\mathbf{B}}_2^{H}\tilde{\mathbf{R}}_2^{-1}\tilde{\mathbf{B}}_2\right]\right)^{-1}. \tag{6.131}$$

The matrix $\mathbb{E}\left[\tilde{\mathbf{B}}_2^{H}\tilde{\mathbf{R}}_2^{-1}\tilde{\mathbf{B}}_2\right]$ consists of two parts, one part corresponding to the pilot carriers and one part corresponding to the pilot symbols transmitted during the guard interval: $\mathbb{E}\left[\tilde{\mathbf{B}}_2^{H}\tilde{\mathbf{R}}_2^{-1}\tilde{\mathbf{B}}_2\right] = \mathbb{E}\left[\tilde{\mathbf{B}}_{2,\mathrm{c}}^{H}\tilde{\mathbf{R}}_2^{-1}\tilde{\mathbf{B}}_{2,\mathrm{c}}\right] + \mathbb{E}\left[\tilde{\mathbf{B}}_{2,\mathrm{g}}^{H}\tilde{\mathbf{R}}_2^{-1}\tilde{\mathbf{B}}_{2,g}\right]$. For the computation of $\mathbb{E}\left[\tilde{\mathbf{B}}_{2,\mathrm{c}}^{H}\tilde{\mathbf{R}}_2^{-1}\tilde{\mathbf{B}}_{2,\mathrm{c}}\right]$ and $\mathbb{E}\left[\tilde{\mathbf{B}}_{2,\mathrm{g}}^{H}\tilde{\mathbf{R}}_2^{-1}\tilde{\mathbf{B}}_{2,g}\right]$, we approximate $\tilde{\mathbf{R}}_2$ (6.60) by a diagonal matrix:

$$\tilde{\mathbf{R}}_2 \approx N_0\frac{N+\nu}{N}\mathbf{I}_{M-\nu} \tag{6.132}$$

because typically $\left(\tilde{\mathbf{R}}_2\right)_{n,m} \ll \left(\tilde{\mathbf{R}}_2\right)_{n,n}$ for $m \neq n$. In that case, $\mathbb{E}\left[\tilde{\mathbf{B}}_{2,\mathrm{c}}^{H}\tilde{\mathbf{R}}_2^{-1}\tilde{\mathbf{B}}_{2,\mathrm{c}}\right]$ and $\mathbb{E}\left[\tilde{\mathbf{B}}_{2,\mathrm{g}}^{H}\tilde{\mathbf{R}}_2^{-1}\tilde{\mathbf{B}}_{2,\mathrm{g}}\right]$ are essentially equal to

$$\left(\mathbb{E}\left[\tilde{\mathbf{B}}_{2,\mathrm{c}}^{H}\tilde{\mathbf{R}}_2^{-1}\tilde{\mathbf{B}}_{2,\mathrm{c}}\right]\right)_{l,l'} = \left(\frac{N}{N+\nu}\right)^2\frac{E_{\mathrm{s}}}{N_0}\sum_{\alpha_n\in S_{\mathrm{p}}}e^{j2\pi\frac{\alpha_n\left(l-l'\right)}{N}} \tag{6.133}$$

and

$$\left(\mathbb{E}\left[\tilde{\mathbf{B}}_{2,\mathrm{g}}^{H}\tilde{\mathbf{R}}_{2}^{-1}\tilde{\mathbf{B}}_{2,\mathrm{g}}\right]\right)_{l,l'} = \left(\frac{N}{N+\nu}\right)^{2}\frac{1}{N_0}\frac{1}{N}\sum_{\alpha_n\in S_{\mathrm{p}}}\sum_{m,m'=0}^{\nu-1}e^{j2\pi\frac{\alpha_n(m-m')}{N}}R_{\mathrm{g}}\left(m,m',l,l'\right), \quad (6.134)$$

where $R_{\mathrm{g}}\left(m,m',l,l'\right) = \mathbb{E}\left[a_{\mathrm{g}}^{*}\left(|m-l|_{\nu}\right)a_{\mathrm{g}}\left(|m'-l'|_{\nu}\right)\right]$. When the comb-type pilot structure (see section 5.2) is considered and $M-\nu$ divides N, the contribution of the pilot carriers (6.133) reduces to a diagonal matrix:

$$\left(\mathbb{E}\left[\tilde{\mathbf{B}}_{2,\mathrm{c}}^{H}\tilde{\mathbf{R}}_{2}^{-1}\tilde{\mathbf{B}}_{2,\mathrm{c}}\right]\right)_{l,l'} = \left(\frac{N}{N+\nu}\right)^{2}\left(M-\nu\right)\frac{E_{\mathrm{s}}}{N_0}\delta_{l-l'}. \quad (6.135)$$

The computation of the contribution from the guard interval pilot symbols (6.134) is somewhat more complex: when the pilot symbols are equally spaced over the carriers (comb-type pilot structure), $\left(\mathbb{E}\left[\tilde{\mathbf{B}}_{2,\mathrm{g}}^{H}\tilde{\mathbf{R}}_{2}^{-1}\tilde{\mathbf{B}}_{2,\mathrm{g}}\right]\right)_{l,l'}$ can be expressed as

$$\left(\mathbb{E}\left[\tilde{\mathbf{B}}_{2,\mathrm{g}}^{H}\tilde{\mathbf{R}}_{2}^{-1}\tilde{\mathbf{B}}_{2,\mathrm{g}}\right]\right)_{l,l'} = \left(\frac{N}{N+\nu}\right)^{2}\frac{1}{N_0}\frac{1}{N}\sum_{m,m'=0}^{\nu-1}\frac{1-e^{j2\pi(m-m')}}{1-e^{j2\pi\frac{(m-m')}{M-\nu}}}R_{\mathrm{g}}\left(m,m',l,l'\right). \quad (6.136)$$

When the total number of pilot carriers $M-\nu$ is larger than $\nu-1$, (6.136) reduces to

$$\left(\mathbb{E}\left[\tilde{\mathbf{B}}_{2,\mathrm{g}}^{H}\tilde{\mathbf{R}}_{2}^{-1}\tilde{\mathbf{B}}_{2,\mathrm{g}}\right]\right)_{l,l'} = \left(\frac{N}{N+\nu}\right)^{2}\frac{1}{N_0}\frac{M-\nu}{N}\sum_{m=0}^{\nu-1}\mathbb{E}\left[a_{\mathrm{g}}^{*}\left(|m-l|_{\nu}\right)a_{\mathrm{g}}\left(\left|m-l'\right|_{\nu}\right)\right]. \quad (6.137)$$

Now we only need to determine for which values of the indices l and l' the expectation $\mathbb{E}\left[a_{\mathrm{g}}^{*}\left(|m-l|_{\nu}\right)a_{\mathrm{g}}\left(|m-l'|_{\nu}\right)\right]$ is not equal to zero: When the duration of the guard interval ν exceeds the length of the channel impulse response (which is normally the case), $\nu > L-1$, we only have a contribution for $l=l'$, so $\mathbb{E}\left[\tilde{\mathbf{B}}_{2,\mathrm{g}}^{H}\tilde{\mathbf{R}}_{2}^{-1}\tilde{\mathbf{B}}_{2,\mathrm{g}}\right]$ reduces to a diagonal matrix with diagonal elements given by

$$\left(\mathbb{E}\left[\tilde{\mathbf{B}}_{2,\mathrm{g}}^{H}\tilde{\mathbf{R}}_{2}^{-1}\tilde{\mathbf{B}}_{2,\mathrm{g}}\right]\right)_{l,l} = \frac{\nu}{N}\left(\frac{N}{N+\nu}\right)^{2}\left(M-\nu\right)\frac{E_{\mathrm{s}}}{N_0}. \quad (6.138)$$

In the special case that $\nu = L-1$, we have besides contributions for $l=l'$, also contributions for $(l,l') = (0,L-1)$ and $(L-1,0)$ which are also equal to the right hand side of (6.138).

Combining the obtained results for the expectations $\mathbb{E}\left[\tilde{\mathbf{B}}_{2,\mathrm{c}}^H \tilde{\mathbf{R}}_2^{-1} \tilde{\mathbf{B}}_{2,\mathrm{c}}\right]$ and $\mathbb{E}\left[\tilde{\mathbf{B}}_{2,\mathrm{g}}^H \tilde{\mathbf{R}}_2^{-1} \tilde{\mathbf{B}}_{2,\mathrm{g}}\right]$ yields[1]

$$\begin{aligned} \mathbb{E}\left[\tilde{\mathbf{B}}_2^H \tilde{\mathbf{R}}_2^{-1} \tilde{\mathbf{B}}_2\right] &= \left(\frac{N}{N+\nu}\right)^2 (M-\nu)\left(1+\frac{\nu}{N}\right)\frac{E_\mathrm{s}}{N_0}\mathbf{I}_L \\ &= \frac{N}{N+\nu}(M-\nu)\frac{E_\mathrm{s}}{N_0}\mathbf{I}_L \end{aligned} \tag{6.139}$$

[1]For the special case that $\nu = L-1$ we have shown that $\mathbb{E}\left[\tilde{\mathbf{B}}_2^H \tilde{\mathbf{R}}_2^{-1} \tilde{\mathbf{B}}_2\right]$ has 2 non-diagonal elements, but they are much smaller than the diagonal elements, so we neglect them.

7
Time Delay Offset Estimation

In this chapter, we tackle the problem of time delay offset estimation. First we start with investigating what the influence is of a time delay offset on the performance of a KSP-OFDM system. Then some algorithms which are proposed in a TDS-OFDM context, are reviewed. These algorithms, however, are not suitable for transmission over dispersive channels as their performance degrades rapidly. This is mainly caused by the fact that they cannot exploit pilots in the frequency domain and the presence of the unknown data symbols. Therefore, we propose some new algorithms to estimate the time delay offset, that exploit both the pilots in the time domain and in the frequency domain. These new algorithms turn out to be more robust to channel dispersion. First we start from the observations in the TD and derive two low SNR approaches to obtain an estimation algorithm for the time delay offset. These two algorithms exhibit a high computational load, so we also search for alternatives. An important simplification is obtained by only considering the TD pilot symbols which are transmitted during the guard interval. However, the performance of this algorithm is not always satisfactory. Finally a pilot aided FD algorithm is derived. The performance of the proposed algorithms is illustrated by means of some simulation results in the last part of this chapter.

A comparison is made with the existing algorithms for TDS-OFDM.

7.1 Problem Statement

An important issue with OFDM systems is their much higher sensitivity to some synchronization errors, as compared to single carrier systems [36], so synchronizing with the transmitter is an important task of the OFDM receiver. One aspect of the synchronization process comprises finding the starting point of the OFDM symbol; accurate time delay offset estimation is necessary because time offsets can cause inter carrier interference (ICI) and IBI [61,62].

Several time delay offset estimation algorithms for CP-OFDM have already been proposed in the literature. In [63], the maximum likelihood (ML) estimator for a time delay offset in the presence of additive white Gaussian noise (AWGN) has been presented. The algorithm exploits the redundancy of the cyclic prefix and the pilot symbols on the carriers. The blind estimator of [64] is a special case of the previous estimator and only makes use of the correlation of the cyclic prefix and the last samples of the transmitted OFDM block. In [65], the authors propose an estimator that uses a specially designed training symbol for the AWGN channel. However in multipath fading environments, the mentioned estimators result in inaccurate estimates with large biases. To guarantee a good performance, the OFDM system would require a long guard interval, which reduces efficiency in terms of both power and bandwidth. In the literature, there are some contributions which take the frequency selective nature of the channel in to account. In [66] and [36], the performance from the algorithm from [65] is improved by optimizing the specially designed training symbol, but the performance can still suffer from a dispersive channel. The authors of [67] derive the ML time delay offset estimator in the case of a dispersive channel under the assumption of perfect channel knowledge. The algorithm exploits the cyclic prefix only. However, in practice this estimator is not applicable because it is very difficult to obtain a channel estimate without knowledge about the time delay. This estimator can only serve as a lower bound on the performance of an estimator which does not assume any knowledge about the channel. In [68], a coarse time delay estimate is obtained by applying the correlation method based on the cyclic prefix. The coarse estimate is further improved by employing a path time delay estimation method. The algorithm can show a good performance, but it is not that easy to choose an appropriate value for the correction term of the coarse estimate and for the threshold value for the path time delay estimation.

For a TDS-OFDM system, there exist some time delay offset estimation algorithms [69,70]. They estimate the time delay offset exploiting the presence of the pseudo noise sequence which is transmitted during the guard interval. A drawback of both algorithms, is that they both assume a flat fading channel for the transmission, which is hardly realistic in an OFDM context. As a result,

their performance is deteriorated in a dispersive environment.

For a KSP-OFDM system, no research has been done about algorithms for the timing synchronization problem. Our goal is to develop timing offset estimation algorithms that exploit both the pilot symbols in the guard interval and the pilot symbols on the pilot carriers.

As already briefly mentioned in chapter 2 (see subsection 2.4.1), the receiver considers a vector $\mathbf{z}'$ consisting of $2(N+\nu)+L-2$ received samples. This vector $\mathbf{z}'$ contains the total contribution from only one OFDM block, i.e. the OFDM block with index $i=0$, besides partial contributions from adjacent blocks (see figure 2.10). Regardless of which detector is used, the receiver needs a window of $N+\nu$ received samples that contains all the $N+L-1$ samples that depend on the data symbols transmitted during block $i=0$. The starting point k_0 of this block of samples in the received signal vector $\mathbf{z}'$ is unknown and has to be estimated. Note that we consider here, in this chapter, time delay offset estimation under the assumption that there is no carrier frequency offset.

7.2 Effect of a Timing Error

Suppose we have obtained an estimate of the time delay, which we denote as $\hat{k}_0$. First of all we investigate the influence of a timing error $k_0-\hat{k}_0$ on the performance of a KSP-OFDM system. We consider two types of timing errors: i) timing errors to the right ($\hat{k}_0 > k_0$) and ii) timing errors to the left ($\hat{k}_0 < k_0$).

Case of timing errors to the right ($\hat{k}_0 > k_0$)

In this case, an error of $\hat{k}_0-k_0$ samples to the right side has occurred. The receiver selects the $N+\nu$ samples $z'(\hat{k}_0),\ \ldots,\ z'(\hat{k}_0+N+\nu-1)$. As a result, the $\hat{k}_0-k_0$ samples $z'(k_0),\ \ldots,\ z'(\hat{k}_0-1)$ are missed by the receiver for the detection of the OFDM block $i=0$. For the FD detector, this means that the orthogonality of the carriers can not be restored anymore, which causes ICI. For the other detectors, this means that some information about the data symbols from block $i=0$ is lost.
Furthermore, the block of $N+\nu$ samples contains the $\hat{k}_0-k_0$ samples $z'(k_0+N+\nu),\ \ldots,\ z'(\hat{k}_0+N+\nu-1)$, which depend on the data symbols from the next OFDM block $i=1$. The presence of these samples causes IBI.

Case of timing errors to the left ($\hat{k}_0 < k_0$)

In this case, an error of $\hat{k}_0-k_0$ samples to the left side has occurred. The consequences are not always as bad as for timing errors to the right. The receiver considers the $N+\nu$ samples $z'(\hat{k}_0),\ \ldots,\ z'(\hat{k}_0+N+\nu-1)$ for data detection. The channel impulse response has a duration of L samples, so the last received sample which depends on the data symbols of the previous OFDM block $i=-1$ is $z'(k_0-\nu+L-2)$. This means that if $\hat{k}_0$ is greater

than or equal to $k_0 - \nu + L - 1$, the $N + \nu$ samples that are selected to perform the data detection are independent of any other OFDM blocks. After the removal of the guard interval contributions, we have in fact a shifted version of the original samples $\mathbf{z}_0$. The data symbols can be detected without any interference from other OFDM blocks. For the FD data detector, we see a similar effect as for a CP-OFDM system [4,62]: the timing error $k_0 - \hat{k}_0$ causes a phase rotation equal to $\exp\left(-j2\pi n(k_0 - \hat{k}_0)/N\right)$ at the output of the FFT at the n-th carrier, but does not cause any ICI in the FD. This phase rotation in the frequency domain still has to be corrected. However, its estimation can be incorporated in the channel estimation. Let us define the $\nu + 1$ size vector $\mathbf{h}'$, given by:

$$\mathbf{h}' = \begin{pmatrix} \mathbf{0}_{(k_0 - \hat{k}_0)\times 1} \\ \mathbf{h} \\ \mathbf{0}_{(\nu - L + 1 + \hat{k}_0 - k_0)\times 1} \end{pmatrix}, \tag{7.1}$$

where we have added zeroes before and / or after the actual channel impulse response $\mathbf{h}$, and $\hat{k}_0$ is asumed to be in the desired range $[k_0 - \nu + L - 1,\ k_0]$. It is easily verified that the N-point FFT of $\mathbf{h}'$ is given by

$$\begin{aligned} H'(n) &= \sum_{l'=0}^{\nu} h'\left(l'\right) e^{-j2\pi \frac{nl'}{N}} \\ &= \sum_{l=0}^{L-1} h(l)\, e^{-j2\pi \frac{nl}{N}} e^{-j2\pi \frac{(k_0 - \hat{k}_0)n}{N}} \\ &= H(n)\, e^{-j2\pi \frac{(k_0 - \hat{k}_0)n}{N}}, \end{aligned} \tag{7.2}$$

where we have used the definition of $H(n)$ (4.17). We see indeed that every new FD channel coefficient $H'(n)$ is given by a rotation of $H(n)$ over an angle $-2\pi n(k_0 - \hat{k}_0)/N$. Hence, by incorporating the phase rotation estimation in the channel estimation, the KSP-OFDM system is made robust to these small timing errors.

When the time delay estimate $\hat{k}_0$ is less than $k_0 - \nu + L - 1$, the receiver selects the $N + \nu$ samples $z'(\hat{k}_0),\ \ldots,\ z'(\hat{k}_0 + N + \nu - 1)$. The $k_0 - \nu + L - 1 - \hat{k}_0$ samples $z'(\hat{k}_0 + N + \nu),\ \ldots,\ z'(k_0 + N + L - 2)$ are missing for the detection of the OFDM block $i = 0$. This results in similar effects as for timing errors to the right. In the FD, the orthogonality of the carriers is lost, which causes ICI. The missing samples are replaced by the samples $z'(\hat{k}_0),\ \ldots,\ z(k_0 - \nu + L - 2)$, which all depend on the previous OFDM block $i = -1$. As a consequence, the performance of the considered detectors is degraded by IBI.

Figure 7.1 shows the three possibilities for $\hat{k}_0$: i) when the estimate of k_0 is larger then k_0 ($\hat{k}_0 > k_0$), some samples which are necessary to be able to restore the orthogonality of the carriers are missed and they are replaced

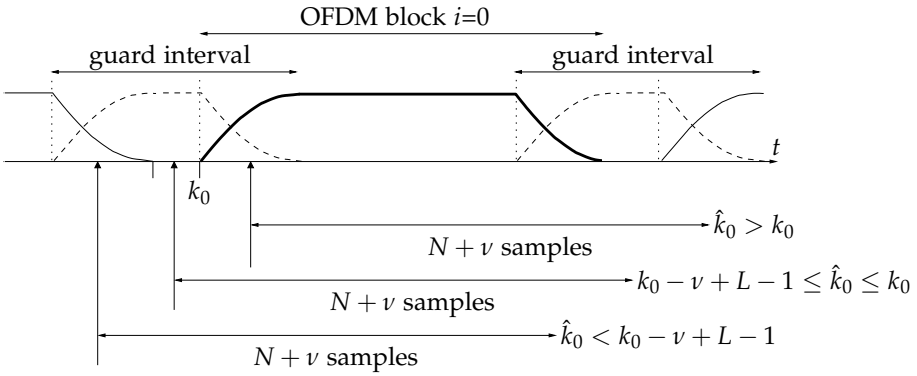

Figure 7.1: The earliest and latest possible timing instants for the detection of the OFDM block $i = 0$.

with samples belonging to the next OFDM block. This results in IBI and ICI. ii) When the estimate of k_0 is in the interval $[k_0 - \nu + L - 1,\ k_0]$, there is no IBI and no ICI in the FD. As a result, optimal detection without any loss in performance can be guaranteed. iii) Finally when the estimate of k_0 is too small ($\hat{k}_0 < k_0 - \nu + L - 1$), again some essential samples are missed which are necessary to be able to restore the orthogonality of the carriers and additionally some samples belonging to the previous OFDM block are considered for the detection of the OFDM block $i = 0$, which results in ICI and IBI.

7.3 Existing algorithms for TDS-OFDM

In this section we review some existing algorithms for TDS-OFDM because with some modifications they can also be applied in a KSP-OFDM system.

7.3.1 Slide Auto-correlation Based Estimation

In [69], the authors show that the guard interval in a TDS-OFDM can be interpreted as an OFDM block preceded by a CP (see also figure 7.2). As we have already described in subsection 1.4.1, the guard interval consists of a pseudo noise sequence of length N_{PN} preceded by a prefix consisting of the N_{pre} last samples of the pseudo noise sequence, and followed by a postfix consisting of the N_{post} first samples of the pseudo noise sequence. This means that the first $N_{\mathrm{CP}} = N_{\mathrm{pre}} + N_{\mathrm{post}}$ samples of the guard interval are equal to the last N_{CP} samples. The last N_{PN} samples of the guard interval are denoted as $\mathbf{a}'_{\mathrm{PN}}$ and they form a cyclic shifted version of the original pseudo noise sequence (which we denote as $\mathbf{a}_{\mathrm{PN}}$). If we regard $\mathbf{a}'_{\mathrm{PN}}$ as an OFDM block of an OFDM

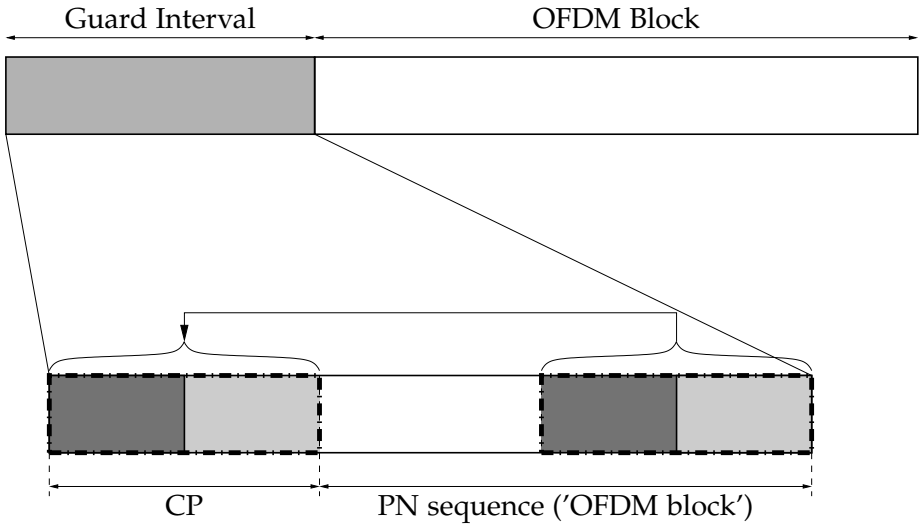

Figure 7.2: The interpretation of the TDS-OFDM guard interval as a CP-OFDM block.

system with N_{PN} carriers, then $\mathbf{a}'_{\mathrm{PN}}$ corresponds to the N_{PN}-point IFFT outputs of the frequency domain symbols $\tilde{\mathbf{a}}'_{\mathrm{PN}}$, where the samples $\tilde{a}'_{\mathrm{PN}}(n)$ are defined as

$$\tilde{a}'_{\mathrm{PN}}(n) = \frac{1}{\sqrt{N_{\mathrm{PN}}}} \sum_{k=0}^{N_{PN}-1} a'_{\mathrm{PN}}(k)\, e^{-j2\pi \frac{kn}{N_{\mathrm{PN}}}},\ n = 0,\ \ldots,\ N_{\mathrm{PN}}. \tag{7.3}$$

Now we can consider the guard interval as a CP-OFDM training block. The authors from [69] apply a slide auto-correlation based on the cyclic prefix from [64] to obtain an estimate of the time delay offset. For our KSP-OFDM system, the slide auto-correlation is given by

$$R_{\mathrm{c}}(k) = \sum_{l=0}^{N_{\mathrm{CP}}-1} \left(z'(k_0+N+l)\right)^* z'(k_0+N+l+N_{\mathrm{PN}}). \tag{7.4}$$

For a flat fading channel, the magnitude of $R_{\mathrm{c}}(k)$ shows a maximum at the correct value of the time delay offset, so the estimate of k_0 is then given by

$$\hat{k}_0 = \arg\max_k |R_{\mathrm{c}}(k)|. \tag{7.5}$$

The major drawback of this algorithm is that it has been derived under the assumption that the channel is nondispersive and that the transmitted signal is only affected by AWGN [64]. A multipath fading channel degrades the performance of the algorithm. This algorithm can only be applied in a KSP-OFDM system if the time domain pilot sequence $\mathbf{a}_{\mathrm{g}}$ is adjusted so that it has

the same structure as the guard interval in a TDS-OFDM system (see figure 7.2).

7.3.2 Composite PN-Correlation (CPC) Based Estimation

A second existing time delay estimation method can be found in [70]. This algorithm computes the pointwise conjugate product of two phase-shifted PN-correlations. First we compute the N_{PN}-length linear correlation between the received signal samples of $\mathbf{z}'$ and $\mathbf{a}_{\text{PN}}$:

$$C_1(k) = \sum_{l=0}^{N_{\text{PN}}-1} z'\left(k + N + N_{\text{pre}} + l\right) \left(a_{\text{PN}}(l)\right)^* . \tag{7.6}$$

For a flat fading channel and at the correct timing instant k_0, the correlation $C_1(k_0)$ reaches a peak. This peak corresponds to the correlation of the PN sequence and the part of the received block $i = 0$ which corresponds tot the transmitted PN sequence $\mathbf{a}_{\text{PN}}$. Similarly, another N_{PN}-length linear correlation is computed but this time with another part of the received signal:

$$C_2(k) = \sum_{l=0}^{N_{\text{PN}}-1} z'\left(k + 2N + \nu + N_{\text{pre}} + l\right) \left(a_{\text{PN}}(l)\right)^* . \tag{7.7}$$

This second correlation also reaches the same peak at the correct timing instant k_0 for a flat fading channel, and is the correlation of the PN sequence $\mathbf{a}_{\text{PN}}$ and the part of the received block $i = 1$ which corresponds to the transmitted PN sequence $\mathbf{a}_{\text{PN}}$.

The Composite PN-correlation is now defined as

$$R_{\text{CPC}}(k) = \left(C_1(k)\right)^* C_2(k) . \tag{7.8}$$

The time delay estimate is then given by

$$\hat{k}_0 = \arg\max_k \left|R_{\text{CPC}}(k)\right| . \tag{7.9}$$

This algorithm has the same major drawback as the previous one: it is derived under the assumption that the channel is a flat fading channel. Again, the multipath fading will degrade the performance of the algorithm. This algorithm can be applied in a KSP-OFDM without any modifications.

7.4 Pilot Aided TD Estimation

In this section an estimator for k_0 is derived, starting from the joint likelihood function of k_0 and $\mathbf{h}$ for the observation $\mathbf{z}'$. To simplify the derivation of the estimator, we assume that $\mathbf{z}'$ only contains noise besides the contribution of the considered transmitted OFDM block $\mathbf{s}_0$[1]. We define the vector

[1]This assumption is only used to derive the estimator, for the simulations we will consider a continuous transmission of consecutive OFDM blocks.

$\mathbf{z}_{0,\mathrm{EXT}}$ as the subvector of $\mathbf{z}'$ that contains the contributions from $\mathbf{s}_0$: $\mathbf{z}_{0,\mathrm{EXT}} = (z'(k_0), \ldots, z'(k_0+N+\nu+L-2))^T$. Because of the already made assumption, the vector $\mathbf{z}_{0,\mathrm{EXT}}$ can be expressed as

$$\mathbf{z}_{0,\mathrm{EXT}} = \mathbf{H}_{\mathrm{EXT}}\mathbf{s}_0 + \mathbf{w}_{0,\mathrm{EXT}} \tag{7.10}$$

where $\mathbf{s}_0$ is defined in (2.8) (with $i = 0$), and the vector $\mathbf{w}_0$ collects the noise samples: $\mathbf{w}_{0,\mathrm{EXT}} = (w(k_0), \ldots, w(k_0+N+\nu+L-2))^T$. The samples $w(k)$ are i.i.d. additive white Gaussian noise samples with variance N_0 and zero mean, and $\mathbf{H}_{\mathrm{EXT}}$ is an $(N+\nu+L-1)\times(N+\nu)$ matrix with a Toeplitz structure. The entries of $\mathbf{H}_{\mathrm{EXT}}$ are given by $(\mathbf{H}_{\mathrm{EXT}})_{l:l+L-1,l} = \mathbf{h}$; $l = 0, \ldots, L-1$. The useful signal in (7.10) can be written as the sum of the contribution of the data symbols and the pilot symbols:

$$\mathbf{H}_{\mathrm{EXT}}\mathbf{s}_0 = \mathbf{B}_{\mathrm{EXT}}\mathbf{h} + \mathbf{A}_{\mathrm{EXT}}\mathbf{h} \tag{7.11}$$

where $\mathbf{B}_{\mathrm{EXT}}$ and $\mathbf{A}_{\mathrm{EXT}}$ are the $(N+\nu+L-1)\times L$ Toeplitz matrices with respective entries $(\mathbf{B}_{\mathrm{EXT}})_{l:l+N+\nu-1,l} = \mathbf{s}_{\mathrm{p}}$ and $(\mathbf{A}_{\mathrm{EXT}})_{l:l+N+\nu-1,l} = \mathbf{s}_{\mathrm{d}}^{(0)}$; $l = 0, \ldots, L-1$ (See equations (5.7) and (5.8) (with $i = 0$) for the definition of $\mathbf{s}_{\mathrm{p}}$ and $\mathbf{s}_{\mathrm{d}}^{(0)}$ respectively).

The joint probability of the samples of $\mathbf{z}'$ given k_0, the channel impulse response $\mathbf{h}$, and the data symbol vector $\mathbf{a}_{\mathrm{d}}^{(0)}$ is given by

$$p\left(\mathbf{z}' \middle| k_0,\ \mathbf{h},\ \mathbf{a}_{\mathrm{d}}^{(0)}\right) = C\exp\left\{-\frac{1}{N_0}\left(\sum_{k=0}^{k_0-1}\left|z'(k)\right|^2 + \sum_{k=k_0+N+\nu+L-1}^{2(N+\nu)+L-3}\left|z'(k)\right|^2\right)\right\}\cdot \exp\left\{-\frac{1}{N_0}\left\|\mathbf{z}_{0,\mathrm{EXT}} - (\mathbf{B}_{\mathrm{EXT}}+\mathbf{A}_{\mathrm{EXT}})\,\mathbf{h}\right\|^2\right\} \tag{7.12}$$

where C is some irrelevant constant. This expression is still dependent on the unknown data symbols $\mathbf{a}_{\mathrm{d}}^{(0)}$. To get rid of this vector, we have two options. A first option is to average (7.12) over the unknown data symbols. This averaging is rather complicated so (7.12) has to be simplified first. For small values of x, $\exp(x)$ can be approximated by the first two terms of its Taylor series, i.e.

$$\exp(x) = 1 + x \quad \text{for } |x| \ll 1. \tag{7.13}$$

So for low signal to noise ratios ($E_s/N_0 \ll 1$), expression (7.12) can be approximated by

$$p\left(\mathbf{z}' \middle| k_0,\ \mathbf{h},\ \mathbf{a}_{\mathrm{d}}^{(0)}\right) = C - \frac{C}{N_0}\left(\sum_{k=0}^{k_0-1}\left|z'(k)\right|^2 + \sum_{k=k_0+N+\nu+L-1}^{2(N+\nu)+L-3}\left|z'(k)\right|^2\right) - \frac{C}{N_0}\left\|\mathbf{z}_{0,\mathrm{EXT}} - (\mathbf{B}_{\mathrm{EXT}}+\mathbf{A}_{\mathrm{EXT}})\,\mathbf{h}\right\|^2. \tag{7.14}$$

The averaging of (7.14) over the unknown data symbols is easily performed as we only need to compute the averages of $\mathbf{A}_{\text{EXT}}$ and $\mathbf{A}_{\text{EXT}}^H\mathbf{A}_{\text{EXT}}$. Using the equalities from (2.2) and the definition of $\mathbf{A}_{\text{EXT}}$, it is not hard to show that $\mathbb{E}\left[\mathbf{A}_{\text{EXT}}\right] = 0$. The average of $\mathbf{A}_{\text{EXT}}^H\mathbf{A}_{\text{EXT}}$ is a little bit more difficult to obtain. First of all we introduce the notation $\mathbf{R}_{\mathbf{A}}$ for this average: $\mathbf{R}_{\mathbf{A}} = \mathbb{E}\left[\mathbf{A}_{\text{EXT}}^H\mathbf{A}_{\text{EXT}}\right]$.

Note that $\mathbf{A}_{\text{EXT}}^H\mathbf{A}_{\text{EXT}}$ is a Hermitian symmetric matrix, so it is sufficient to only consider the elements $(k,\ l)$ with $l \geq k$ for the computation of $\mathbf{R}_{\mathbf{A}}$. The elements of $\mathbf{A}_{\text{EXT}}^H\mathbf{A}_{\text{EXT}}$ are given by

$$\left(\mathbf{A}_{\text{EXT}}^H\mathbf{A}_{\text{EXT}}\right)_{k,l} = \sum_{m=0}^{N-1-(l-k)} s_{\text{d}}^{(0)}(m)\left(s_{\text{d}}^{(0)}(m+l-k)\right)^* \tag{7.15}$$

$$l \geq k,\ k = 0,\ \ldots,\ L-1 \tag{7.16}$$

where $s_{\text{d}}^{(0)}(m)$ are the elements of the vector $\mathbf{s}_{\text{d}}^{(0)}$, defined in (5.8). To obtain the (k,l)-th member of $\mathbf{R}_{\mathbf{A}}$, e.g. $(\mathbf{R}_{\mathbf{A}})_{k,l} = \mathbb{E}\left[\left(\mathbf{A}_{\text{EXT}}^H\mathbf{A}_{\text{EXT}}\right)_{k,l}\right]$, we need to compute the average of $\left(s_{\text{d}}^{(0)}(m_2)\right)^* s_{\text{d}}^{(0)}(m_1)$:

$$\mathbb{E}\left[s_{\text{d}}^{(0)}(m_1)\left(s_{\text{d}}^{(0)}(m_2)\right)^*\right] = \frac{1}{N+\nu}\sum_{n_1,n_2=0}^{N-M+\nu-1}\mathbb{E}\left[a_{\text{d}}(n_1)\,a_{\text{d}}^*(n_2)\right]e^{j2\pi\frac{m_1\beta_{n_1}-m_2\beta_{n_2}}{N}},$$

$$m_1,\ m_2 = 0,\ \ldots,\ N-1, \tag{7.17}$$

where β_{n_1} and β_{n_2} are members of the set of data carriers S_{d}. Exploiting (2.2), this expression can be further simplified, yielding

$$\mathbb{E}\left[s_{\text{d}}^{(0)}(m_1)\left(s_{\text{d}}^{(0)}(m_2)\right)^*\right] = \frac{E_{\text{s}}}{N+\nu}\sum_{n_1=0}^{N-M+\nu-1}e^{j2\pi\frac{(m_1-m_2)\beta_{n_1}}{N}}. \tag{7.18}$$

Replacing $\mathbb{E}\left[s_{\text{d}}^{(0)}(m_1)\left(s_{\text{d}}^{(0)}(m_2)\right)^*\right]$ by equation (7.18) in the expression for the elements $(\mathbf{R}_{\mathbf{A}})_{k,l}$ results in

$$(\mathbf{R}_{\mathbf{A}})_{k,l} = \frac{E_{\text{s}}}{N+\nu}\sum_{m=0}^{N-1+k-l}\sum_{n_1=0}^{N-M+\nu-1}e^{j2\pi\frac{(k-l)\beta_{n_1}}{N}} \tag{7.19}$$

$$= \frac{N+k-l}{N+\nu}E_{\text{s}}\sum_{n_1=0}^{N-M+\nu-1}e^{j2\pi\frac{(k-l)\beta_{n_1}}{N}} \tag{7.20}$$

$$l \geq k,\ k = 0,\ \ldots,\ L-1 \tag{7.21}$$

This yields for $p\left(\mathbf{z}' \left| k_0,\ \mathbf{h}\right.\right)$

$$p\left(\mathbf{z}' \left| k_0,\ \mathbf{h}\right.\right) = C\left\{1 - \frac{1}{N_0}\left[\mathbf{z}'^H\mathbf{z}' - \mathbf{z}_{0,\mathrm{EXT}}^H\mathbf{B}_{\mathrm{EXT}}\mathbf{h} - \mathbf{h}^H\mathbf{B}_{\mathrm{EXT}}^H\mathbf{z}_{0,\mathrm{EXT}}\right]\right.$$
$$\left. - \frac{1}{N_0}\mathbf{h}^H\left(\mathbf{B}_{\mathrm{EXT}}^H\mathbf{B}_{\mathrm{EXT}} + \mathbf{R}_{\mathbf{A}}\right)\mathbf{h}\right\}. \quad (7.22)$$

The ML estimates of k_0 and $\mathbf{h}$ are obtained by maximizing the probability function of (7.22) with respect to k_0 and $\mathbf{h}$. The estimate $\hat{\mathbf{h}}$ can be expressed as a function of k_0 and is obtained by equating to zero the derivative of (7.22) with respect to $\mathbf{h}$. This yields

$$\hat{\mathbf{h}}\left(k_0\right) = \left(\mathbf{B}_{\mathrm{EXT}}^H\mathbf{B}_{\mathrm{EXT}} + \mathbf{R}_{\mathbf{A}}\right)^{-1}\mathbf{B}_{\mathrm{EXT}}^H\mathbf{z}_{0,\mathrm{EXT}}. \quad (7.23)$$

To find the ML estimate of k_0, we substitute the estimate of $\mathbf{h}$ in (7.22), where we neglect the terms that are independent of $\mathbf{h}$ and k_0. We obtain the function $\Gamma_{TD,1}\left(k_0\right)$, which only depends on k_0:

$$\Gamma_{\mathrm{TD},1}\left(k_0\right) = \frac{1}{N_0}\mathbf{z}_{0,\mathrm{EXT}}^H\mathbf{B}_{\mathrm{EXT}}\left(\mathbf{B}_{\mathrm{EXT}}^H\mathbf{B}_{\mathrm{EXT}} + \mathbf{R}_{\mathbf{A}}\right)^{-1}\mathbf{B}_{\mathrm{EXT}}^H\mathbf{z}_{0,\mathrm{EXT}}. \quad (7.24)$$

The estimate of k_0 is then given by

$$\hat{k}_0 = \arg\max_{k_0}\left\{\Gamma_{\mathrm{TD},1}\left(k_0\right)\right\}. \quad (7.25)$$

A second approach to cope with the presence of the unknown data symbols in (7.12) is to totally neglect the contributions of the unknown data symbols. This means that we neglect $\mathbf{A}$ in (7.12) and (7.14). In that case the estimate of $\mathbf{h}$ given k_0 is given by

$$\hat{\mathbf{h}}\left(k_0\right) = \left(\mathbf{B}_{\mathrm{EXT}}^H\mathbf{B}_{\mathrm{EXT}}\right)^{-1}\mathbf{B}_{\mathrm{EXT}}^H\mathbf{z}_{0,\mathrm{EXT}} \quad (7.26)$$

and the estimate of k_0 is then given by

$$\hat{k}_0 = \arg\max_{k_0}\left\{\Gamma_{\mathrm{TD},2}\left(k_0\right)\right\} \quad (7.27)$$

with

$$\Gamma_{\mathrm{TD},2}\left(k_0\right) = \frac{1}{N_0}\mathbf{z}_{0,\mathrm{EXT}}^H\mathbf{B}_{\mathrm{EXT}}\left(\mathbf{B}_{\mathrm{EXT}}^H\mathbf{B}_{\mathrm{EXT}}\right)^{-1}\mathbf{B}_{\mathrm{EXT}}^H\mathbf{z}_{0,\mathrm{EXT}}. \quad (7.28)$$

The maximization of both $\Gamma_{\mathrm{TD},1}\left(k_0\right)$ and $\Gamma_{\mathrm{TD},2}\left(k_0\right)$ cannot be performed analytically and therefore a one-dimensional search procedure is applied.

Although (7.23) or (7.26) can be used to estimate $\mathbf{h}$ after obtaining $\hat{k}_0$ from (7.25) or (7.27), we are not planning to use these estimates, because they perform poorly at high E_s/N_0. Neglecting the contributions from the data symbols in (7.14) and (7.22) results in an error floor in the MSE of $\mathbf{h}$ and in the

resulting BER (see chapter 6 and [23, 27]). Actually, (7.23) and (7.26) are only used to obtain $\Gamma_{\mathrm{TD},1}(k_0)$ and $\Gamma_{\mathrm{TD},2}(k_0)$. We have derived channel estimators that perform better at high E_s/N_0 than the estimators (7.23) and (7.26) in chapter 6.

The functions $\Gamma_{\mathrm{TD},1}(k_0)$ and $\Gamma_{\mathrm{TD},2}(k_0)$ have a similar structure. They both compute the correlation between the received signal and the pilot vector $\mathbf{s}_p$ at L successive time instants as be can be concluded from the matrix product $\mathbf{B}_{\mathrm{EXT}}^H \mathbf{z}_{0,\mathrm{EXT}}$:

$$\left(\mathbf{B}_{\mathrm{EXT}}^H \mathbf{z}_{0,\mathrm{EXT}}\right)_l = \sum_{k=0}^{N-1} z'(k_0+l+k)\left(s_p(k)\right)^* + \sqrt{\frac{N}{N+\nu}}\left(\sum_{k=0}^{\nu-1} z'(k_0+l+N+k)\left(a_g(k)\right)^*\right) \quad (7.29)$$

where $l = 0, \ldots, L-1$. Both the estimators (7.25) and (7.27) search for the $\hat{k}_0$ that maximizes a function of the L successive correlations between the received signal and the pilot vector.

The major drawbacks of these two proposed algorithms are: *i*) the fact that eventually at high E_s/N_0 they will suffer from the interference of the unknown data symbols, and *ii*) the high computational load.

To illustrate this we are going to compute the number of complex multiplications and the number of complex additions that are necessary to obtain the functions $\Gamma_{\mathrm{TD},1}(k_0)$ and $\Gamma_{\mathrm{TD},2}(k_0)$ for one test value of k_0. The functions $\Gamma_{\mathrm{TD},1}(k_0)$ and $\Gamma_{\mathrm{TD},2}(k_0)$ have both a similar structure: they consist of the multiplication of the Hermitian conjugate of a vector of length $N+\nu+L-1$ with the product of a square matrix with dimensions equal to $N+\nu+L-1$ and the same vector. This results in a total number of $(N+\nu+L)\cdot(N+\nu+L-1)$ complex multiplications and $(N+\nu+L-1)\cdot(N+\nu+L-2)+1$ complex additions. We can conclude that it takes $\mathcal{O}\left(N^2\right)$ arithmetical operations per test value of k_0.

7.5 Estimation Based on the Guard Interval

The time delay estimator from the previous section exhibits a high computational load, so we are looking for alternatives. The first solution that comes in mind is to try to estimate the time delay offset only based on the time domain pilot symbols that are transmitted during the guard interval. We consider again the block of samples $\mathbf{z}'$, but this time we assume that $\mathbf{z}'$ only contains the contribution of the guard interval pilot symbols $\mathbf{a}_g$ and noise. The subvector of $\mathbf{z}'$ that contains the contributions from $\mathbf{a}_g$ is denoted as $\mathbf{z}_g$: $\mathbf{z}_g = \left(z'(k_0+N), \ldots, z'(k_0+N+\nu+L-2)\right)^T$. Based on the assumption

made in this section, the received signal vector $\mathbf{z}_{\mathrm{g}}$ can be written as

$$\mathbf{z}_{\mathrm{g}} = \mathbf{B}'_{\mathrm{g}}\mathbf{h} + \mathbf{w}_{\mathrm{g}}, \tag{7.30}$$

where $\mathbf{B}'_{\mathrm{g}}$ is a $(\nu + L - 1) \times L$ Toeplitz matrix which is defined as

$$\left(\mathbf{B}'_{\mathrm{g}}\right)_{l:l+\nu-1,l} = \sqrt{\frac{N}{N+\nu}}\mathbf{a}_{\mathrm{g}};\ l = 0,\ \ldots,\ L-1, \tag{7.31}$$

and the noise vector $\mathbf{w}_{\mathrm{g}}$ contains the corresponding noise samples $w\left(k_0 + N\right)$, ..., $w\left(k_0 + N + \nu + L - 2\right)$.

The joint log likelihood function $\ln p\left(\mathbf{z}'|k_0, \mathbf{h}\right)$ of the time delay offset k_0 and the channel impulse response $\mathbf{h}$ given the observation vector $\mathbf{z}'$ is expressed as

$$\ln p\left(\mathbf{z}'|k_0, \mathbf{h}\right) = -\frac{1}{N_0}\left(\sum_{k=0}^{k_0+N-1}\left|z'\left(k\right)\right|^2 + \sum_{k=k_0+N+\nu+L-1}^{2(N+\nu)+L-3}\left|z'\left(k\right)\right|^2\right) - \frac{1}{N_0}\left\|\mathbf{z}_{\mathrm{g}} - \mathbf{B}'_{g}\mathbf{h}\right\|^2. \tag{7.32}$$

Rearranging the terms in (7.32) yields for the joint log likelihood function $\ln p\left(\mathbf{z}'|k_0, \mathbf{h}\right)$

$$\ln p\left(\mathbf{z}'|k_0, \mathbf{h}\right) = -\frac{1}{N_0}\left(\mathbf{z}'^H\mathbf{z}' - \mathbf{z}_{\mathrm{g}}^H\mathbf{B}'_{\mathrm{g}}\mathbf{h} - \mathbf{h}^H\mathbf{B}'^H_{\mathrm{g}}\mathbf{z}_{\mathrm{g}}^H + \mathbf{h}^H\mathbf{B}'^H_{\mathrm{g}}\mathbf{B}'_{\mathrm{g}}\mathbf{h}\right). \tag{7.33}$$

The first term does not depend on k_0 nor on $\mathbf{h}$, so it has no further use for the estimation problem and can be neglected. The joint maximization of $\ln p\left(\mathbf{z}'|k_0, \mathbf{h}\right)$ with respect to k_0 and $\mathbf{h}$ can be reduced to a one dimensional search over the time delay offset k_0 followed by the analytical computation of the estimate of the channel impulse response $\mathbf{h}$: the estimate of $\mathbf{h}$ is given by

$$\hat{\mathbf{h}}\left(k_0\right) = \left(\mathbf{B}'^H_{\mathrm{g}}\mathbf{B}'_{\mathrm{g}}\right)^{-1}\mathbf{B}'^H_{\mathrm{g}}\mathbf{z}_{\mathrm{g}}. \tag{7.34}$$

Substituting $\mathbf{h}$ by its estimate $\hat{\mathbf{h}}\left(k_0\right)$ in the log likelihood function $\ln p\left(\mathbf{z}'|k_0, \mathbf{h}\right)$ yields a function that only depends on k_0

$$\Gamma_{\mathrm{g}}\left(k_0\right) = \frac{1}{N_0}\mathbf{z}_{\mathrm{g}}^H\mathbf{B}'_{\mathrm{g}}\left(\mathbf{B}'^H_{\mathrm{g}}\mathbf{B}'_{\mathrm{g}}\right)^{-1}\mathbf{B}'^H_{\mathrm{g}}\mathbf{z}_{\mathrm{g}}. \tag{7.35}$$

The maximization of this function with respect to k_0 yields the estimate of k_0:

$$\hat{k}_0 = \arg\max_{k_0}\Gamma_{\mathrm{g}}\left(k_0\right). \tag{7.36}$$

The function $\Gamma_{\mathrm{g}}\left(k_0\right)$ has a similar structure as the functions $\Gamma_{\mathrm{TD},1}\left(k_0\right)$ (7.24) and $\Gamma_{\mathrm{TD},2}\left(k_0\right)$ (7.28) from the previous section: $\Gamma_{\mathrm{g}}\left(k_0\right)$ consists of the multiplication of the Hermitian conjugate of a vector with the product of a square

matrix and the same vector. The only difference are the dimensions: the vector has a length equal to $\nu + L - 1$ and the square matrix is a $(\nu + L - 1) \times (\nu + L - 1)$ matrix. This results in a total number of $(\nu + L) \cdot (\nu + L - 1)$ complex multiplications and $(\nu + L - 1) \cdot (\nu + L - 2) + 1$ complex additions. Usually the length of the guard interval ν and the length of the channel impulse response L are of the same order, so we can conclude that it takes $\mathcal{O}\left(\nu^2\right)$ arithmetical operations.

Note that we have neglected the contributions from the actual OFDM blocks in $\mathbf{z}_g$ (7.30), so depending on how the guard interval length ν compares to length L of the channel impulse response, the performance might be different. When ν is somewhat larger than L, this algorithm will work better than when ν has the minimal length of $\nu = L - 1$.

7.6 Pilot Aided FD Estimation

We have seen that the estimators which operate in the time domain (see section 7.4 and section 7.5) suffer from the interference from the unknown data symbols, because it is impossible to separate the data symbol contributions from the pilot symbol contributions.

However, in the frequency domain we have a whole different story. We know already from the previous chapter (see section 6.6) that when the block of $N + \nu$ samples $\mathbf{z}_0 = (z'(k_0),\ \ldots,\ z'(k_0 + N + \nu - 1)\ )^T$ are transformed to the FD, by adding the last ν samples of $\mathbf{z}_0$ to the first ν samples of $\mathbf{z}_0$ and applying the resulting N samples to an FFT, the orthogonality of the carriers is restored. This means that the FFT outputs at the pilot carrier positions only depend on both the TD pilot symbols $\mathbf{a}_g$ and the pilot symbols transmitted on the pilot carriers $\mathbf{a}_c$. We can exploit this fact to derive an estimator for the time delay k_0.

We consider the $N + \nu$ samples $\mathbf{z}_0$, which are needed for the detection of the OFDM block $i = 0$. To restore the orthogonality of the carriers, the last ν samples are added to the first ν samples, resulting in N samples that are fed to the FFT:

$$\tilde{\mathbf{z}}\left(k_0\right) = \mathbf{F}\mathbf{\Omega}\mathbf{z}_0. \tag{7.37}$$

The $M - \nu$ FFT outputs at the pilot carrier positions are given by

$$\tilde{\mathbf{z}}_2\left(k_0\right) = \tilde{\mathbf{B}}_2\mathbf{h} + \tilde{\mathbf{w}}_2, \tag{7.38}$$

where we have used the same notations as in section 6.6, see (6.61) for the definition of the pilot matrix $\tilde{\mathbf{B}}_2$. The noise samples $\tilde{\mathbf{w}}_2$ are zero mean Gaussian distributed with an autocorrelation matrix $\tilde{\mathbf{R}}_2$ given by (6.60). Note that the data symbols are not present in $\tilde{\mathbf{z}}_2\left(k_0\right)$. The log likelihood function of k_0 and $\mathbf{h}$ given the observation $\tilde{\mathbf{z}}_2\left(k_0\right)$ is defined as

$$\ln p\left(\tilde{\mathbf{z}}_2\left(k_0\right)|k_0,\ \mathbf{h}\right) = -\left(\tilde{\mathbf{z}}_2\left(k_0\right) - \tilde{\mathbf{B}}_2\mathbf{h}\right)^H \tilde{\mathbf{R}}_2^{-1}\left(\tilde{\mathbf{z}}_2\left(k_0\right) - \tilde{\mathbf{B}}_2\mathbf{h}\right). \tag{7.39}$$

Now the ML estimates of k_0 and $\mathbf{h}$ based on the log likelihood function $\ln p\left(\tilde{\mathbf{z}}_2\left(k_0\right)|k_0,\ \mathbf{h}\right)$ can be obtained. The ML estimate of $\mathbf{h}$ can be written as a closed form expression and is a function of k_0 and is given by

$$\hat{\mathbf{h}}\left(k_0\right) = \left(\tilde{\mathbf{B}}_2^H \tilde{\mathbf{R}}_2^{-1} \tilde{\mathbf{B}}_2\right)^{-1} \tilde{\mathbf{B}}_2^H \tilde{\mathbf{R}}_2^{-1} \tilde{\mathbf{z}}_2\left(k_0\right). \tag{7.40}$$

This estimate $\tilde{\mathbf{h}}$ corresponds to the estimate obtained with the FD DA estimator from section 6.6. When we use (7.40) to replace $\mathbf{h}$ in the log likelihood function (7.39), a function is obtained that only depends on the time delay k_0:

$$\Gamma_{\mathrm{FD}}\left(k_0\right) = -\tilde{\mathbf{z}}_2^H\left(k_0\right)\left(\tilde{\mathbf{R}}_2^{-1} - \tilde{\mathbf{R}}_2^{-1}\tilde{\mathbf{B}}_2\left(\tilde{\mathbf{B}}_2^H\tilde{\mathbf{R}}_2^{-1}\tilde{\mathbf{B}}_2\right)^{-1}\tilde{\mathbf{B}}_2^H\tilde{\mathbf{R}}_2^{-1}\right)\tilde{\mathbf{z}}_2\left(k_0\right). \tag{7.41}$$

The ML estimate of k_0 is given by

$$\hat{k}_0 = \arg\max_{k_0} \Gamma_{\mathrm{FD}}\left(k_0\right). \tag{7.42}$$

The maximization of $\Gamma_{\mathrm{FD}}\left(k_0\right)$ can not be carried out analytically, so we have to resort to a linear search.

This estimation algorithm is very similar to the TD algorithms we have proposed in section 7.4. But in contrast with the TD algorithms, there is no interference from the unknown data symbols and also the computational load has been reduced: The function $\Gamma_{\mathrm{FD}}\left(k_0\right)$ has the same structure as the TD functions $\Gamma_{\mathrm{TD},1}\left(k_0\right)$ and $\Gamma_{\mathrm{TD},2}\left(k_0\right)$, it consists also of the multiplication of the Hermitian conjugate of a vector of observations $\tilde{\mathbf{z}}_2^H\left(k_0\right)$ with the product of a square matrix and the same vector of observations $\tilde{\mathbf{z}}_2\left(k_0\right)$, but the dimensions are different. The observation vector $\tilde{\mathbf{z}}_2\left(k_0\right)$ has a length of $M-\nu$ samples. So the computation of $\Gamma_{\mathrm{FD}}\left(k_0\right)$ needs $(M-\nu+1)\cdot(M-\nu)$ complex multiplications and $(M-\nu)\cdot(M-\nu-1)+1$ complex additions for every test value of k_0. Additionally, the algorithm operates in the FD, so to obtain the observation vector $\tilde{\mathbf{z}}_2\left(k_0\right)$, we need an extra $(N/2)\cdot\log_2 N$ complex multiplications and $\nu + N\cdot\log_2 N$ complex additions. Usually the number of pilot carriers $M-\nu$ is much smaller than the total number of carriers N, so the computational complexity is dominated by the transformation to the FD. As a result, the complexity is of the order $\mathcal{O}\left(N\cdot\log_2 N\right)$ per test value of k_0.

When we compare this with the computational complexity of the TD algorithms, we can conclude that despite the extra FFT, which is needed for every test value of k_0, the FD estimation algorithm relaxes the number of necessary arithmetical operations.

7.7 Numerical Results

This section shows some numerical results for the proposed algorithms and compares them with the existing algorithms from the literature.

We consider $N = 1024$ carriers and a guard interval of length $\nu = 100$. The total number of pilot symbols M is equal to 200, which means that there are 100 pilot symbols who are transmitted during the guard interval and 100 carriers are selected as pilot carrier. For the pilot symbol placement in the frequency domain, we have considered the comb-type pilot scheme (see section 5.2 for more details). The transmitted symbols consist of randomly generated QPSK symbols. The channel impulse response is given by (4.29). We consider Rayleigh fading [3], so the samples of $\mathbf{h}$ are Gaussian distributed with zero mean and variance equal to $1/L$, with $L = 50$. Note that this channel model corresponds to a worst case scenario: the power delay profile is uniform which means that there is a lot of multipath interference, even from paths with high delays. This interference makes the estimation of the time delay offset based on the TD received signal more difficult. For an exponentially decaying power delay profile the interference from the paths with the highest delays will be less severe.

7.7.1 Histogram of the estimation error $\hat{k}_0 - k_0$

The performance of the time delay offset estimation algorithms is first shown by means of a histogram of the probability of the estimation error $P(\hat{k}_0 - k_0)$ as a function of the estimation error $\hat{k}_0 - k_0$. To obtain the histograms of the different considered estimation algorithms, the SNR E_s/N_0 is set to 20 dB. An estimate $\hat{k}_0$ of the time delay offset k_0 is considered sufficiently accurate if it is in the desirable range $[k_0 - \nu + L - 1,\ k_0]$. Furthermore a small variance on the estimate $\hat{k}_0$ is also considered an advantageous property because in that case we can reduce the length of the guard interval ν.

To obtain the numerical results for the TDS-OFDM algorithms from section 7.4, we have generated a guard interval with the structure given / shown in subsection 1.4.1. The pseudo noise sequence is given by a maximum length sequence of $N_{\text{PN}} = 2^6 - 1$ symbols. The prefix has a length of $N_{\text{pre}} = 19$ samples and consists of the last 19 symbols of the maximum length sequence, while the postfix has a length of $N_{\text{post}} = 18$ samples and consists of the first 18 symbols of the maximum length sequence. The guard interval has a length of $\nu = N_{\text{pre}} + N_{\text{PN}} + N_{\text{post}} = 100$ samples.

First of all we review the performance of the existing methods for TDS-OFDM that are described in section 7.3. Figure 7.3 and 7.4 show the results for the slide auto-correlation based algorithm from subsection 7.3.1 and the composite PN-correlation based algorithm from subsection 7.3.2 respectively. The performances of the TDS-OFDM algorithms are poor in the considered KSP-OFDM system: less than 3% of the estimates of the time delay offset are in the desirable range $[k_0 - \nu + L - 1,\ k_0]$ for the composite PN-correlation based algorithm from subsection 7.3.2, while for the slide auto-correlation based algorithm from subsection 7.3.1 it is even worse: less than 1% of the

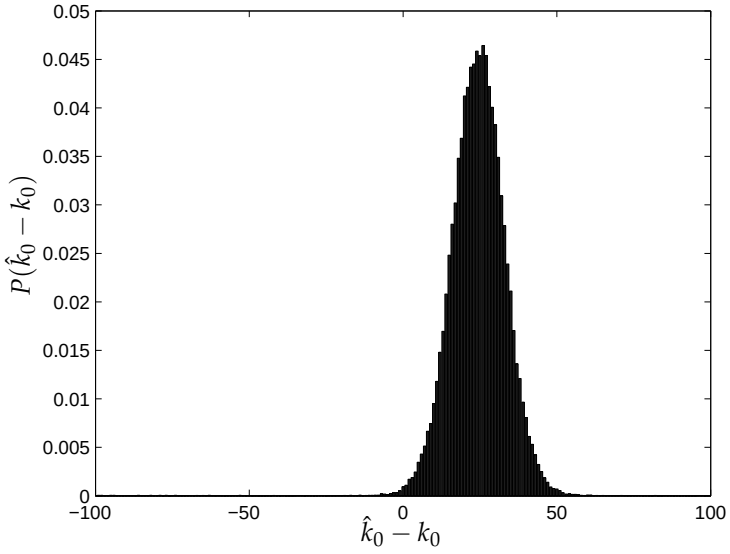

Figure 7.3: Histogram of the time delay estimation error for the TDS-OFDM estimator from subsection 7.3.1, $N = 1024$, $\nu = 100$, $L = 50$, $M = 200$, $E_s/N_0 = 20$ dB.

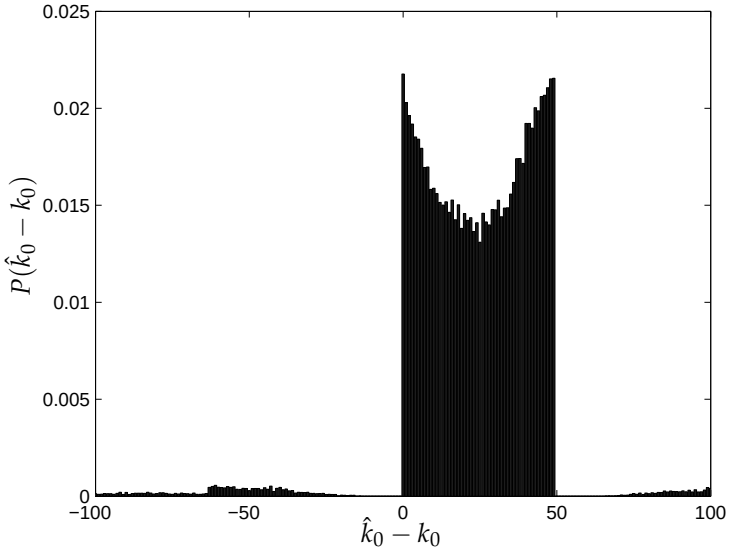

Figure 7.4: Histogram of the time delay estimation error for the TDS-OFDM estimator from subsection 7.3.2, $N = 1024$, $\nu = 100$, $L = 50$, $M = 200$, $E_s/N_0 = 20$ dB.

estimates are in the desirable range. Applying these algorithms will result in interference from other OFDM blocks and additionally when the estimation error is to large, the orthogonality of the carriers can not be restored, which gives rise to ICI (see also section 7.2). Both algorithms are derived under the assumption of a flat fading channel, which is clearly not the case in the simulated system. This explains their bad performance. However it can be noticed that the main part of the estimates is in the interval $[k_0,\ k_0 + L]$: for the algorithm from subsection 7.3.2 this is true in approximately 83% of the cases (see figure 7.4), while for the algorithm from subsection 7.3.1 it is true in 96% of the considered cases (see also figure 7.3). The performances of the two considered algorithms can be improved by exploiting the fact that most of their estimates are in the range $[k_0,\ k_0 + L]$: the new estimate is then given by $\hat{k}_0' = \hat{k}_0 - L$, where $\hat{k}_0$ is given by (7.5) or (7.9) . Unfortunately, this solution is only an ad hoc method to improve the estimate. For another channel model, this solution might not work that well.

Next, we consider the pilot aided time domain algorithms from section 7.4. The results are shown in figure 7.5. The left side of the figure corresponds to the estimator (7.25) which averages the likelihood function first over the unknown data symbols. The results for the algorithm (7.27) which ignores the unknown data symbols, are shown in the right part. Both algorithms have a similar performance: their histogram shows a pronounced peak for $\hat{k}_0 = k_0$. The TD estimator which takes the data symbols into account (7.25), finds the correct value of k_0 in more than 83% of the simulated cases, while the TD estimator which ignores the data symbols (7.27) yields an estimation error equal to zero in more than 92% of the cases. Furthermore, the estimates are all in the neighborhood of the correct value k_0: the estimation error $\left|\hat{k}_0 - k_0\right|$ is smaller than two samples in more than 97% of all simulated cases for the TD estimator which takes the data symbols into consideration (7.25), while for the TD estimator which ignores the data symbols, the probability that the estimation error $\left|\hat{k}_0 - k_0\right|$ is smaller than two samples, is more than 99%.

Compared to the existing TDS-OFDM estimators from section 7.3, these two KSP-OFDM TD estimation algorithms yield a much better performance. The variance of the estimates obtained with the KSP TD estimators (7.25) and (7.27), is much smaller than the variance of the estimates obtained with the two TDS-OFDM estimators from section 7.3. However, the KSP-OFDM TD algorithms have a much higher computational complexity, i.e. $\mathcal{O}\left(N^2\right)$, than the TDS-OFDM algorithms.

To lower the computational complexity, we now consider time delay offset estimation based on the guard interval with the estimator (7.36) from section 7.5. The histogram of the estimation error $\hat{k}_0 - k_0$ is shown in figure 7.6. The correct value of k_0 is found by the estimator in more than 60% of all the cases. Compared to the TDS-OFDM estimators from section 7.3, the performance of the guard interval based estimator (7.36) is significantly better, as most of the

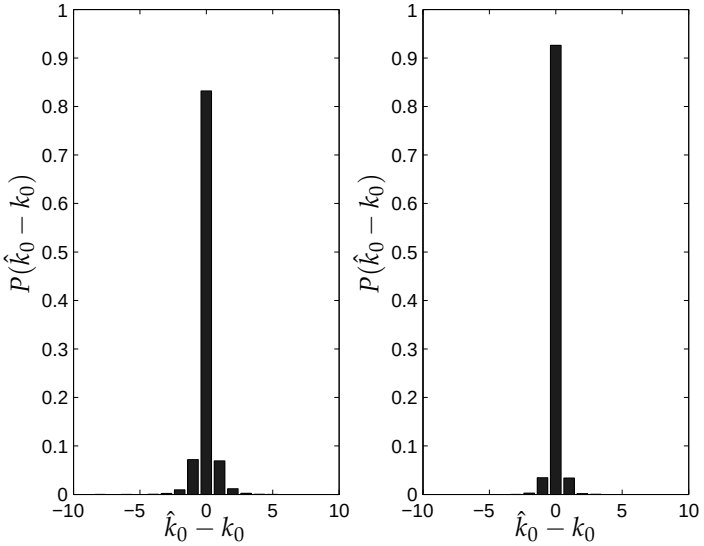

Figure 7.5: Histogram of the time delay estimation error for the KSP-OFDM TD estimators, $N = 1024$, $\nu = 100$, $L = 50$, $M = 200$, $E_s/N_0 = 20$ dB, (a): (7.25) (b): (7.27).

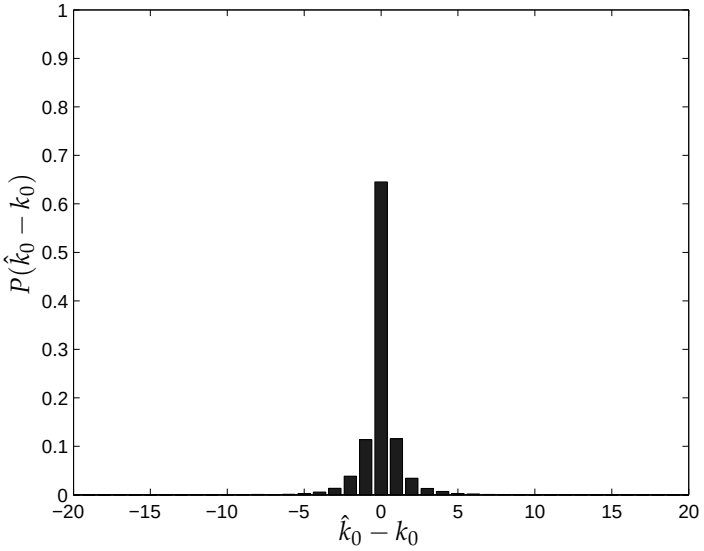

Figure 7.6: Histogram of the time delay estimation error for the KSP-OFDM guard interval based estimator, $N = 1024$, $\nu = 100$, $L = 50$, $M = 200$, $E_s/N_0 = 20$ dB.

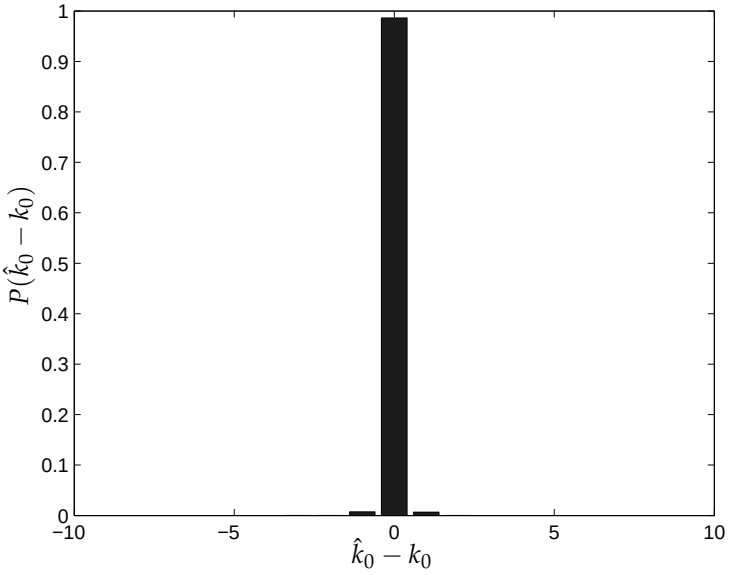

Figure 7.7: Histogram of the time delay estimation error for the KSP-OFDM FD estimator, $N = 1024$, $\nu = 100$, $L = 50$, $M = 200$, $E_s/N_0 = 20$ dB.

estimates are close to the correct value of k_0: the estimation error $\left|\hat{k}_0 - k_0\right|$ is 5 samples or less in 99% of all the simulated cases. If we compare this result with the histograms for TDS-OFDM (figures 7.3 and 7.4), we see that this is clearly not the case for the TDS-OFDM estimators. However, the performance of the KSP-OFDM TD estimators (7.25) and (7.27) from section 7.4 is not reached. The lower computational complexity has its price: estimating the time delay offset, based on the presence of the pilot symbols in the time domain only results in a degradation, or in other words exploiting the presence of the pilot symbols on the pilot carriers gives rise to a performance improvement (besides a higher computational complexity).

Finally we consider the frequency domain based time delay offset estimation (7.42) from section 7.6. Figure 7.7 shows the results for the histogram of the estimation error $\hat{k}_0 - k_0$. The FD estimator yields by far the best result. The correct value of the time delay offset k_0 is found in 98% of all the simulated cases. The variance of the estimate $\hat{k}_0$ is very small: almost every time, the estimation error $\left|\hat{k}_0 - k_0\right|$, if it is not equal to zero, is no more than 1 sample, i.e. $P\left(\left|\hat{k}_0 - k_0\right| \leq 1\right) > 99\%$. Compared to the TDS-OFDM estimators from section 7.3, this estimator gives rise to a way better performance. This KSP-OFDM FD estimator (7.42) even outperforms the KSP-OFDM TD estima-

tors (7.25) and (7.27) (not to mention the TD guard interval based estimator (7.36) from section 7.5). However, in terms of the computational complexity, the KSP-OFDM FD estimator (7.42) needs a lot more arithmetical operations than the TDS-OFDM estimators or the KSP-OFDM guard interval based estimator. Compared to the KSP-OFDM TD estimators, it still yields a significant improvement.

Now, if we review the obtained results for the KSP-OFDM TD estimators (see figure 7.5) and the results for the KSP-OFDM FD estimator (see figure 7.7), we conclude that the variance on the estimate of the time delay offset is rather small. We have selected the length of the guard interval ν equal to 100 samples resulting in a 'safety' margin of $\nu - L = 50$ samples for the estimate of the time delay offset, or in other words, the estimate of the time delay offset must be in the range $[k_0 - \nu + L - 1,\ k_0]$. For the KSP-OFDM TD estimators (7.25) and (7.27), and the FD estimator (7.42), this margin might be a bit to large. On the other hand we want to keep the duration of the guard interval as short as possible, because a longer guard interval reduces the throughput of the system. Therefore, it is interesting to investigate the performance of the different proposed algorithms in an OFDM system with a much shorter guard interval. Now we set ν equal to 50 samples, the other parameters remain the same except the total number of pilot symbols M which reduces to 100 pilot carriers plus 50 time domain pilot symbols, e.g. $M = 150$.

First of all we start with the existing estimators for TDS-OFDM from section 7.3. To obtain the results for the TDS-OFDM algorithms, we have kept the structure of the guard interval of a TDS-OFDM system (see subsection 1.4.1 for more details), but we have scaled the different lengths N_{PN}, N_{pre} and N_{post}: the pseudo noise sequence is now given by a maximum length sequence of $N_{\text{PN}} = 2^5 - 1 = 31$ symbols. The prefix has a length of $N_{\text{pre}} = 10$ samples and consists of the last 10 symbols of the maximum length sequence, while the postfix has a length of $N_{\text{post}} = 9$ samples and consists of the first 9 symbols of the maximum length sequence. This results in a guard interval length equal to $\nu = N_{\text{pre}} + N_{\text{PN}} + N_{\text{post}} = 50$ samples.

The histogram of the estimation error $\hat{k}_0 - k_0$ for the slide auto-correlation based estimator (7.5) and the one for the composite PN-correlation based estimator (7.9) are shown in figures 7.8 and 7.9 respectively. The results are pretty bad, the estimation variance is very large and as a result the performance of the receiver is deteriorated by IBI and ICI. Due to the short guard interval there is almost no tolerance: to guarantee optimal performance, the time delay estimate $\hat{k}_0$ must be in the range $[k_0 - \nu + L - 1,\ k_0]$ (see section 7.2). For the simulated system, this means that the estimate $\hat{k}_0$ must be equal to k_0 or $k_0 - 1$. So the TDS-algorithms are not reliable for the time delay offset estimation in this case. For the previous system with the larger guard interval, there was an ad hoc solution to improve the estimate, but here in this case, applying the same solution does not lead to a significant improvement.

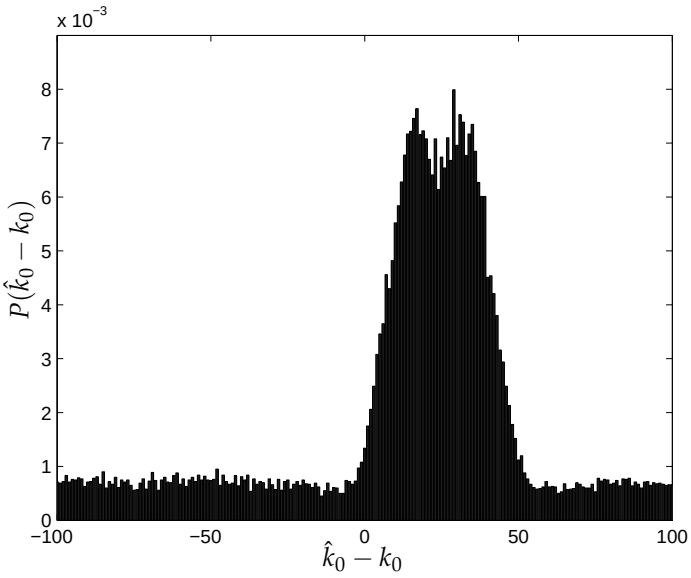

Figure 7.8: Histogram of the time delay estimation error for the TDS-OFDM estimator from subsection 7.3.1, $N = 1024$, $\nu = 50$, $L = 50$, $M = 150$, $E_s/N_0 =$ 20 dB.

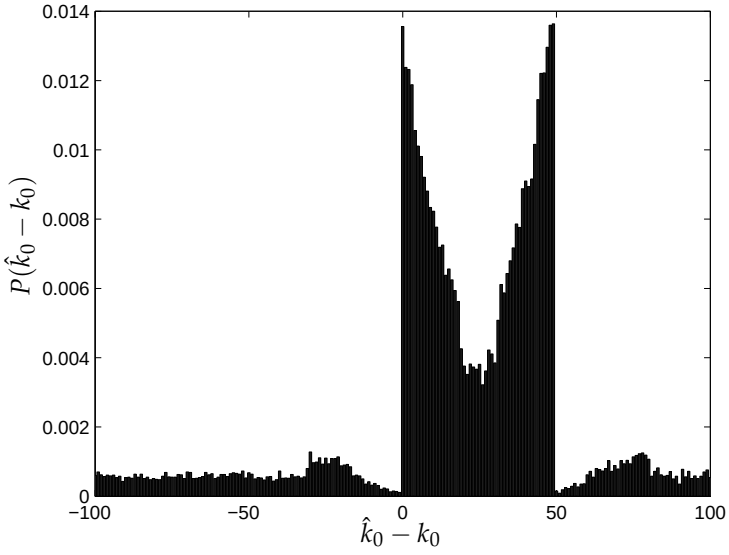

Figure 7.9: Histogram of the time delay estimation error for the TDS-OFDM estimator from subsection 7.3.2, $N = 1024$, $\nu = 50$, $L = 50$, $M = 150$, $E_s/N_0 =$ 20 dB.

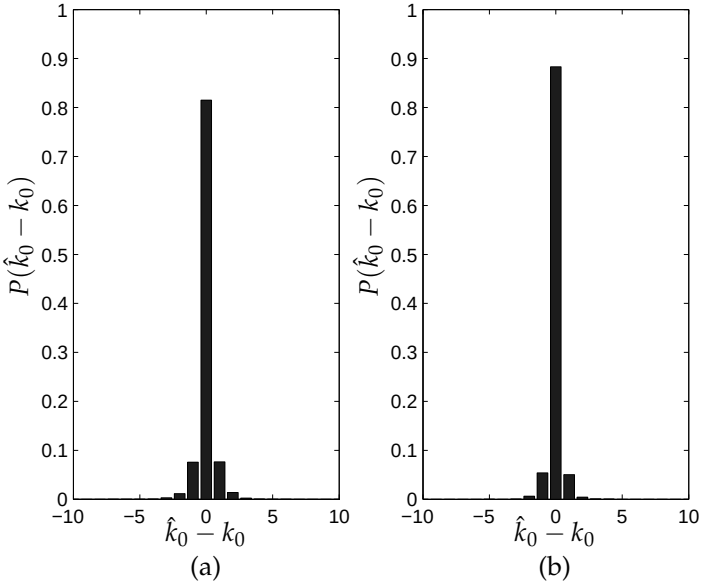

Figure 7.10: Histogram of the time delay estimation error for the KSP-OFDM guard interval based estimator, $N = 1024$, $\nu = 50$, $L = 50$, $M = 150$, $E_s/N_0 =$ 20 dB, (a): (7.25) (b): (7.27).

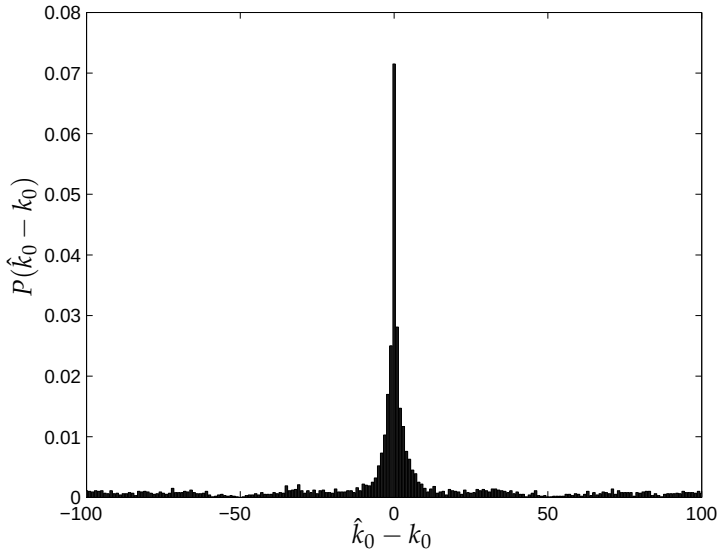

Figure 7.11: Histogram of the time delay estimation error for the KSP-OFDM guard interval based estimator, $N = 1024$, $\nu = 50$, $L = 50$, $M = 150$, $E_s/N_0 =$ 20 dB.

Next we look at the performance of the pilot aided TD estimation algorithms (7.25) and (7.27) for KSP-OFDM from section 7.4. Figure 7.10 shows the histogram of the estimation error $\hat{k}_0 - k_0$ for the KSP-OFDM with a shorter guard interval ($\nu = 50$). The results are very similar to the ones displayed in figure 7.5 ($\nu = 100$). The performance is a little bit worse: the KSP-OFDM TD estimator which averages first over the data symbols (7.25) finds the correct value of the time delay offset in more than 81% of all simulated cases and for the KSP-OFDM TD estimator which ignores the data symbols (7.27), this is true for more than 88% of the estimates. The variance of the estimation error remains small for both estimators: the estimation error $\left|\hat{k}_0 - k_0\right|$ is smaller than two samples in more than 96 % of the simulated cases for the KSP-OFDM TD estimator which takes the presence of the data symbols into account (7.25), while for the KSP-OFDM TD estimator which ignores the data symbols, the estimation error $\left|\hat{k}_0 - k_0\right|$ remains smaller than two samples in more than 98 % of all simulated cases. So we can conclude that both KSP-OFDM TD estimators do not suffer from the shortening of the guard interval. This means that it is not necessary to make the guard interval much longer than the duration of the channel impulse response in order to guarantee interference free data detection.

Next we investigate the influence of the length of the guard interval on the performance of the low complexity alternative for the KSP-OFDM TD estimators: the guard interval based estimator (7.36) from section 7.5. Figure 7.11 shows the histogram of the estimation error $\hat{k}_0 - k_0$. Similarly to the existing TDS-OFDM algorithms from section 7.3, the guard interval based KSP-OFDM estimator does not function well when the guard interval becomes very short. As a consequence the correct value of the time delay offset k_0 is only found in approximately 7% of all the simulated cases and the estimation error exhibits a large variance. Only considering the guard interval pilot symbols for the estimation of the time delay offset is only applicable when the guard interval is sufficiently longer than the duration of the channel impulse response.

Finally, the KSP-OFDM FD estimator (7.42) from section 7.6 is used for the time delay offset estimation in the KSP-OFDM with the short guard interval ($\nu = 50$). The results for the histogram of the estimation error $\hat{k}_0 - k_0$ are shown in figure 7.12. The results are almost identical as the results for $\nu = 100$. The correct value of k_0 is found in 98% of all the simulated cases. The estimation error $\left|\hat{k}_0 - k_0\right|$ is almost always smaller or equal than 1 sample: $P\left(\left|\hat{k}_0 - k_0\right| \leq 1\right) > 99\%$.

The KSP-OFDM FD estimator yields again the best performance of all the considered time delay offset estimation algorithms. Furthermore, it does not show any loss in performance at all when the guard interval is shortened. The KSP-OFDM TD estimators also exhibit a good performance but there is small degradation compared to the case where $\nu = 100$.

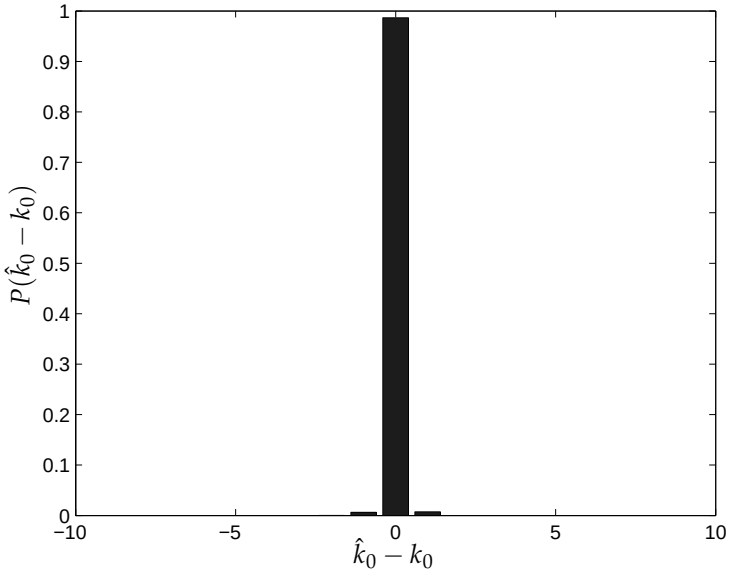

Figure 7.12: Histogram of the time delay estimation error for the KSP-OFDM FD estimator, $N = 1024$, $\nu = 50$, $L = 50$, $M = 150$, $E_s/N_0 = 20$ dB.

7.7.2 BER

In this subsection we investigate the BER performance of a receiver which applies one of the proposed time delay estimation algorithms. We have used the frequency domain symbol by symbol detector (4.28) from section 4.2 to determine the transmitted data symbols.

In figure 7.13, the BER results are shown as a function of E_b/N_0 for a KSP-OFDM with a guard interval length ν of 100 symbols. Two benchmark curves are also added to the figure for comparison. One curve corresponds to the BER of a receiver which is perfectly synchronized and has perfect knowledge about the channel state information (CSI) and is the absolute lower bound for receivers which have to estimate both the channel impulse response and the time delay offset. The other reference curve shows the BER for a receiver which is perfectly synchronized but estimates the channel impulse response. There is already a degradation caused by the imperfect channel estimation.

First of all, we consider the TD time delay offset estimation algorithms (7.25) and (7.27) from section 7.4. For low to moderate SNR ($E_b/N_0 \leq 20$ dB), the performance of a receiver which uses one of the TD estimators is equal to the performance of a receiver with perfect knowledge of the time delay offset. Only for high SNR, there is an error floor visible. The error floor is caused by the assumptions made for the derivation of the estimators: we have

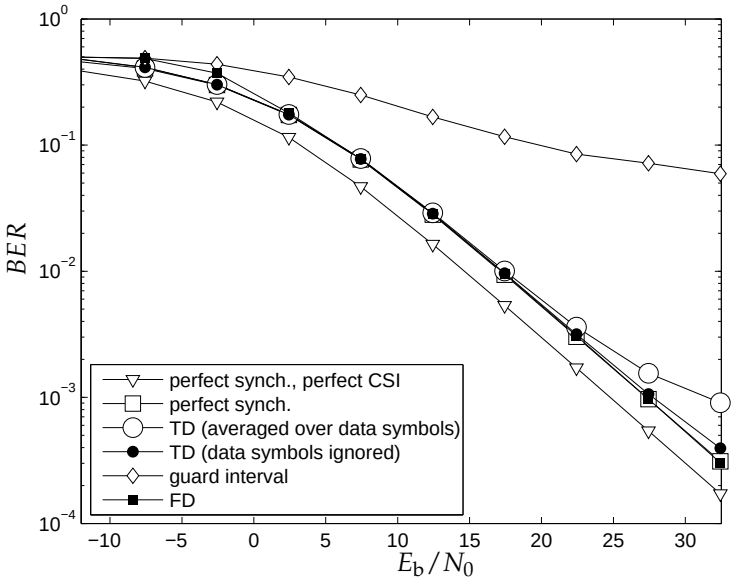

Figure 7.13: BER as a function of E_b/N_0, $N = 1024$, $\nu = 100$, $L = 50$, $M = 200$.

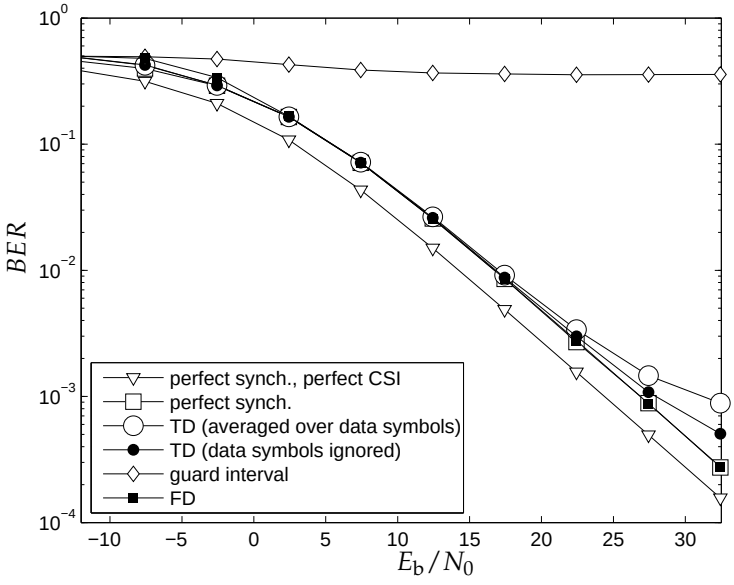

Figure 7.14: BER as a function of E_b/N_0, $N = 1024$, $\nu = 50$, $L = 50$, $M = 150$.

assumed a low SNR to obtain (7.14). The TD time delay offset estimator which averages first over the data symbols (7.25), results in the highest error floor. For the TD time delay estimation algorithm which ignores the presence of the data symbols (7.27) the error floor is not yet fully visible and the degradation remains small for the considered range of E_b/N_0 values. Next we look at the BER of a receiver which applies the guard interval based algorithm (7.36) from section 7.5 to estimate the time delay offset. The BER results clearly indicate that the time delay offsets obtained with the guard interval based algorithm are not accurate enough, and as a result there is a large degradation for the BER. Finally, the BER results of a receiver which employs the FD estimator (7.42) from section 7.6 to obtain the time delay offset estimate are shown. For low E_b/N_0, there is a small degradation compared to the performance of the receiver with perfect knowledge about the time delay offset and the receiver with one of the TD estimation algorithms from section 7.4. Note that we have shown in section 7.2 that, to make the system robust against certain timing errors, the channel impulse response needs to be extended to the whole duration of the guard interval. In that case, the timing error is taken into account in the estimate of the channel impulse response $\mathbf{h}'$ (7.1). A drawback of this extension is the increase of the number of channel impulse response samples that need to be estimated: the actual channel impulse response $\mathbf{h}$ consists of L samples, while $\mathbf{h}'$ has a length of ν samples (of which only L of them are not equal to zero). From (6.67), it can be seen that the MSE of the channel estimate depends on the number of estimated samples. So estimating $\mathbf{h}'$ results in a higher estimation variance and a loss in performance. For the guard interval based time delay offset estimator (7.36), the variance of the time delay offset estimate is very large so for this estimator we have to estimate the samples of $\mathbf{h}'$ to be able to make the system robust against some timing errors. The TD time delay offset estimators (7.25) and (7.27), and the FD time delay offset estimation algorithm (7.42), exhibit a very small variance of the estimate. So in most of the cases, they provide the correct value of the time delay offset. If the time delay offset estimate $\hat{k}_0$ is equal to the correct k_0, then the extended channel impulse response $\mathbf{h}'$ is given by

$$\mathbf{h}' = \begin{pmatrix} \mathbf{h} \\ \mathbf{0}_{(\nu-L)\times 1} \end{pmatrix}. \tag{7.43}$$

The last $\nu - L$ samples of $\mathbf{h}'$ are always equal to zero if $\hat{k}_0 = k_0$. So for time delay offset estimators which almost always deliver the correct k_0, it is not necessary to estimate ν channel impulse response samples, we only have to estimate the L samples of $\mathbf{h}$. We see from figure 7.13 that this approach works perfect for the FD time delay offset estimation algorithm: the BER of a receiver with perfect knowledge about the time delay offset is reached (except for very low E_b/N_0). For the TD time delay offset estimators, there is a degradation for (very) high values of E_b/N_0. As we can see from figure 7.5, there are some

timing errors to the left ($\hat{k}_0 < k_0$), which could be incorporated in the channel estimation by estimating the ν samples of $\mathbf{h}'$. As we did not estimate $\mathbf{h}'$ but $\mathbf{h}$, these timing errors are partially responsible for the error floor in the BER (timing errors to the right always result in a degradation).

To illustrate the influence of a short guard interval on the performance of the proposed estimation algorithms, figure 7.14 shows the BER results for a KSP-OFDM system with a guard interval consisting of $\nu = 50$ pilot symbols. The BER of a receiver which estimates the time delay offset with one of the TD estimators from 7.4 is equal to the BER of a receiver which is perfectly synchronized for low to moderate SNR ($E_b/N_0 \leq 15$ dB), although the error floor appears already at lower E_b/N_0 than in figure 7.13. Again the TD estimator (7.27) which ignores the presence of the data symbols results in the lowest error floor, but the degradation is also in this case a bit larger than for the KSP-OFDM system with $\nu = 100$. The BER results for a receiver which is equipped with the guard interval based time delay estimation algorithm, are catastrophic. These results are of course as expected if we look at the histogram of the estimation error $\hat{k}_0 - k_0$ (see figure 7.11). For the receiver with the FD estimator, there is no difference with the situation $\nu = 100$. There is a small gap with the BER performance of a receiver with perfect knowledge about the time delay offset for low E_b/N_0. For moderate to high SNR ($E_b/N_0 \geq 0$ dB) the receiver with the FD estimator reaches the performance of the perfectly synchronized receiver.

We can conclude that of all proposed time delay offset estimators, the FD estimator yields the best performance in terms of the BER for moderate to high SNR. Only for low SNR, the TD estimators result in a lower BER. For the time delay offset estimation we can not really rely on the guard interval based estimator as its BER performance is really poor.

7.8 Conclusion

In this chapter the problem of time delay offset estimation has been tackled. First we have shown that there is a range of tolerable time delay offsets which do not cause any degradation. Outside this range, a time delay offset causes IBI and ICI. Secondly two existing time delay offset estimation algorithms for TDS-OFDM have been discussed: the slide auto-correlation based estimator and the composite PN-correlation based estimator. They both try to estimate the time delay offset by exploiting the presence of the pseudo noise sequence in the guard interval. With some small modifications they can easily be applied in a KSP-OFDM context. Their major drawback is their assumption of a flat fading channel. The numerical results show that especially for short guard intervals, their performance is rather poor.

Next we have proposed some TD pilot aided estimation algorithms. They both estimate the time delay offset using the pilot symbols transmitted during

the guard interval and the pilot symbols transmitted on the pilot carriers. For their derivation we have assumed the low SNR regime. In a first approach we have replaced the likelihood function by the first two terms of its Taylor series expansion and have averaged them over the unknown data symbols. The joint estimation problem of the channel impulse response and the time delay offset reduces to a linear search for the time delay offset (and an analytical computation of the channel impulse response estimate). In a second approach, the contribution of the data symbols is ignored. Following a similar reasoning as for the first approach, the estimate of the time delay offset is found by performing a one-dimensional search. The numerical results show that they both exhibit good performance even for a very short guard interval. The major drawback of both algorithms is their high computational load. Therefore we have searched for a low complexity alternative. A major reduction in computational complexity is achieved by only considering the pilot symbols in the guard interval (but still assuming a frequency selective channel). However, for very short guard intervals, the performance of the guard interval based time delay offset estimator is not that great.

The last proposed time delay offset estimator operates in the frequency domain. It estimates the time delay offset based on the FFT outputs of the received signal at the pilot carrier positions. The algorithm exploits both the time domain pilot symbols from the guard interval and the pilot symbols on the pilot carriers. The log likelihood function shows a sharp peak for the correct value of the time delay offset; there is no interference from the data symbols. This algorithm exhibits an excellent performance independent of the length of the guard interval. The computational complexity is somewhere between the one of the KSP-OFDM TD estimators and the one from the guard interval based estimator.

The derivation of the TD pilot aided estimation algorithms and the study of their performance has been carried out in [35,71,72].

8

Carrier Frequency Offset Estimation

In this chapter we look into the estimation of an unknown carrier frequency offset. After the problem formulation, we first review existing algorithms for TDS-OFDM. These algorithms are based on the pilot symbols in the guard interval and thus they can be applied (possibly after some modifications) in a KSP-OFDM system. In a second part of this chapter, we propose some new carrier frequency offset estimation algorithms that make use of the presence of the pilot symbols in the guard interval and the pilot symbols on the pilot carriers, and that are robust against an unknown channel impulse response. A first algorithm operates in the time domain and is robust against possible time delay offsets. The algorithm is based on the periodicity of the transmitted pilot signal. A second class of algorithms that we introduce in this work, operates in the frequency domain. One algorithm estimates the carrier frequency offset based on the FFT outputs at pilot carrier positions. A second algorithm is a decision aided algorithm and makes use of all of the FFT outputs. The last part of the chapter consists of numerical results to illustrate the performance

of the proposed estimation algorithms and to compare them with the existing algorithms for TDS-OFDM.

8.1 Problem Statement

One of the problems in an OFDM system is its sensitivity to a frequency mismatch between the oscillators in the transmitter and the receiver. The BER performance can be severely degraded when the carrier frequency offset (CFO) (even of the order of a fraction of the carrier spacing) is not compensated [73], because the frequency offset causes inter carrier interference and attenuates the useful signal. For CP-OFDM several algorithms have been proposed. The authors from [64] propose the joint ML estimation of the time delay offset and the CFO. The estimator provides the time delay offset and the CFO estimates exploiting the presence of the CP. However, the authors assume an AWGN channel and as a result the performance of the proposed algorithm is degraded in a frequency selective fading channel. The same authors propose a CFO estimation algorithm for CP-OFDM in a frequency selective channel in [67], but they assume that the channel impulse response is known by the receiver, which is hardly realistic. In [74] and [65], estimation algorithms are proposed which are based on specially designed training symbols.

In the literature several CFO estimation algorithms for TDS-OFDM can be found. In [69], it is shown that the content of the guard interval in a TDS-OFDM system can be viewed as a CP-OFDM training symbol. The algorithm from [64] is then applied. The authors from [70] introduce the method of composite PN-correlation, which computes the point-wise conjugate product of two phase-shifted PN-correlations. The obtained product is then used to estimate the CFO. In [75], the guard interval of one OFDM block is correlated with the guard interval of the next OFDM block. When the channel only varies slowly during two consecutive OFDM blocks, there is only a difference in phase between the two received guard intervals. This phase difference depends linearly on the CFO and so it can be exploited to estimate the CFO. These algorithms for a TDS-OFDM system only make use of the time domain pilot symbols while there are also frequency domain pilot symbols available in a general KSP-OFDM system. We want to develop some frequency offset estimation algorithms that exploit both the time domain pilot symbols and the frequency domain pilot symbols, but we do not use specially designed training symbols that do not contain any data symbols.

The OFDM blocks are transmitted over a frequency selective channel with impulse response $\mathbf{h}$ with L channel taps. The $N + \nu$ received time-domain samples corresponding to the observation interval of the i-th OFDM block shown in figure 4.1.b are given by

$$\mathbf{z}_i = e^{j2\pi \frac{i(N+\nu)}{N}\epsilon} \mathbf{E}(\epsilon)\, \mathbf{H}_{\text{ch}} \mathbf{s}_i + \mathbf{w}_i. \tag{8.1}$$

The unknown parameter ϵ denotes the CFO between the transmitter and receiver oscillators expressed as a fraction of the intercarrier spacing. We assume that the timing offset is already known and that a rough estimation of the frequency offset already has been performed so that $|\epsilon| < 0.5N/(N+\nu)$. The $(N+\nu) \times (N+\nu)$ channel matrix $\mathbf{H}_{\text{ch}}$ is given by (4.2). The diagonal matrix $\mathbf{E}(\epsilon)$ depends on the unknown CFO and is defined as:

$$(\mathbf{E}(\epsilon))_{k,k} = e^{j2\pi\frac{k}{N}\epsilon}. \tag{8.2}$$

As the performance strongly degrades if a (small) CFO is present, we need an estimate of the CFO ϵ to compensate its presence in (8.1), before the data detection can be performed.

8.2 Existing TDS-OFDM Algorithms

First, we start with reviewing some existing frequency offset estimation algorithms for TDS-OFDM.

8.2.1 Slide Auto-correlation Based Estimation

The principle of the slide auto-correlation has already been introduced in the previous chapter for the time delay offset estimation (see subsection 7.3.1). The authors from [69] also propose a frequency offset estimation algorithm in the same paper. As we have shown in figure 7.2 (and subsection 7.3.1), the guard interval in a TDS-OFDM system can be seen as a CP-OFDM block: the first N_{CP} time domain pilot symbols transmitted during the guard interval are equal to the last N_{CP} time domain pilot symbols $a'_{\text{PN}}(N_{\text{PN}} - N_{\text{CP}})$, ..., $a'_{\text{PN}}(N_{\text{PN}} - 1)$. If we assume transmission over a flat fading channel with channel coefficient h, then we can express the received samples $z_i(N)$, ..., $z_i(N + N_{\text{CP}} - 1)$ corresponding to the first N_{CP} pilot symbols as

$$z_i(N+k) = \sqrt{\frac{N}{N+\nu}} h \cdot a'_{\text{PN}}(N_{\text{PN}} - N_{\text{CP}} + k)\, e^{j2\pi\frac{i(N+\nu)+N+k}{N}\epsilon} + w_i(N+k), \quad k = 0,\ ...,\ N_{\text{CP}} - 1, \tag{8.3}$$

while the received samples $z_i(N + \nu - N_{\text{CP}})$, ..., $z_i(N + \nu - 1)$ corresponding to last N_{CP} pilot symbols of the guard interval are given by

$$z_i(N+\nu-N_{\text{CP}}+k) = \sqrt{\frac{N}{N+\nu}} h \cdot a'_{\text{PN}}(N_{\text{PN}} - N_{\text{CP}} + k)\, e^{j2\pi\frac{(i+1)(N+\nu)-N_{\text{CP}}+k}{N}\epsilon} + w_i(N+\nu-N_{\text{CP}}+k),\ k = 0,\ ...,\ N_{\text{CP}} - 1, \tag{8.4}$$

where h denotes the unknown channel coefficient. The slide auto-correlation is defined as[1]

$$R_{\mathrm{C},i} = \sum_{k=0}^{N_{\mathrm{CP}}-1} \left(z_i\left(N+k\right)\right)^* z_i\left(N+\nu-N_{\mathrm{CP}}+k\right). \tag{8.5}$$

It is easily seen that the useful part of $R_{\mathrm{C},i}$ is given by

$$N_{\mathrm{CP}} \frac{N}{N+\nu} \left|h\right|^2 E_{\mathrm{s}} e^{j2\pi \frac{N_{\mathrm{PN}}}{N}\epsilon} \tag{8.6}$$

and its phase is equal to $2\pi\frac{N_{\mathrm{PN}}}{N}\epsilon$, so the estimate of the frequency offset is given by

$$\hat{\epsilon} = \frac{N}{N_{\mathrm{PN}}} \frac{1}{2\pi} \angle\left(R_{\mathrm{C},i}\right), \tag{8.7}$$

where $\angle\left(x\right)$ returns the argument of the complex number x in the interval $[-\pi,\ \pi[$. From (8.8) it can be seen that the slide auto-correlation based CFO estimator can return the correct value of ϵ, if ϵ is in the interval $\left[-\frac{1}{2}\frac{N}{N_{\mathrm{PN}}},\ \frac{1}{2}\frac{N}{N_{\mathrm{PN}}}\right[$. Now if we have K received OFDM blocks, the estimator is easily extended to use all of the K received blocks:

$$\hat{\epsilon} = \frac{N}{N_{\mathrm{PN}}} \frac{1}{2\pi} \angle\left(\sum_{i=0}^{K-1} R_{\mathrm{C},i}\right) \tag{8.8}$$

8.2.2 Composite PN-Correlation Based Estimation

As already discussed in subsection 7.3.2, the composite PN-correlation based estimation algorithm computes the point wise conjugate product of two phase-shifted PN-correlations.

After transmission over a flat fading channel with channel impulse response equal to h, the received signal samples which correspond to the guard interval of the i-th OFDM block are given by

$$z_i\left(N+k\right) = \sqrt{\frac{N}{N+\nu}} h \cdot a_{\mathrm{g}}\left(k\right) e^{j2\pi\frac{i(N+\nu)+N+k}{N}\epsilon} + w_i\left(N+k\right), \qquad k = 0,\ \ldots,\ \nu-1 \tag{8.9}$$

The N_{PN}-length linear correlation C_1 is computed between the received sam-

[1] unlike the slide auto-correlation (7.4) from subsection 7.3.1, there is no time delay offset

ples $z_i\left(N+N_{\text{pre}}+k\right)$ from (8.9) and $\mathbf{a}_{\text{PN}}$:

$$\begin{aligned} C_{1,i} &= \sum_{k=0}^{N_{\text{PN}}-1} z_i\left(N+N_{\text{pre}}+k\right)\left(a_{\text{PN}}\left(k\right)\right)^* \\ &= \sqrt{\frac{N}{N+\nu}}h\cdot E_{\text{s}}e^{j2\pi\frac{i(N+\nu)+N+N_{\text{pre}}}{N}\epsilon}\sum_{k=0}^{N_{\text{PN}}-1}e^{j2\pi\frac{k}{N}\epsilon} \\ &\quad + \sum_{k=0}^{N_{\text{PN}}-1} w_i\left(N+N_{\text{pre}}+k\right)\left(a_{\text{PN}}\left(k\right)\right)^*. \end{aligned} \tag{8.10}$$

A second N_{PN}-length linear correlation C_2 is obtained by correlating the samples from the next OFDM block $z_{i+1}\left(N+N_{\text{pre}}+k\right)$ with $\mathbf{a}_{\text{PN}}$:

$$\begin{aligned} C_{2,i} &= \sum_{k=0}^{N_{\text{PN}}-1} z_{i+1}\left(N+N_{\text{pre}}+k\right)\left(a_{\text{PN}}\left(k\right)\right)^* \\ &= \sqrt{\frac{N}{N+\nu}}h\cdot E_{\text{s}}e^{j2\pi\frac{(i+1)(N+\nu)+N+N_{\text{pre}}}{N}\epsilon}\sum_{k=0}^{N_{\text{PN}}-1}e^{j2\pi\frac{k}{N}\epsilon} \\ &\quad + \sum_{k=0}^{N_{\text{PN}}-1} w_{i+1}\left(N+N_{\text{pre}}+k\right)\left(a_{\text{PN}}\left(k\right)\right)^*, \end{aligned} \tag{8.11}$$

where the samples $z_{i+1}\left(N+k\right)$ are given by (8.9) (where we have replaced i by $i+1$).

The composite PN-correlation is defined as

$$R_{\text{CPC},i} = \left(C_{1,i}\right)^* C_{2,i}. \tag{8.12}$$

From the definition of $C_{1,i}$ (8.10) and $C_{2,i}$ (8.11) it is easily seen that the useful part of $R_{\text{CPC},i}$ is given by

$$\frac{N}{N+\nu}E_{\text{s}}^2\left|h\right|^2\left|\sum_{k=0}^{N_{\text{PN}}-1}e^{j2\pi\frac{k}{N}\epsilon}\right|^2 e^{j2\pi\frac{N+\nu}{N}\epsilon}. \tag{8.13}$$

The estimate of the CFO is given by

$$\hat{\epsilon} = \frac{N}{(N+\nu)}\frac{1}{2\pi}\angle\left(R_{\text{CPC},i}\right). \tag{8.14}$$

The estimation range of the composite PN-correlation based CFO estimator is equal to: $\epsilon\in\left[-\frac{1}{2}\frac{N}{N+\nu},\ \frac{1}{2}\frac{N}{N+\nu}\right[$. For K received OFDM blocks the extension of the estimator to consider all of the received OFDM blocks is given by

$$\hat{\epsilon} = \frac{N}{(N+\nu)}\frac{1}{2\pi}\angle\left(\sum_{i=0}^{K-1}R_{\text{CPC},i}\right). \tag{8.15}$$

8.2.3 Guard Interval Based Estimation

In [75], the authors propose a CFO estimation algorithm based on the pilot symbols transmitted during the guard intervals of two consecutively transmitted OFDM blocks. It was initially proposed for a TDS-OFDM system, but because it does not demand any special structure about the time domain pilot sequence, it can be applied directly without any changes in a KSP-OFDM system.

The algorithm starts from the received signal vectors corresponding to the guard interval $\mathbf{z}_{\mathrm{g},i} = \left(z_i\left(N\right),\, \ldots,\, z_i\left(N+\nu-1\right)\right)^T$, $i = 0,\, \ldots,\, K-1$, which are defined as

$$\mathbf{z}_{\mathrm{g},i} = e^{j2\pi\frac{i(N+\nu)}{N}\epsilon}\mathbf{E}_{\mathrm{g}}\left(\epsilon\right)\mathbf{H}_{\mathrm{g}}\left(\mathbf{s}_{\mathrm{p}} + \mathbf{s}_{\mathrm{d}}^{(i)}\right) + \mathbf{w}_{\mathrm{g},i} \tag{8.16}$$

where $\mathbf{H}_{\mathrm{g}}$ consists of the last ν rows of $\mathbf{H}_{\mathrm{ch}}$, $\mathbf{E}_{\mathrm{g}}\left(\epsilon\right)$ is a diagonal matrix given by $\left(\mathbf{E}_{\mathrm{g}}\left(\epsilon\right)\right)_{k,k} = e^{j2\pi\frac{N+k}{N}\epsilon}$ and $\mathbf{w}_{\mathrm{g},i}$ collects the noise samples. We compute the product of the Hermitian transpose of $\mathbf{z}_{\mathrm{g},i}$ with $\mathbf{z}_{\mathrm{g},i+1}$ for $i = 0,\, \ldots,\, K-2$. The summation of those $K-1$ products yields a quantity that can be used to estimate ϵ:

$$\sum_{i=0}^{K-2}\mathbf{z}_{\mathrm{g},i}^H\mathbf{z}_{\mathrm{g},i+1} = \left(K-1\right)e^{j2\pi\frac{N+\nu}{N}\epsilon}\left|\mathbf{H}_{\mathrm{g}}\mathbf{s}_{\mathrm{p}}\right|^2 + \sum_{i=0}^{K-2}n_i \tag{8.17}$$

where n_i collects the contributions from $\mathbf{z}_{\mathrm{g},i}^H\mathbf{z}_{\mathrm{g},i+1}$ which depend on the noise and the unknown data symbols. The estimate of ϵ is then given by

$$\hat{\epsilon} = \frac{1}{2\pi}\frac{N}{N+\nu}\angle\left\{\sum_{i=0}^{K-2}\mathbf{z}_{\mathrm{g},i}^H\mathbf{z}_{\mathrm{g},i+1}\right\}. \tag{8.18}$$

Inspecting (8.18), we conclude that the guard interval based estimator can estimate the CFO, if ϵ is in the range $\left[-\frac{1}{2}\frac{N}{N+\nu},\, \frac{1}{2}\frac{N}{N+\nu}\right[$.

8.3 TD Pilot Aided Estimation

Besides the TD pilot symbols in the guard interval there are also pilot symbols available on the pilot carriers, which can be used to estimate the CFO. Therefore, a straightforward extension of the algorithm from [75] would be to consider the total received symbol vectors $\mathbf{z}_i$ instead of only the parts corresponding to the guard intervals $\mathbf{z}_{\mathrm{g},i}$. However, the received signals that correspond to the actual OFDM blocks depend on the unknown data symbols. As a result the contributions of the pilot carriers are heavily contaminated by the interference from the data symbols. So extending the algorithm from [75], so that it also takes the received signals corresponding to the pilot carriers into

account and estimates the CFO based on the total received symbol vectors $\mathbf{z}_i$ results in a less accurate estimate and a degradation of the performance. The advantage of this approach is its robustness against a time delay offset.

The proposed algorithm starts from the received signal samples of K (≥ 2) consecutively transmitted OFDM blocks. We denote them as $\mathbf{z}_0$, ..., $\mathbf{z}_{K-1}$, which are given by (8.1):

$$\begin{aligned} \mathbf{z}_0 &= \mathbf{E}(\epsilon)\,\mathbf{H}_{\text{ch}}\left(\mathbf{s}_{\text{p}} + \mathbf{s}_{\text{d}}^{(0)}\right) + \mathbf{w}_0 \\ &\vdots \\ \mathbf{z}_{K-1} &= e^{j2\pi\frac{(K-1)(N+\nu)}{N}\epsilon}\mathbf{E}(\epsilon)\,\mathbf{H}_{\text{ch}}\left(\mathbf{s}_{\text{p}} + \mathbf{s}_{\text{d}}^{(K-1)}\right) + \mathbf{w}_{K-1} \end{aligned} \tag{8.19}$$

We see that the pilot symbol contribution in $\mathbf{z}_i$ (i.e. the contribution from $\mathbf{s}_{\text{p}}$), with $i = 1, \ldots, K-1$, is equal to the pilot symbol contribution in $\mathbf{z}_0$ multiplied by a factor $e^{j2\pi\frac{i(N+\nu)}{N}\epsilon}$. We will exploit this fact to estimate the CFO ϵ.

When we consider the unknown data symbols as an additional noise term in (8.19), we can rewrite (8.19) as

$$\mathbf{z}_i = e^{j2\pi\frac{i(N+\nu)}{N}\epsilon}\mathbf{E}(\epsilon)\,\mathbf{H}_{\text{ch}}\mathbf{s}_{\text{p}} + \bar{\mathbf{w}}_i \tag{8.20}$$

where $\bar{\mathbf{w}}_i = \mathbf{w}_i + e^{j2\pi\frac{i(N+\nu)}{N}\epsilon}\mathbf{E}(\epsilon)\,\mathbf{H}_{\text{ch}}\mathbf{s}_{\text{d}}^{(i)}$. We compute the product of the Hermitian transpose of $\mathbf{z}_i$ with $\mathbf{z}_{i+1}$ for $i = 0, \ldots, K-2$. The summation of those K products yields a quantity that can be used to estimate ϵ:

$$\sum_{i=0}^{K-2}\mathbf{z}_i^H\mathbf{z}_{i+1} = (K-1)\,e^{j2\pi\frac{N+\nu}{N}\epsilon}\left|\mathbf{H}_{\text{ch}}\mathbf{s}_{\text{p}}\right|^2 + \sum_{i=0}^{K-2} n_i \tag{8.21}$$

where n_i collects the contributions from $\mathbf{z}_i^H\mathbf{z}_{i+1}$ which depend on the noise and the unknown data symbols. The estimate of ϵ is then given by

$$\hat{\epsilon} = \frac{1}{2\pi}\frac{N}{N+\nu}\angle\left\{\sum_{i=0}^{K-2}\mathbf{z}_i^H\mathbf{z}_{i+1}\right\}. \tag{8.22}$$

This CFO estimator has the same estimation range as the guard interval based estimator (8.18): the correct estimate of ϵ if $\epsilon \in \left[-\frac{1}{2}\frac{N}{N+\nu},\ \frac{1}{2}\frac{N}{N+\nu}\right[$.

The main advantage of this approach is that we do not need to know the time delay offset (TDO) first before an estimate of the CFO can be obtained. Figure 8.1 shows the transmitted time domain pilot signal (top of the figure): in every block, the first N samples correspond to the IFFT outputs of the pilot carriers and they are followed by the ν pilot symbols of the guard interval. The bottom of the figure shows the received pilot signal. When there is no TDO, the received vectors corresponding to every OFDM block are given by

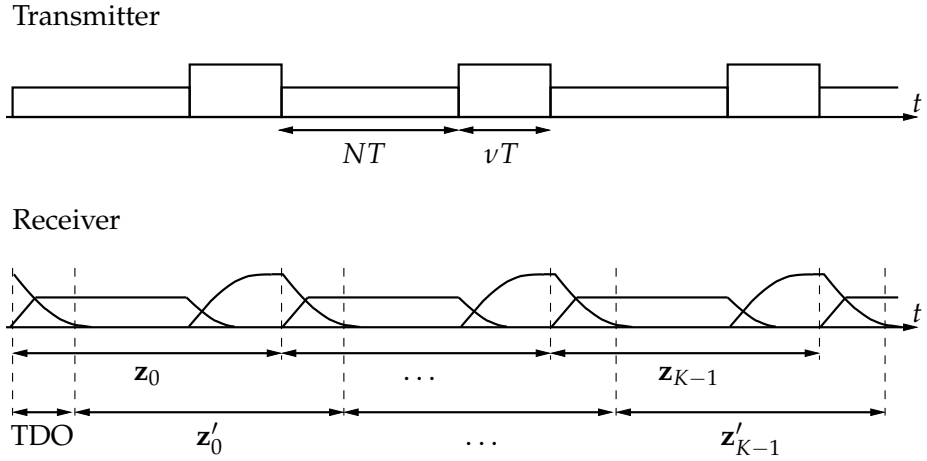

Figure 8.1: The time domain pilot signal: transmitted signal (top), and received signal without and with a TDO (bottom).

$\mathbf{z}_i$ (8.20). In the presence of a TDO, the pilot contribution in each received vector, which we now denote as $\mathbf{z}'_i$, is a circular shift of the pilot contribution in $\mathbf{z}_i$ with $i = 0, \ldots, K-1$ as can be seen from figure 8.1. The summation of the K products of the Hermitian transpose of $\mathbf{z}'_i$ with $\mathbf{z}'_{i+1}$, $i = 0, \ldots, K-2$, yields

$$\sum_{i=0}^{K-2} \left(\mathbf{z}'_i\right)^H \mathbf{z}'_{i+1} = (K-1)\, e^{j2\pi\frac{N+\nu}{N}\epsilon} \left|\mathbf{H}_{\text{ch}}\mathbf{s}_{\text{p}}\right|^2 + \sum_{i=0}^{K-2} n'_i, \tag{8.23}$$

which is equal to (8.21) except for the contribution from the unknown data symbols and the noise. The estimate of the CFO ϵ is obtained by applying the same estimator as (8.22) where $\mathbf{z}_i$ is replaced by $\mathbf{z}'_i$ for $i = 0, \ldots, K-2$.

8.4 FD Pilot Aided Estimation

From the previous section we know that estimating the CFO based on the time domain received signal corresponding to both the guard interval pilot symbols and the pilot symbols transmitted on the pilot carriers, results in a degradation compared to estimating the CFO based on the received signal corresponding to the guard interval pilot symbols only. From the previous chapters, we know that for channel estimation (see section 6.6) and TDO estimation (see section 7.6) and without the presence of an CFO, the best results are obtained if the received signal is transformed to the frequency domain. The main problem in the time domain is the presence of the unknown data symbols, which interfere with the pilot contributions. In the frequency domain, the pilot carriers and the data carriers are orthogonal, so the FFT outputs of the received signal

on the pilot carriers only depend on the pilot symbols (from both the pilot carriers and the guard interval). However, in the presence of an CFO it has been shown in [73], that the orthogonality of the carriers is destroyed so there exists ICI. But fortunately for small values the CFO, the interference of the unknown data symbols which is present in the received signal on the pilot carriers is a lot weaker than in the time domain signal. So in this section we are deriving an CFO estimation algorithm that operates in the frequency domain.

First we transform the received vectors $\mathbf{z}_i,\ i = 0,\ \ldots,\ K-1$ to the frequency domain. To achieve this, the last ν samples of each vector are added to the first ν samples, and the first N samples of the resulting vectors are then applied to an FFT:

$$\tilde{\mathbf{z}}_i = \mathbf{F}\mathbf{\Omega}\mathbf{z}_i, \tag{8.24}$$

where the matrix $\mathbf{\Omega}$ represents the summation and is defined as (4.19) and $\mathbf{F}$ is the FFT matrix (2.5). The output of the FFT of OFDM block i at carrier n is given by

$$\begin{aligned}\tilde{z}_i(n) = e^{j2\pi\frac{i(N+\nu)}{N}\epsilon}\Bigg(&\tilde{H}(\epsilon,\ n)\sum_{m=0}^{M-\nu-1} a_{\mathrm{c}}(m)\, I_{\alpha_m-n}(\epsilon) \\ &+ \tilde{H}(\epsilon,\ n)\sum_{m=0}^{N+\nu-M-1} a_{\mathrm{d}}^{(i)}(m)\, I_{\beta_m-n}(\epsilon) + \sum_{l=0}^{L-1}\tilde{B}_{\mathrm{g}}(\epsilon,\ n,\ l)\, h(l)\Bigg) + \tilde{w}_i(n)\end{aligned} \tag{8.25}$$

where $I_m(\epsilon)$, $\tilde{H}(\epsilon,\ n)$ and $\tilde{B}_{\mathrm{g}}(\epsilon,\ n,\ l)$ are given by

$$I_m(\epsilon) = \frac{1}{\sqrt{N(N+\nu)}}\frac{1-e^{j2\pi(m+\epsilon)}}{1-e^{j2\pi\frac{(m+\epsilon)}{N}}} \tag{8.26}$$

$$\tilde{H}(\epsilon,\ n) = \sum_{l=0}^{L-1} h(l)\, e^{-j2\pi\frac{(n-\epsilon)l}{N}} \tag{8.27}$$

$$\begin{aligned}\tilde{B}_{\mathrm{g}}(\epsilon,\ n,\ l) = \frac{1}{\sqrt{N+\nu}}\Bigg(&\sum_{k=0}^{l-1} a_{\mathrm{g}}(\nu+k-l)\, e^{j2\pi\frac{(\epsilon-n)k}{N}} \\ &+ \sum_{k=l}^{\nu-1} a_{\mathrm{g}}(k-l)\, e^{j2\pi\frac{(\epsilon-n)(N+k)}{N}}\Bigg)\end{aligned} \tag{8.28}$$

and α_m and β_m are carrier indices belonging to the subsets of carrier indices of pilot carriers S_{p} (5.1) and data carriers S_{d} (5.2), respectively. The samples $\tilde{w}_i(n)$, with $n = 0,\ \ldots,\ N-1$, are Gaussian noise samples with zero mean and have an autocorrelation matrix $\mathbf{R}_{\tilde{\mathbf{w}}}$ defined as (4.25). The factor $I_0(\epsilon)$ shows that the useful component in (8.25) is attenuated due to the CFO. Secondly, it

is also clear from (8.25) that the CFO also causes ICI as in (8.25), not only the symbol $a_{\mathrm{c}}(n)$ (or $a_{\mathrm{d}}^{(i)}(n)$) occurs but also other symbols.

We collect the $M-\nu$ FFT outputs from the K OFDM blocks corresponding to the set S_{p} of pilot carriers in the vectors $\tilde{\mathbf{z}}_{2,i}$, $i = 0, \ldots, K-1$:

$$\begin{aligned} \tilde{\mathbf{z}}_{2,i} &= \left[\tilde{z}_i(\alpha_0), \ldots, \tilde{z}_i(\alpha_{M-\nu-1})\right]^T \\ &= e^{j2\pi\frac{i(N+\nu)}{N}\epsilon}\left(\tilde{\mathbf{b}} + \tilde{\mathbf{a}}_i\right) + \tilde{\mathbf{w}}_{2i} \end{aligned} \tag{8.29}$$

where $\tilde{\mathbf{b}}$ is the vector that contains the contributions from both time and frequency domain pilots. The vector $\tilde{\mathbf{a}}_i$ collects the contributions from the unknown data symbols from the i-th OFDM block and $\tilde{\mathbf{w}}_{2i}$ contains the noise samples at the pilot carrier positions of the i-th block. The summation over i of the multiplications of the Hermitian transpose of $\tilde{\mathbf{z}}_{2,i}$ with $\tilde{\mathbf{z}}_{2,i+1}$ for $i = 0, \ldots, K-2$ results in a function which is used to estimate ϵ:

$$\sum_{i=0}^{K-2} \tilde{\mathbf{z}}_{2,i}^H \tilde{\mathbf{z}}_{2,i+1} = (K-1)\, e^{j2\pi\frac{(N+\nu)}{N}\epsilon}\left|\tilde{\mathbf{b}}\right|^2 + \sum_{i=0}^{K-2} \tilde{n}_i \tag{8.30}$$

where $\tilde{n}_i$ collects the contributions from the unknown data symbols and the noise samples from $\tilde{\mathbf{z}}_{2,i}^H \tilde{\mathbf{z}}_{2,i+1}$. The estimate of ϵ is then given by

$$\hat{\epsilon} = \frac{1}{2\pi}\frac{N}{N+\nu}\angle\left\{\sum_{i=0}^{K-2} \tilde{\mathbf{z}}_{2,i}^H \tilde{\mathbf{z}}_{2,i+1}\right\}. \tag{8.31}$$

The provided estimate is limited to the range $\left[-\frac{1}{2}\frac{N}{N+\nu}, \frac{1}{2}\frac{N}{N+\nu}\right[$, so the CFO than can be estimated with the FD pilot aided CFO estimator (8.31) should also be in this range.

8.5 Decision Aided FD Estimation

In a decision aided CFO estimator, initially, an estimate of the CFO $\hat{\epsilon}_0$ is obtained by applying some CFO estimation algorithm (see for example the proposed algorithms in sections 8.3 and 8.4). This estimate is used to compensate the CFO in the time domain. In a second step, an estimate of the channel impulse response $\mathbf{h}$ is obtained based on the received signal with compensated CFO using one of the data aided algorithms from chapter 6. The channel estimate is then used to detect the transmitted data symbols with one of the detectors from chapter 4.

We can now start an iterative algorithm which iterates between the following processes: i) obtaining an estimate of the remaining CFO using the latest obtained estimate of $\mathbf{h}$ and the detected data symbols, ii) re-estimating $\mathbf{h}$ after compensation of the remaining CFO and using the detected data symbols, and

iii) refining the detection of the data symbols using the most recent estimate of $\mathbf{h}$. This procedure is repeated until convergence is reached.

In the p-th iteration, the remaining CFO is denoted as ϵ_p and defined as

$$\epsilon_p = \epsilon - \sum_{k=0}^{p-1} \hat{\epsilon}_k. \tag{8.32}$$

The FFT output of OFDM block i at carrier n, denoted by $\tilde{z}_i^{(p)}(n)$, is given by (8.25) where we replace ϵ by ϵ_p. When ϵ_p is very small, the inter carrier interference can be neglected: $I_m(\epsilon_p) \approx 0$, for $m \neq 0$, which yields for the FFT outputs at the pilot carrier positions:

$$\tilde{z}_i^{(p)}(\alpha_m) = e^{j2\pi \frac{i(N+\nu)}{N}\epsilon_p} \Bigg(a_{\mathrm{c}}(m)\, I_0(\epsilon_p)\, \tilde{H}(\epsilon_p,\, \alpha_m) + \sum_{l=0}^{L-1} \tilde{B}_{\mathrm{g}}(\epsilon_p,\, \alpha_m,\, l)\, h(l) \Bigg) + \tilde{w}_i^{(p)}(\alpha_m), \tag{8.33}$$

where $\alpha_m \in S_{\mathrm{p}}$ (5.1). The signal component of $\tilde{z}_i^{(p)}(\alpha_m)$ is the product of a factor which depends on the block index i and a factor which is independent of the block index. We exploit this fact to estimate the CFO.

The FFT outputs at the data carrier positions are given by

$$\tilde{z}_i^{(p)}(\beta_m) = e^{j2\pi \frac{i(N+\nu)}{N}\epsilon_p} \Bigg(a_{\mathrm{d}}^{(i)}(m)\, I_0(\epsilon_p)\, \tilde{H}(\epsilon_p,\, \beta_m) + \sum_{l=0}^{L-1} \tilde{B}_{\mathrm{g}}(\epsilon_p,\, \beta_m,\, l)\, h(l) \Bigg) + \tilde{w}_i^{(p)}(\beta_m), \tag{8.34}$$

where $\beta_m \in S_{\mathrm{d}}$ (5.2). The signal component of $\tilde{z}_i^{(p)}(\beta_m)$ can not be expressed as the product of a factor which depends on the block index and a factor which is independent of it, so we have to modify the FFT outputs at the data carrier positions first. There are two possible options: i) subtracting the contribution from the unknown data symbols (using previous data symbol decisions), or ii) subtracting the contribution from the guard interval pilot symbols and multiplying the results with the complex conjugates of the latest obtained hard decisions of the unknown data symbols. One can prove that the power ratio of the data symbols and the guard interval pilot symbols contribution in $\tilde{z}_i^{(p)}(\beta_m)$ is approximately equal to N/ν. Since N is usually much larger than ν, the second option will lead to a smaller estimation error as the variance of the guard interval pilots is smaller. Therefore, we choose to follow the second option.

First, we have to subtract the contribution from the guard interval pilot samples[2]:

$$\tilde{y}_i^{(p)}(\beta_m) = \tilde{z}_i^{(p)}(\beta_m) - \sum_{l=0}^{L-1} B_{\mathrm{g}}(0,\ \beta_m,\ l)\,\hat{h}^{(p-1)}(l), \tag{8.35}$$

where $\beta_m \in S_{\mathrm{d}}$ and $\hat{h}^{(p-1)}(l)$, $l = 0, \ldots, L-1$, are the estimated samples of the channel impulse response obtained in the $(p-1)$-th iteration. We assume that the estimated channel impulse response is equal to the true channel impulse response so that the contribution of the guard interval can be perfectly removed:

$$\tilde{y}_i^{(p)}(\beta_m) = e^{j2\pi\frac{i(N+\nu)}{N}\epsilon_p} a_{\mathrm{d}}^{(i)}(m)\, I_0(\epsilon_p)\, \tilde{H}(\epsilon_p,\ \beta_m) + \tilde{w}_i^{(p)}(\beta_m). \tag{8.36}$$

Multiplying $\tilde{y}_i^{(p)}(\beta_m)$ by the complex conjugate of the latest obtained hard decision of $a_{\mathrm{d}}^{(i)}(m)$ results in

$$\tilde{y}_i^{'(p)}(\beta_m) = e^{j2\pi\frac{i(N+\nu)}{N}\epsilon_p} E_{\mathrm{s}} I_0(\epsilon_p)\, \tilde{H}(\epsilon_p,\ \beta_m) + \left(a_{\mathrm{d}}^{(i)}(m)\right)^* \tilde{w}_i^{(p)}(\beta_m), \tag{8.37}$$

where we assume that the decisions about the data symbols from the previous iteration are correct, and x^* denotes the complex conjugate of the complex number x.

Both (8.33) and (8.37) correspond to the following model:

$$\tilde{u}_i(n) = e^{j2\pi\frac{i(N+\nu)}{N}\epsilon_p} x(n) + \tilde{v}_i(n), \tag{8.38}$$

where $x(n)$ can be seen as an unknown coefficient and $\tilde{v}_i(n)$ is the additive noise. Note that $x(n)$ does not depend on the block index i. Rewriting everything in vector format yields

$$\tilde{\mathbf{u}}_i = e^{j2\pi\frac{i(N+\nu)}{N}\epsilon_p}\mathbf{x} + \tilde{\mathbf{v}}_i, \tag{8.39}$$

where the components of the vector $\tilde{\mathbf{u}}_i$ are defined as:

$$\tilde{u}_i(n) = \begin{cases} \tilde{z}_i^{(p)}(n) & n \in S_{\mathrm{p}} \\ \tilde{y}_i^{'(p)}(n) & n \in S_{\mathrm{d}}. \end{cases} \tag{8.40}$$

The next step is to obtain the least squares estimates of ϵ_p and $\mathbf{x}$, which are given by

$$(\hat{\epsilon}_p,\ \hat{\mathbf{x}}) = \arg\min_{\epsilon_p,\ \mathbf{x}} \sum_{i=0}^{K-1} \left| \tilde{\mathbf{u}}_i - e^{j2\pi\frac{i(N+\nu)}{N}\epsilon_p}\mathbf{x} \right|^2. \tag{8.41}$$

[2]Again we assume that ϵ_p is small: $e^{j2\pi\frac{i(N+\nu)}{N}\epsilon_p} B_{\mathrm{g}}(\epsilon_p,\ \beta_m,\ l) \approx B_{\mathrm{g}}(0,\ \beta_m,\ l)$.

The estimate $\hat{\mathbf{x}}$ can be expressed as a function of ϵ_p and is given by

$$\hat{\mathbf{x}}\left(\epsilon_p\right) = \frac{1}{K}\sum_{i=0}^{K-1} e^{-j2\pi\frac{i(N+\nu)}{N}\epsilon_p}\tilde{\mathbf{u}}_i. \tag{8.42}$$

Substituting $\mathbf{x}$ by its estimate (8.42) in (8.41) yields

$$\hat{\epsilon}_p = \arg\max_{\epsilon_p}\left|\sum_{i=0}^{K-1} e^{-j2\pi\frac{i(N+\nu)}{N}\epsilon_p}\tilde{\mathbf{u}}_i\right|^2 \tag{8.43}$$

where we have omitted irrelevant terms and a constant factor. Rearranging the terms of the squared magnitude results in[3]

$$\hat{\epsilon}_p = \arg\max_{\epsilon_p}\sum_{m=1}^{K-1} 2\Re\left\{R\left(m\right)e^{-j2\pi\frac{N+\nu}{N}m\epsilon_p}\right\} \tag{8.44}$$

where the correlation $R\left(m\right)$ is defined as

$$R\left(m\right) = \sum_{i=0}^{K-1-m}\tilde{\mathbf{u}}_i^H\tilde{\mathbf{u}}_{i+m}. \tag{8.45}$$

Now we differentiate the function from (8.44) with respect to ϵ_p and equate the derivative to zero:

$$\sum_{m=1}^{K-1}\Im\left\{mR\left(m\right)e^{-j2\pi\frac{N+\nu}{N}m\epsilon_p}\right\} = 0. \tag{8.46}$$

This can be rewritten as

$$\sum_{m=1}^{K-1} m\left|R\left(m\right)\right|\sin\left[\angle\left(R\left(m\right)\right) - 2\pi\frac{N+\nu}{N}m\epsilon_p\right] = 0, \tag{8.47}$$

We assume that the noise contribution in $R\left(m\right)$ is small so that $\angle\left(R\left(m\right)\right) \approx 2\pi\frac{N+\nu}{N}m\epsilon_p$. As the function $\sin\left(x\right)$ can be approximated by x for small values of x, expression (8.47) can be approximated by

$$\sum_{m=1}^{K-1} m\left|R\left(m\right)\right|\left(\angle\left(R\left(m\right)\right) - 2\pi\frac{N+\nu}{N}m\epsilon_p\right) = 0. \tag{8.48}$$

Solving this equation with respect to ϵ_p yields the estimate of ϵ_p. To expand the range of possible values of ϵ_p that can be estimated and / or to reduce the numerical complexity of the algorithm, the interval of the summation index

[3] We have omitted the term for $m = 0$, because it is independent of ϵ_p

m can be limited to $[P_1,\ P_2]$, where $1 \leq P_1 \leq P_2 \leq K-1$. The estimate of ϵ_p is then given by

$$\hat{\epsilon}_p = \frac{1}{2\pi} \frac{N}{N+\nu} \frac{\sum_{m=P_1}^{P_2} m \left|R\left(m\right)\right| \angle\left(R\left(m\right)\right)}{\sum_{m=P_1}^{P_2} m^2 \left|R\left(m\right)\right|}. \tag{8.49}$$

This estimator can only be applied when the remaining CFO ϵ_p (8.32) is in the range $\left[-\frac{1}{2}\frac{1}{P_2}\frac{N}{N+\nu},\ \frac{1}{2}\frac{1}{P_2}\frac{N}{N+\nu}\right[$, which does not necessarily have to be a big problem provided that the previous estimates are good enough so that the remaining CFO is small enough. A special case of the algorithm is obtained when we choose $P_1 = P_2 = P$, where $1 \leq P \leq K-1$. This yields for the estimate of ϵ_p:

$$\hat{\epsilon}_p = \frac{1}{2\pi} \frac{N}{N+\nu} \frac{\angle\left(R\left(P\right)\right)}{P}. \tag{8.50}$$

The estimation range of this special case is equal to: $\left[-\frac{1}{2}\frac{1}{P}\frac{N}{N+\nu},\ \frac{1}{2}\frac{1}{P}\frac{N}{N+\nu}\right[$.

In the next step, the remaining CFO is compensated using the estimate $\hat{\epsilon}_p$ and a new estimate of the channel is obtained per block i, $i = 0,\ \ldots,\ K-1$, by applying one of the decision aided channel estimation algorithms from chapter 6.

8.6 Numerical Results

In this section, the performance of the proposed frequency estimation algorithms is evaluated. We consider a frequency selective Rayleigh fading channel with an impulse response given by (4.29) consisting of $L = 50$ channel taps with equal variance $1/L$. For the pilots transmitted on the carriers, the comb-type pilot arrangement from section 5.2 is used. The pilot symbols are randomly selected QPSK symbols. We consider an OFDM system with $N = 1024$ carriers and a guard interval of length $\nu = 100$. Besides the 100 pilot symbols in the guard interval, an additional 100 carriers are selected as pilot carriers which means that a total number of $M = 200$ pilot symbols are transmitted. The data symbols are QPSK symbols. For the CFO estimation we consider $K = 10$ received OFDM blocks. It is assumed that the simulated OFDM system does not suffer from any time delay offset.

8.6.1 MSE

First we evaluate the performance of the estimators in terms of the MSE, which is defined as:

$$MSE\left(p\right) = \mathbb{E}\left[\left|\epsilon - \sum_{k=0}^{p} \hat{\epsilon}_k\right|^2\right], \tag{8.51}$$

where p is the iteration number, so $MSE\left(p\right)$ is the MSE after p iterations.

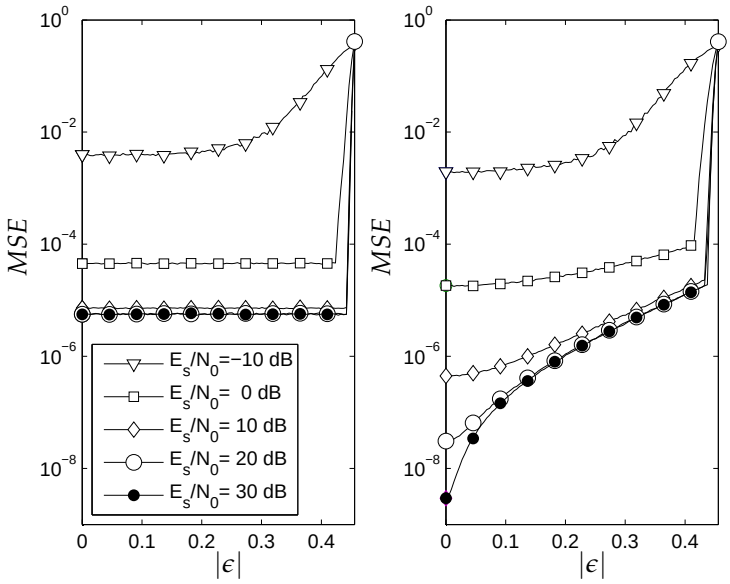

Figure 8.2: MSE versus $|\epsilon|$ for guard interval based (left) and FD pilot aided (right) estimation, $\nu = 100$, $N = 1024$, $M = 200$, $K = 10$.

8.6.1.1 Pilot Aided Estimation

Figure 8.2 shows the MSE of the guard interval based algorithm (8.18) which was originally proposed for TDS-OFDM in [75] and our proposed FD pilot aided estimation algorithm (8.31) versus $|\epsilon|$ for different values of E_s/N_0. In theory, the estimators should be able to estimate the CFO if $|\epsilon| < 0.5\frac{N}{(N+\nu)}$. However, the simulations results indicate that both estimators have a bad performance for $|\epsilon| > 0.3$. This is due to the discontinuity of the $\angle\{.\}$ function in (8.18) and (8.31): when x is slightly bigger than π, the argument $\angle\{\exp(jx)\}$ returns $x - 2\pi$ instead of x (a similar effect can be seen for x slightly less than $-\pi$). Therefore for the following simulation results we will assume that ϵ is randomly selected from a uniform distribution over the interval $[-0.3,\ 0.3]$. For the guard interval based estimation algorithm (8.18) we see that the MSE is only weakly dependent on ϵ and for high SNR, it even becomes independent of the SNR. The performance of the FD pilot aided estimator (8.4) on the contrary, strongly depends on the value of ϵ, especially for high SNR. For large values of ϵ, the MSE becomes independent of the SNR. As a result, both estimators will show an error floor for high values of the SNR when we plot the MSE versus the E_s/N_0, as can be seen from figure 8.3. This error floor is caused by the unknown data symbols, which interfere with the pilot symbols. We see that the FD pilot aided estimation algorithm (8.4) yields the lowest

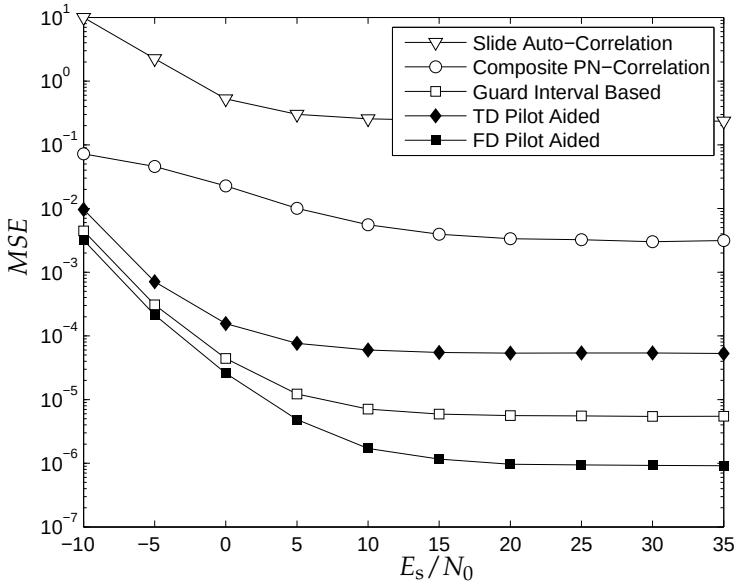

Figure 8.3: MSE as a function of E_s/N_0, $\nu = 100$, $N = 1024$, $M = 200$, $K = 10$.

error floor. This behavior is expected because for very small values of ϵ, the ICI will become very small. In that case the interference of the unknown data symbols will be much smaller in the FD (and eventually equal to zero when $\epsilon = 0$), while the interference in the TD is independent of the presence of a CFO.

Figure 8.3 compares the performance of the two proposed pilot aided estimators (8.22) and (8.31) in terms of the MSE. Additionally the results for the discussed TDS-OFDM estimators from section 8.2 are also shown. The performance of both the slide auto-correlation based estimation algorithm (8.8) and composite PN-correlation based estimation algorithm (8.15) is rather poor. Both show an error floor which is rather high, making the obtained CFO estimates not accurate enough. Both estimators were derived in a flat fading channel and as a result they do not function very well in a multipath fading environment. The third considered TDS-OFDM estimation algorithm, e.g. the guard interval based estimator (8.18), exhibits a much better performance. The error floor is several orders of magnitude lower than the error floors of the two other considered TDS-OFDM estimation algorithms. Next we consider the TD pilot aided estimator (8.22) for the estimation of the CFO. The results for the TD pilot aided estimator (8.22) confirm what was already said in section 8.3: exploiting the contribution of the pilot carriers in the TD received signal besides the contribution of the guard interval pilot symbols de-

teriorates the quality of the CFO estimate compared to only considering the received signal corresponding to the guard interval pilot symbols as is done in the guard interval based estimation algorithm (8.18). The received signal corresponding to the actual OFDM blocks is dominated by the contributions of the unknown data symbols besides the contribution from the pilot carriers and the channel noise. The interference from the data symbols prevents the estimation algorithm from obtaining any improvement in performance compared to the guard interval based estimator (8.18). The TD pilot aided estimator still results in a large improvement in terms of the MSE compared with the other two discussed TDS-OFDM CFO estimation algorithms (8.8) and (8.15). Finally, the FD pilot aided estimation algorithm (8.31) is applied to estimate the CFO. It outperforms all of the previously discussed CFO estimation algorithms in terms of the MSE. For high values of E_s/N_0 this algorithm still exhibits an error floor, which is caused by the presence of the unknown data symbols. However, contrary to the situation in the time domain, the interference strongly depends on the CFO: for small values of $|\epsilon|$ the ICI becomes small. Furthermore, the interference at one pilot carrier is mainly caused by the adjacent data carriers, which results in a less severe interference compared to the received signal in the time domain.

8.6.1.2 Decision Aided Estimation

In this part we investigate how the decision aided CFO estimation algorithm from section 8.5 performs. The initial estimate of the CFO is provided by the guard interval based estimator (8.18) or the FD pilot aided estimator (8.31)[4]. Every iteration, after the CFO estimation, the channel needs to be estimated and the data symbols need to be detected. In this section we consider the FD pilot aided estimator (6.64) to provide the initial channel estimate. In the other iterations, the FD hard decision aided channel estimator (6.77) is used. The data symbols are detected using the symbol by symbol FD data detector (4.28). We recapitulate from section 8.5 that the proposed estimators (8.49) and (8.50) still depend on some parameters, i.e. P_1 and P_2, which have an influence on the performance and on the computational complexity.

To properly evaluate the performance of the decision-directed CFO estimation algorithm, we have first done some simulations to determine the optimal values for the parameter P_1 and P_2. We have considered two situations:

- in the first case we select P_1 equal to 1 and let P_2 be equal to P (which corresponds to the more general estimator (8.49)),
- in the second case we select both P_1 and P_2 equal to P (which is the special case (8.50)).

[4]In the figures, 'GI init.' means 'initialized by the guard interval based estimator', while 'FD init.' means 'initialized by the FD pilot aided estimator'.

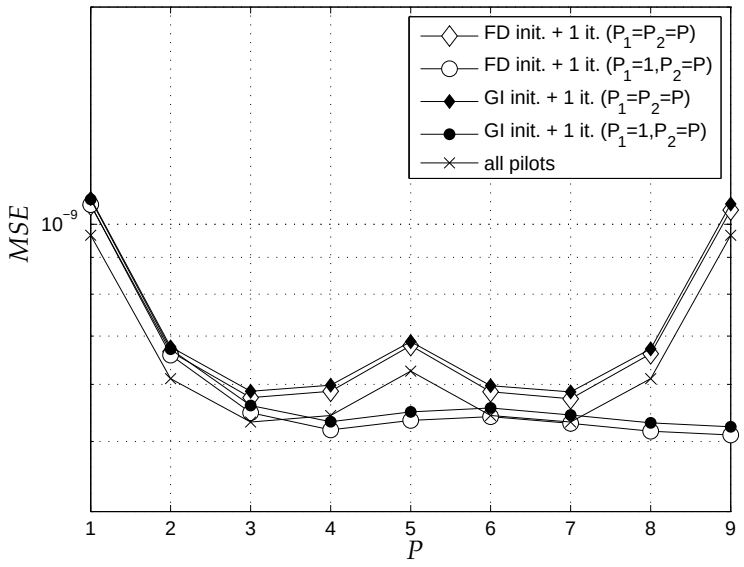

Figure 8.4: MSE of the decision aided estimators versus P, for $E_s/N_0 = 25$ dB, $\nu = 100$, $N = 1024$, $M = 200$, $K = 10$.

The parameter P varies from 1 to $K-1$. Figure 8.4 shows the MSE as a function of P for $E_s/N_0 = 25$ dB. For the first case, the lowest MSE is reached when P_2 is chosen equal to $K-1 = 9$, which is the maximum value for P_2. The curve is almost independent of P when $P \in [4,\ 9]$, so choosing $P_2 = 4$ instead of the optimal value 9, results in a near optimum performance in terms of the MSE while the numerical complexity is reduced.

For the second case, the performance of the estimator is optimal when both P_1 and P_2 are chosen equal to 3 or 7. To explain this, we look at the performance of the all pilots case of the estimator (8.50), also shown in the figure. It can be observed that the performance in the second case is (essentially) proportional to that of the all pilots case. In the appendix 8.A the MSE performance of the all pilots case is derived. The minimum value of the MSE is indeed reached for P_1 and P_2 both equal to the values of P closest to $10/3$ and $20/3$. The second case gives rise to a small loss in performance compared to the first case. These results are independent of which initialization algorithm was used first: from figure 8.4 it can be seen that initializing the decision aided algorithms with the guard interval based estimator (8.18) or the FD pilot aided estimator (8.31) yields essentially the same result, with the decision aided estimator initialized by the FD pilot aided estimator resulting in a marginally lower MSE.

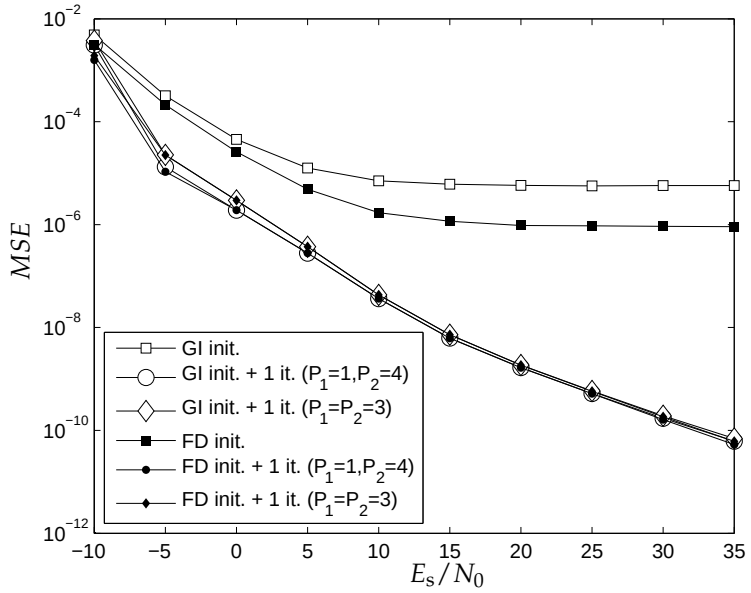

Figure 8.5: MSE as a function of E_s/N_0 (decision aided estimation) , $N = 1024$, $\nu = 100$, $L = 50$, $M = 200$, $K = 10$.

The MSE of the decision aided CFO estimation algorithms (8.49) and (8.50) as a function of E_s/N_0 is shown in figure 8.5. Both the guard interval based estimator (8.18) and the FD pilot aided estimator (8.31) are used to obtain the initial estimate of the CFO to start up the decision aided estimators. The results are shown for $P_1 = 1$ and $P_2 = 4$; and $P_1 = P_2 = 3$. Compared to the results of the initialization algorithms, a large improvement is reached after only one iteration: independent of which initialization algorithm has been applied, the error floor is completely removed for the considered range of E_s/N_0 values.

Both considered choices for P_1 and P_2 result in a similar performance. Only for low to moderate E_s/N_0, there is a small difference in performance.

8.6.2 BER

Figure 8.6 shows the BER results as a function of E_b/N_0, where E_b is the transmitted energy per information bit. We have only considered our proposed FD initialization method to obtain these results, because we have seen from figure 8.5 that the performance of our proposed decision-directed algorithm is independent of the initialization algorithm.

For the uncoded case, the BER of a receiver which only applies the FD

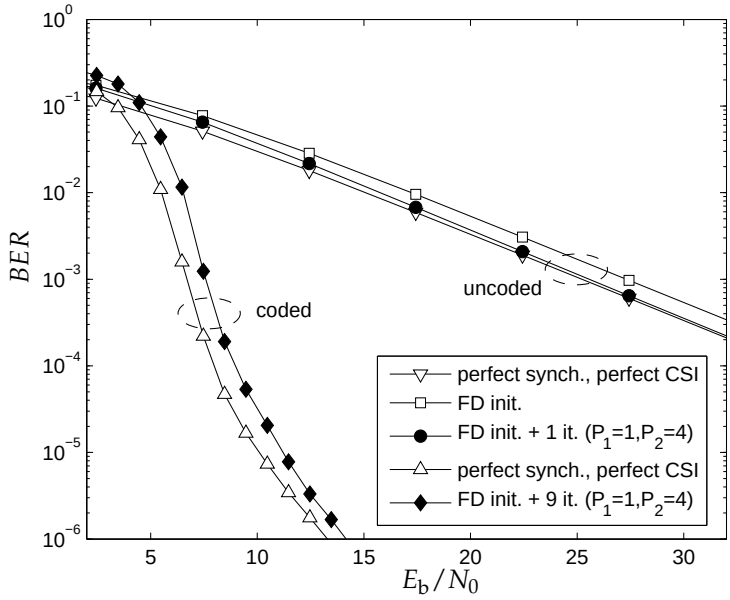

Figure 8.6: BER as a function of E_b/N_0, $N = 1024$, $\nu = 100$, $L = 50$, $M = 200$, $K = 10$.

initialization algorithm for the estimation of the CFO, is close to the case when the CFO is perfectly known (but with estimated channel). After one iteration, the decision aided algorithm results in a performance which is close to the BER of a receiver with perfect knowledge about the channel and the CFO. Additional iterations are not necessary for the considered range of E_b/N_0 values, as the potential improvements are negligible.

During the derivation of the iterative decision-directed algorithm, we have made some assumptions concerning the correctness of the obtained channel impulse response estimate and the hard decisions of the data symbols. The BER curves show that those assumptions are valid for moderate to high E_b/N_0 values.

For the coded case, we have considered a turbo code that consists of two identical 16-state rate 1/2 recursive systematic convolutional codes with generator sequences $(37)_8$ and $(21)_8$ (in octal notation) through a uniform random interleaver of length 919. The parity bits are punctured to obtain an overall rate of approximately 1/2. Every iteration, the turbo decoder provides the necessary hard decisions on the transmitted data symbols for the decision aided CFO and channel estimation algorithm. After 9 iterations convergence is reached. The application of our proposed algorithm results in a small loss in performance compared to the BER of a receiver with perfect knowledge

about the channel and the CFO. The gap between the two curves is less than 1 dB.

8.6.3 Optimization of the Number of Iterations

To obtain the results from figure 8.6, every iteration the decision aided CFO estimation algorithm provides an estimate of the remaining CFO, which is then used to update the channel estimate and finally a new decoding iteration is performed. Similarly like in subsection 6.10.4, the number of the iterations of the decision aided CFO estimator (8.49) can be reduced to lower the computational complexity in a coded system. Here we consider the same turbo code as in the previous subsection 8.6.2. To determine the number of necessary iterations of the decision aided CFO estimator we first look at the MSE versus the number of performed iterations. Here we consider the decision aided FD estimator (8.49) with parameters P_1 and P_2 equal to 1 and 4 respectively. Figure 8.7 shows the MSE of the decision directed CFO estimation algorithm as a function of the number of iterations, for two different values of E_s/N_0, i.e. 7 dB and 10 dB. Iteration '0' corresponds to the MSE of the FD pilot aided CFO estimator (8.31). For both values of E_s/N_0, the FD decision aided CFO estimator converges after 4-5 iterations, so it is not necessary to do more than 5 estimator iterations. After only 1 iteration of the FD decision aided CFO estimator, we see that the MSE is already very low ($\leq 5e-7$), so we are going to reduce the number of CFO estimation iterations to 1 (which means that after the pilot aided initial estimate, only 1 extra estimate of the remaining CFO is obtained). The resulting BER for the FD detector from section 4.2 is shown in figure 8.8 as a function of E_b/N_0. The BER results for a perfectly synchronized receiver with perfect channel knowledge and for a receiver which updates the CFO estimate every iteration (labeled as 'FD init. + 9 it.' in the figure) are also added. We can see that only updating the CFO estimate once yields essentially the same performance as updating the CFO estimate every iteration. Only when E_b/N_0 is in the range [6 dB, 9 dB], there is a small degradation visible.

8.7 Conclusion

In this chapter we have looked into CFO estimation for a KSP-OFDM system. An uncompensated CFO causes interference between the different carriers and attenuates the useful signal component. As a result the BER performance can be severely degraded. First we have reviewed some existing estimation algorithm that have been proposed in a TDS-OFDM context. Two of them, i.e. the slide auto-correlation based estimator and the composite PN-correlation based estimator assume a flat fading channel and as a result they do not perform very well in a multipath fading environment. The guard interval based

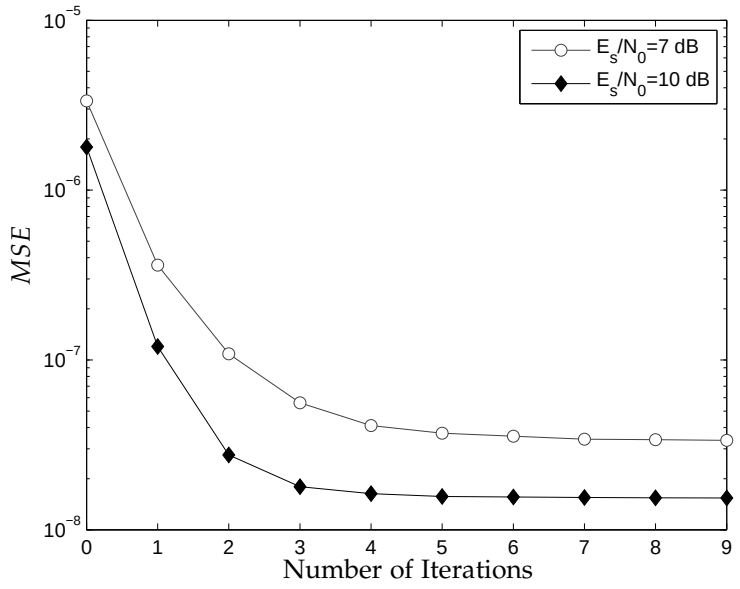

Figure 8.7: MSE as a function of the number of iterations , $N = 1024$, $\nu = 100$, $L = 50$, $M = 200$, K=10.

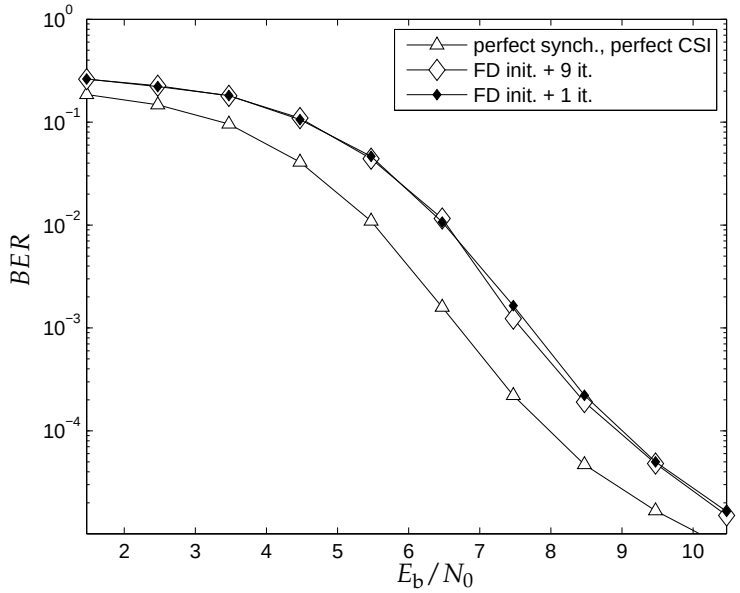

Figure 8.8: BER as a function of E_b/N_0, $N = 1024$, $\nu = 100$, $L = 50$, $M = 200$, $K = 10$.

estimation algorithm, on the other hand, yields a satisfying performance and can be applied in a KSP-OFDM system. As the name already suggests, the CFO is estimated based solely on the presence of the guard interval pilot symbols. To exploit also the pilot carriers which are typically available in a KSP-OFDM system, we have proposed the TD pilot aided estimator. The results show that exploiting the presence of the pilot carriers in the time domain, results in a degradation in performance compared to the guard interval based estimator. However, due to the periodic nature of the TD transmitted pilot signal (guard interval pilot symbols + pilot carriers), the TD pilot aided estimator is robust against any time delay offset and therefore this algorithm can be applied to first obtain an estimate of the CFO before the time delay offset is estimated. Next, the FD pilot aided estimator was derived. This estimator operates in the frequency domain and estimates the CFO based on the FFT outputs of the received signal at pilot carrier positions. This algorithm outperforms all of the other pilot aided CFO estimators that are discussed in this chapter. Nevertheless, the MSE shows an error floor for high SNRs. Finally a decision aided approach is derived to further improve the quality of the CFO estimate. Every iteration the decision aided algorithm estimates the residual CFO based on the FFT outputs of both the pilot carriers and the data carriers. The error floor wich is present for the MSE of the pilot aided CFO estimators is completely removed after one iteration for the considered range of SNR values.

The derivation of both the TD pilot aided CFO estimator (8.22) and the FD pilot aided CFO estimator (8.31), and a study of their performance have been published in [76]. The FD decision aided CFO estimator is derived and studied in [77].

8.A MSE of the all pilots FD estimator

In this appendix we derive the MSE of the CFO estimator from (8.50) in the case that only pilot symbols are transmitted. The estimate is given by

$$\hat{\epsilon}_p = \frac{1}{2\pi}\frac{N}{N+\nu}\frac{\angle\left(R\left(P\right)\right)}{P}. \tag{8.52}$$

where $R\left(P\right)$ is defined in (8.45). First we rewrite $R\left(P\right)$ as

$$R\left(P\right) = \left(K-P\right)e^{j2\pi\frac{P(N+\nu)}{N}\epsilon_p}\mathbf{x}^H\mathbf{x}\Bigg(1+\frac{1}{\left(K-P\right)\mathbf{x}^H\mathbf{x}}\sum_{i=P}^{K-1}\mathbf{x}^H\tilde{\mathbf{v}}_i e^{-j2\pi\frac{i(N+\nu)}{N}\epsilon_p} + \frac{1}{\left(K-P\right)\mathbf{x}^H\mathbf{x}}\sum_{i=0}^{K-P-1}\mathbf{x}^T\tilde{\mathbf{v}}_i^* e^{j2\pi\frac{i(N+\nu)}{N}\epsilon_p}\Bigg), \tag{8.53}$$

where we have neglected the second order noise terms.

For a complex number x with $|x| \ll 1$, the function $\angle\left(1+x\right)$ is approximated by $\Im\left\{x\right\}$. We can use this approximation in (8.50) for sufficiently high E_s/N_0, which yields for $\hat{\epsilon}_p$

$$\hat{\epsilon}_p = \epsilon_p + \frac{1}{\left(K-P\right)P}\frac{1}{2\pi}\frac{N}{N+\nu}\frac{1}{\mathbf{x}^H\mathbf{x}}\Im\Bigg\{\sum_{i=P}^{K-1}\mathbf{x}^H\tilde{\mathbf{v}}_i e^{-j2\pi\frac{i(N+\nu)}{N}\epsilon_p} + \sum_{i=0}^{K-P-1}\mathbf{x}^T\tilde{\mathbf{v}}_i^* e^{j2\pi\frac{i(N+\nu)}{N}\epsilon_p}\Bigg\}. \tag{8.54}$$

The MSE is given by

$$E\left[\left|\hat{\epsilon}_p-\epsilon_p\right|^2\right] = \frac{1}{\left(K-P\right)^2 P}\frac{1}{\left(2\pi\right)^2}\left(\frac{N}{N+\nu}\right)^2\frac{1}{\left(\mathbf{x}^H\mathbf{x}\right)^2}\mathbf{x}^H\mathbf{R}_{\tilde{\mathbf{v}}}\mathbf{x}, \tag{8.55}$$

for $P \leq K/2$ and by

$$E\left[\left|\hat{\epsilon}_p-\epsilon_p\right|^2\right] = \frac{1}{\left(K-P\right)P^2}\frac{1}{\left(2\pi\right)^2}\left(\frac{N}{N+\nu}\right)^2\frac{1}{\left(\mathbf{x}^H\mathbf{x}\right)^2}\mathbf{x}^H\mathbf{R}_{\tilde{\mathbf{v}}}\mathbf{x}, \tag{8.56}$$

for $P \geq K/2$. The matrix $\mathbf{R}_{\tilde{\mathbf{v}}}$ is the autocorrelation matrix of the samples of the noise vector $\tilde{\mathbf{v}}_i$: $\mathbf{R}_{\tilde{\mathbf{v}}} = E\left[\tilde{\mathbf{v}}_i\tilde{\mathbf{v}}_i^H\right]$. The MSE is symmetric with respect to $P = K/2$: the values $P = Q$ and $P = K-Q$ result in the same MSE, and the minimum value of the MSE is reached for the values of P which are closest to $K/3$ and $2K/3$.

9

Conclusions and Ideas for Future Work

In this chapter we give a summary of the most important achievements of this work. Furthermore some directions for future research are briefly introduced.

9.1 Conclusions

In this work we have studied an OFDM system. Between every two transmitted OFDM systems a guard interval is inserted to avoid IBI between adjacent blocks. The most popular guard interval techniques are the CP and the ZP technique. In the CP technique, the last part of the OFDM block is copied and transmitted before the actual OFDM block, while in the ZP technique the guard interval is left empty. Here we have considered a third guard interval technique called KSP. The guard interval is filled with a known pilot sequence. Contrary to the CP and the ZP technique, we have full control about the content of the guard interval. As a result the guard interval can be optimized and exploited for the synchronization and channel estimation purposes. The main

focus of this dissertation is the derivation of synchronization and channel estimation algorithms that exploit both the pilot symbols in the time domain and the pilot symbols that are transmitted on the pilot carriers.

After the introductory chapters in which we have derived a system model (chapter 2) and have reviewed some general techniques for detection and estimation problems (chapter 3), we have discussed data detection for a perfectly synchronized KSP-OFDM system in chapter 4. After the removal of the guard interval contributions, the system under study can be considered as a ZP-OFDM system. Two main categories of detection algorithms exist: i) the ones that operate in the time domain and ii) the ones that operate in the frequency domain. The optimal time domain detection algorithm, i.e. MLSD, exhibits a very high computational complexity. Luckily low complexity alternatives exist: in the time domain we have the ZF detector and in the frequency domain we have the simple symbol-by-symbol detector. The time domain detectors yield the better performance, especially for high SNR because they can exploit the multipath diversity. On the other hand, the FD detector exhibits the lowest computational complexity. For a coded transmission, the ZF detector and the FD detector are good alternatives for the computationally prohibitive optimal detector. There is no big difference in performance between the ZF detector and the FD detector as is the case for uncoded transmission.

In the remaining chapters (5-8), the estimation of several system parameters is carried out. Chapter 5 serves as an introductory chapter. A first major issue in a communication system is the estimation of the channel impulse response. In chapter 6 we investigate several channel estimation algorithms. First a lower bound on the achievable performance, i.e. the GCRB, is derived together with a low and high SNR limit. We have shown that the MSE's of the existing channel estimators which we discuss in this work, suffer from an error floor for moderate to high SNR because they neglect the presence of the unknown data symbols. For the all pilots case, both algorithms are equal and do not show an error floor. This TD all pilots estimator serves as a benchmark for the performance of iterative decision aided algorithms. We have proposed a DA estimator that operates on the FFT outputs of the received signal at the pilot carrier positions. The FFT outputs at the pilot carrier positions only depend on the pilot symbols (both the transmitted time domain as the frequency domain pilot symbols) and as a result, the MSE of our FD DA estimator does not shown any error floor. A minor drawback of this algorithm is the fact the contribution of the guard interval pilot symbols is spread over all of the carriers, so the parts of their contribution which appear at the FFT outputs corresponding to the data carrier positions can not be exploited for channel estimation. As a result for low SNR, the performance of the FD DA estimator is slightly worse compared to the existing estimators. For the all pilots case, the proposed FD DA estimator yields a second all pilots estimator, i.e. the FD all pilots estimator. The channel estimate obtained by only exploiting the pilot

symbols is further improved by applying decision aided channel estimation algorithms. Here we propose three different decision aided algorithms. The first one is the TD hard decision aided estimation algorithm. As the name suggests, it considers the detected data symbols which are provided by the detector /decoder as extra pilot symbols and uses them to improve the channel estimate. The algorithm operates on the time domain received signal. A second algorithm which is closely related to the TD hard decision based estimator, is the EM based channel estimator. Instead of using hard decisions of the data symbols, the algorithm uses soft information about the data symbols. Both algorithms reach the performance of the TD all pilots estimator for sufficiently high SNR values. The third proposed decision aided algorithm is the FD hard decision aided algorithm. It can be seen as an extension of the FD DA estimator: after the detection / decoding of the data symbols, they are considered as extra pilot symbols and their hard decisions are used to estimate the channel by applying the FD DA estimator to all of the carriers. The FD hard decision aided estimator reaches the performance of the FD all pilots estimator provided that the SNR is high enough.

Chapter 7 deals with the estimation of the time delay offset. First of all the consequences of a timing error are discussed. It is shown that there is a range of timing errors which can be tolerated (depending on the length of the guard interval): the receiver can deal with them by incorporating them into the channel estimation without any loss in performance caused by IBI. Then some existing algorithm for TDS-OFDM are discussed. They assume transmission over a flat fading channel, which leads to inaccurate estimates in a multipath fading environment. We have developed two pilot aided time domain estimation algorithms which exploit both the pilot symbols in the guard interval and the pilots transmitted on the pilot carriers. Both of them are derived under the assumption that the SNR is low. A first algorithm considers the first two terms of the Taylor series expansion of the joint likelihood function of the time delay offset and the channel impulse response and averages over the unknown data symbols. The second algorithm ignores the presence of the unknown data symbols. In both algorithms the estimate of the channel impulse response is expressed as a function of the time delay offset and as a result the joint estimation problem reduces to a one-dimensional search for the time delay offset (and an analytical computation of the channel impulse response estimate). Both pilot aided TD algorithms show a good performance even for a very short guard interval. A drawback of both of them is their high computational complexity. We have tried to solve this by only considering the pilot symbols transmitted in the guard interval for the time delay offset estimation. Despite the reduction of the computational complexity, its performance is not that great, especially for short guard intervals. A last time delay offset estimation algorithm operates in the frequency domain. The time delay offset is estimated based on the FFT outputs of the received signal at the pilot

carrier positions. The algorithm considers both the pilot symbols on the pilot carriers and the pilot symbols transmitted in the time domain. For the correct value of the time delay offset, there is no interference from the unknown data symbols. As a result, the algorithm exhibits an excellent performance independent of the length of the guard interval. The computational complexity of this algorithm is somewhere between the computational complexity of the guard interval based estimator and the pilot aided TD estimators.

In chapter 8 we take care of the CFO estimation. A CFO that has not been compensated, attenuates the useful signal component and causes interference between the different carriers. Both effects might result in an unacceptable increase of the BER. First some existing algorithms for TDS-OFDM are reviewed. Two of them assume a flat fading channel, which deteriorates their performance in a multipath fading channel. The third existing algorithm, i.e. the guard interval based estimation algorithm exhibits a good performance and can also be applied in a KSP-OFDM system. Besides the guard interval pilot symbols, also the pilots transmitted on the pilot carriers can be exploited for the CFO estimation. A first possibility is to extend the guard interval based estimation algorithm so that it also considers the time domain contributions of the pilot symbols on the pilot carriers. However the performance of this TD pilot aided estimator is worse than the performance of the guard interval based estimator. This is caused by the interference of the unknown data symbols. The main advantage of this approach is its robustness against a time delay offset due to the periodic nature of the TD transmitted pilot signal (guard interval pilot symbols + pilot carriers). So when the receiver needs to estimate both the CFO and the time delay offset, which is normally the case, the TD pilot aided estimator can provide an estimate of the CFO before the time delay offset is estimated. Next, we have introduced the FD pilot aided estimator. This estimation algorithm operates in the frequency domain and estimates the CFO based on the FFT outputs of the received signal at the pilot carrier positions. Of all the DA CFO estimators that are discussed in this work, the FD pilot aided estimator yields the best performance. However, its MSE still shows an error floor for high SNR. This error floor is caused by the interference from adjacent data carriers. To improve the estimate of the CFO, a decision aided algorithm, which operates in the frequency domain, is derived. Every iteration the decision aided algorithm estimates the residual CFO based on the FFT outputs of both the pilot and the data carriers. As a result, the error floor which is present for the MSE of the pilot aided CFO estimators is removed after one iteration for the considered range of SNR values.

Table 9.1 gives an overview of the different algorithms that we have proposed in this work to estimate the channel impulse response, the TDO and the CFO respectively . The table also indicates whether an algorithm operates on the received signal in the time domain or on the received signal in the frequency domain. For the pilot aided algorithms we have seen that the ones

Parameter	Time Domain	Frequency Domain
Channel	HD Aided (sec. 6.8.1)	Pilot Aided (sec. 6.6)
	EM based (sec. 6.9)	HD Aided (sec. 6.8.2)
TDO	Pilot Aided (sec. 7.4)	Pilot Aided (sec. 7.6)
	GI based (sec. 7.5)	
CFO	Pilot Aided (sec. 8.3)	Pilot Aided (sec. 8.4)
		HD Aided (sec. 8.5)

Table 9.1: Overview of the proposed algorithms.

that operate in the FD yield the better performance. The main advantage of FD pilot aided algorithms is that they do not suffer from interference from the unknown data symbols (or at least the interference is a lot weaker), although they do not fully exploit the TD pilot symbols because some of their contributions appear at the FFT outputs corresponding to the data carrier positions.

9.2 Future Work

In this section we give some possibilities to further improve the proposed estimation algorithms. Furthermore we provide some future research ideas for other systems which are related to our considered KSP-OFDM system.

9.2.1 Joint Frequency and Timing Estimation

For the derivation of the time delay offset estimation algorithms in chapter 7, we have assumed that there was no CFO. In chapter 8, we have assumed that the time delay offset was perfectly known by the receiver for the derivation of the CFO estimation algorithms. In practice, the receiver has to deal with both effects (time delay offset and CFO). The proposed time delay offset estimation algorithms are not robust against a CFO. So when there is a CFO, their performance will be degraded. Same goes for the FD pilot aided CFO estimator (8.31) and the guard interval based CFO estimator (8.18). So the joint estimation of both the time delay offset and the CFO is an important topic that needs to be looked into further. One possible solution might be given by the TD pilot aided CFO estimator (8.22). We have shown that this estimator is robust against any time delay offset but has limited performance. The estimate provided by this estimator can be used to compensate the CFO. Then one of our time delay offset estimators can be applied to estimate the time delay offset and then in a next step another (fine) estimation of the CFO can be performed by applying the guard interval based CFO estimator (8.18)

or the FD pilot aided CFO estimator (8.31).

9.2.2 Selection of the Pilot Carrier Positions

In chapter 5 we have introduced two different pilot schemes, i.e. the comb-type pilot scheme with equally spaced pilot cariers and the random pilot scheme. It has been shown in the literature ([78]) that equally spacing pilot carriers result in the best performance in terms of the MSE on the estimate of the channel impulse response when the number of carriers N is a multiple of the number of pilot carriers $M - \nu$. Figure 6.7 illustrates this. However, when N is not a multiple of $M - \nu$ the equally spaced pilot carriers can result in a very large degradation. See for example figure 6.7, for 593 pilot carriers ($M = 600$), the MSE is much larger than the theoretical lower bound. The results also indicate that when the number of pilot carriers is large, the random pilot scheme might give better results, but this pilot scheme is not very practical.

In [78], the authors give a hint of how to choose the pilot carriers when N is not a multiple of $M - \nu$. Instead of only considering one pilot spacing, as is the case when the pilot carriers are equally spaced, we now have to consider two pilot carrier spacings: one pilot carrier spacing is equal to $\lfloor N/(M-\nu) \rfloor$ and the other one is equal to $\lfloor N/(M-\nu) \rfloor + 1$. We then have to uniformly interleave these two pilot spacings. We will illustrate this with an example: suppose we have $N = 32$ carriers and we want to select $M - \nu = 5$ pilot carriers. Figure 9.1 shows the pilot carrier positions for both the equal spacing and the quasi-uniform spacing of the pilot carriers from [78]. When the pilot carriers are equally spaced, the distance between every two pilot carriers is equal to 6 except for the distance between the last and the first pilot carrier. This distance is equal to 8. When we apply the quasi-uniform spacing, we see that the distance between the pilot carriers alternates between 6 and 7. The distance between the last and the first pilot carrier is equal to 6. The pilot carriers are more evenly spread over all of the carriers. In [79], these results have been confirmed and are further extended. It turns out that the different pilot carrier positions for the quasi-uniform spacing are given by

$$\alpha_m = \text{round}\left(\frac{mN}{M-\nu}\right),\ m = 0,\ \ldots,\ M-\nu-1, \tag{9.1}$$

where round (x) rounds x to the nearest integer.

This new pilot carrier scheme can also be applied in a KSP-OFDM system. To give a first impression of the improvement in performance that can be achieved with quasi-uniform spaced pilot carriers, we review the BER results for the guard interval based TDO estimator (7.36) from chapter 7. In figure 7.13 we have seen that the BER results for the guard interval based TDO estimator were rather poor. These results were obtained with equally spaced

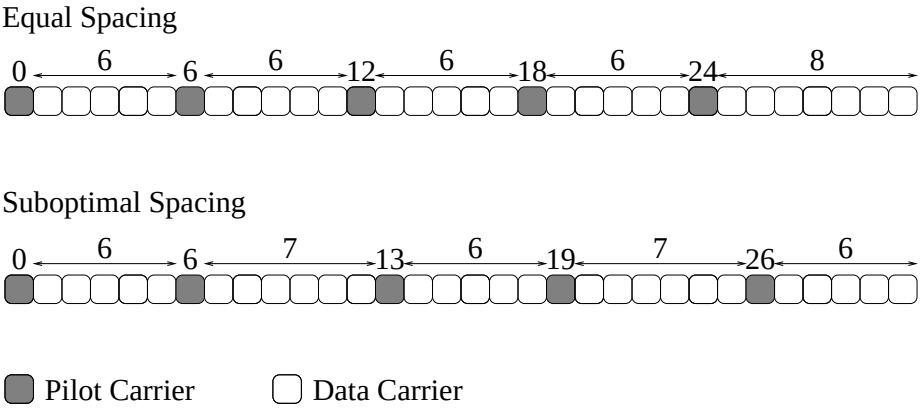

Figure 9.1: Illustration of the equal spacing and the suboptimal spacing of the pilot carriers for $N = 32$ and $M - \nu = 5$.

pilot carriers. When we apply the quasi-uniform spacing of the pilot carriers, we see in figure 9.2 that there is a large improvement especially for high values of E_b/N_0.

9.2.3 Comparison with CP-OFDM

Now that we have derived some estimation algorithms for different parameters, a logical next step is to make an extensive comparison with CP-OFDM. Note that the proposed algorithms that operate in the FD can easily be adopted so that they also can be applied in a CP-OFDM system. It is expected that estimation algorithms which only consider the TD pilot symbols in a KSP-OFDM system will outperform algorithms for CP-OFDM which only exploit the presence of the CP for the estimation process. However for algorithms that operate in the FD the conclusion might be entirely different because we have seen that in a KSP-OFDM the transformation of the received signal to the FD results in an increase of the noise variance, which is not the case in a CP-OFDM system. In [23] it has already been shown that CP-OFDM results in a (slightly) higher SNR of the received signal in the FD in the case of perfect channel knowledge and perfect synchronization.

9.2.4 Parametric Channel Estimation

In chapter 6, we have introduced some channel estimation algorithms which perform well. The decision aided algorithms reach the performance of the all pilots estimators. So in terms of the MSE, the achieved performance is excellent. However if we look at the BER (see for example figures 6.16 and 6.18), there is a gap between the performance of a receiver with perfect channel

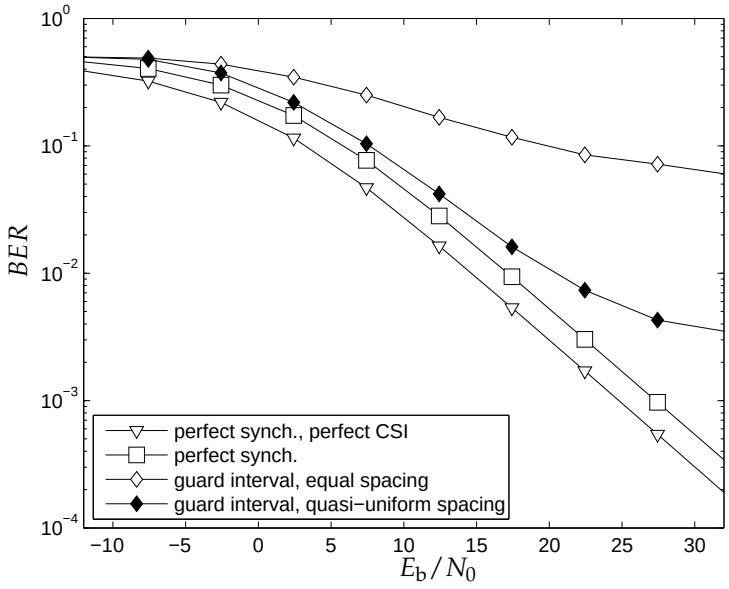

Figure 9.2: BER as a function of E_b/N_0, $N = 1024$, $\nu = 100$, $L = 50$, $M = 100$.

knowledge and a receiver which applies one of the decision aided channel estimation algorithms.

The proposed channel estimation algorithms belong to the class of the non-parametric channel estimators. They estimate the samples of the channel impulse response $h(t)$ (2.25). They do not make any assumptions about the underlying channel model, while we have shown in chapter 2 that $h(t)$ is the cascade of the transmit pulse $p(t)$, the channel impulse response $\tilde{h}(t)$ and the impulse response of the matched filter $p^*(t)$. The transmit pulse $p(t)$ and the impulse response of the matched filter $p^*(t)$ are both known by the receiver. So if the receiver estimates $\tilde{h}(t)$, estimates of the samples of $h(t)$ can be reconstructed. Each path of the channel model $\tilde{h}(t)$ is defined by its parameters, i.e. the complex amplitude $\tilde{h}_l$ and the path delay τ_l. Channel estimation algorithms which provide estimates of the complex path gains $\tilde{h}_l$ and the path delays τ_l, $l = 0, \ldots, \tilde{L} - 1$, are called parametric channel estimation algorithms. Usually, the number of paths $\tilde{L}$ is much smaller than the number of samples in the channel impulse response L. We have shown (see [80,81]) that parametric channel estimation results in a better performance than non parametric channel estimation. So to improve the quality of the estimate of the channel impulse response $\mathbf{h}$, we want to develop parametric channel estimators. For CP-OFDM, there exist already some algorithms in the literature [82, 83] that consider parametric channel estimation. They can not be directly applied in

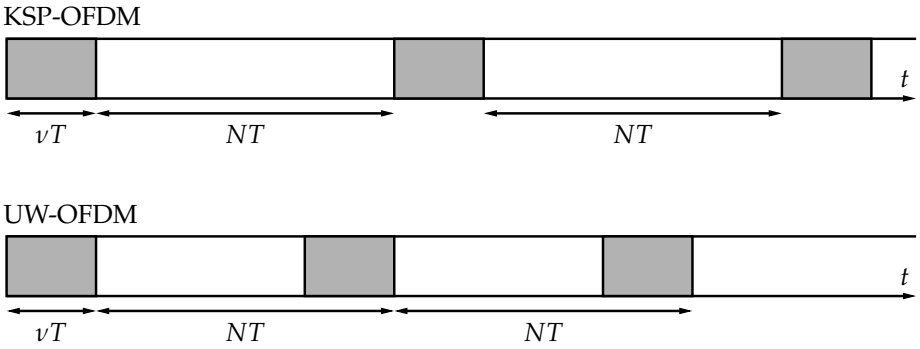

Figure 9.3: The structure of the transmitted signal for KSP-OFDM and UW-OFDM respectively.

a KSP-OFDM system due to the presence of the contributions of the guard interval pilot symbols on the FFT outputs (see 6.59).

9.2.5 Unique Word OFDM

Unique word (UW) OFDM is a novel way of constructing OFDM symbols [84, 85]. In conventional OFDM systems adjacent OFDM blocks are separated by a guard interval (see also sections 1.3 and 1.4). The time domain pilot sequence is part of the FFT interval in UW-OFDM, whereas in KSP-OFDM this is not the case. Figure 9.3 compares the transmit data structures of KSP-OFDM and UW-OFDM. Looking at the transmit structure of UW-OFDM, we can see that cyclicity is also ensured as in CP-OFDM.

To generate a UW-OFDM block there exist two approaches [86], i.e. a direct approach and a two-step approach. Both approaches make use of redundant carriers. In order to obtain the desired pilot sequence at the last ν positions of the IFFT operation, at least ν carriers must be reserved and cannot be used for data transmission. Appropriate values have to be transmitted on these redundant carriers to yield the UW at the output. The direct approach generates the unique word directly at the output of the IFFT. The two-step approach tries to generate an IFFT output of which the last ν samples are equal to zero. In a second step the unique word is added in the time domain. It turns out that the two-step approach needs to transmit much less energy than the direct approach [86], so the two-step approached is preferred. A critical aspect of UW-OFDM is the choice of the redundant carriers. A suboptimal placement may result in very high values of the peak-to-average-power-ratio (PAPR) [84]. A high PAPR is undesirable because it causes poor power efficiency.

The redundant carriers can be used at the receiver side to improve the detection of the data symbols [85]. Especially when the channel frequency response shows deep fades at carrier positions, a UW-OFDM systems exhibits

a better performance in an uncoded system than a CP-OFDM system.

In the literature, there are some publications ([84, 85]) which deal with the data detection in a UW-OFDM. However, they consider a perfectly synchronized receiver which knows the channel impulse response perfectly. To our knowledge there exist no publications which deal with channel estimation nor the estimation of synchronization parameters, so an extension of the work presented in this dissertation, would be to investigate if the here proposed estimation algorithms can be applied (possibly with some adjustments) in a UW-OFDM system.

Publications

The research presented in this dissertation resulted in 1 refereed international journal publications, 9 conference publications, and 1 submitted paper

Journal Publications

- IEEE Transactions on Signal Processing: [34]
- IEEE Transactions on Communications (submitted): [77]

Conference Publications

- Refereed international conference publications: [33,35,58,59,71,72,76,80, 81]

Apart from these, we have published 7 additional works (2 journal papers, 1 book chapter and 4 conference papers) which are related to research that falls outside the scope of this work: [87–93].

Bibliography

[1] H. Meyr, M. Moeneclaey, S.A. Fechtel, *Synchronization, Channel Estimation, and Signal Processing*, ser. Digital Communication Receivers. John Wiley & Sons, 1997, vol. 2.

[2] J.G. Proakis, *Digital Communications*, 4th ed. McGraw-Hill, 2001.

[3] M.K. Simon, M.-S. Alouini , *Digital Communication over Fading Channels*, 2nd ed., ser. Wiley Series in Telecommunications and Signal Processing. Wiley-IEEE Press, 2005.

[4] H. Steendam, "The Effect of Synchronization Errors on Multicarrier Systems," Ph.D. dissertation, Ghent University, Belgium, 2000.

[5] D. Tse, P. Viswanath , *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[6] J. A. C. Bingham, "Multicarrier modulation for data transmission, an idea whose time has come," *IEEE Communications Magazine*, vol. 28, no. 5, pp. 5–14, May 1990.

[7] S. B. Weinstein, "The history of orthogonal frequency-division multiplexing," *IEEE Communications Magazine*, vol. 47, no. 11, pp. 26–35, Nov. 2009.

[8] S. Weinstein and P. Ebert, "Data transmission by frequency-division multiplexing using the discrete Fourier transform," *IEEE Transactions on Communication Technology*, vol. 19, no. 5, pp. 628 –634, 1971.

[9] P. Kyees, R. McConnell, and K. Sistanizadeh, "ADSL: a new twisted-pair access to the information highway," *IEEE Communications Magazine*, vol. 33, no. 4, pp. 52–60, Apr. 1995.

[10] K. Maxwell, "Asymmetric digital subscribers line: interim technology for the next years," *IEEE Communications Magazine*, vol. 34, no. 10, pp. 100–106, Oct. 1996.

[11] J. Cioffi, V. Oksman, J.-J. Werner, T. Pollet, P. Spruyt, J. Chow, and K. Jacobsen, "Very-high-speed digital subscriber lines," *IEEE Communications Magazine*, vol. 37, no. 4, pp. 72–79, Apr. 1999.

[12] European Telecommunications Standards Institute (ETSI), "Digital audio broadcasting (DAB); DAB to mobile portable and fixed receivers," France, ETS 300 401, Feb. 1995.

[13] ——, "Digital video broadcasting (DVB); Framing structure, channel coding and modulation for 11/12 GHz satellite services," France, ETS 300 421, Dec. 1994.

[14] U. Reimers, "Digital video broadcasting," *IEEE Communications Magazine*, vol. 36, no. 6, pp. 104–110, June 1998.

[15] H. Sari, G. Karam, and I. Jeanclaude, "Transmission techniques for digital terrestrial TV broadcasting," *IEEE Communications Magazine*, vol. 33, no. 2, pp. 100–109, Feb. 1995.

[16] J. Song, Z. Yang, L. Yang, K. Gong, C. Pan, J. Wang, and Y. Wu, "Technical review on Chinese digital terrestrial television broadcasting standard and measurements on some working modes," *IEEE Transactions on Broadcasting*, vol. 53, no. 1, pp. 1–7, Mar. 2007.

[17] IEEE 802.11, "Wireless LAN medium access control (MAC) physical layer (PHY) specifications, amendment 1: high-speed physical layer in the 5 GHz band," July 1999.

[18] IEEE 802.16, "IEEE standard for local and metropolitan area networks part 16: air interface for fixed broadband wireless access systems," May 2009.

[19] S. Galli and O. Logvinov, "Recent developments in the standardization of power line communications within the IEEE," *IEEE Communications Magazine*, vol. 46, no. 7, pp. 64 –71, July 2008.

[20] V. Oksman and S. Galli, "G.hn: The new ITU-T home networking standard," *IEEE Communications Magazine*, vol. 47, no. 10, pp. 138 –145, Oct. 2009.

[21] D. Astely, E. Dahlman, A. Furuskar, Y. Jading, M. Lindstrom, and S. Parkvall, "LTE: the evolution of mobile broadband," *IEEE Communications Magazine*, vol. 47, no. 4, pp. 44–51, 2009.

[22] A. Ghosh, R. Ratasuk, B. Mondal, N. Mangalvedhe, and T. Thomas, "LTE-Advanced: next-generation wireless broadband technology [Invited Paper]," *IEEE Wireless Communications Magazine*, vol. 17, no. 3, pp. 10–22, 2010.

[23] H. Steendam and M. Moeneclaey, "Different guard interval techniques for OFDM: performance comparison," in *Proc. from 6th International Workshop on Multi-Carrier Spread Spectrum (MC-SS'07)*, Herrsching, Germany, May 2007, pp. 11–24.

[24] B. Muquet, Z. Wang, G.B. Giannakis, M. de Courville and P. Duhamel, "Cyclic prefixing or zero padding for wireless multicarrier transmissions," *IEEE Transactions on Communications*, vol. 50, no. 12, pp. 2136–2148, Dec. 2002.

[25] Z. Wang and G. Giannakis, "Wireless multicarrier communications: where Fourier meets Shannon," *IEEE Signal Processing Magazine*, vol. 17, no. 3, pp. 29–48, May 2000.

[26] L. Deneire, B. Gyselinckx, and M. Engels, "Training sequence versus cyclic prefix - a new look on single carrier communication," *IEEE Communications Letters*, vol. 5, no. 7, pp. 292–294, July 2001.

[27] O. Rousseaux, G. Leus and M. Moonen, "Estimation and equalization of doubly selective channels using known symbol padding," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 979–990, Mar. 2006.

[28] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: turbo-codes," *IEEE Transactions on Communications*, vol. 44, no. 10, pp. 1261 –1271, Oct. 1996.

[29] R. Gallager, "Low-density parity-check codes," *IRE Transactions on Information Theory*, vol. 8, no. 1, pp. 21 –28, Jan. 1962.

[30] S. Lin, D.J. Costello , *Error Control Coding*, 2nd ed. Prentice-Hall, 2004.

[31] N. Noels, "Synchronization in Digital Communication Systems: Performance Bounds and Practical Algorithms," Ph.D. dissertation, Ghent University, Belgium, 2009.

[32] H. Steendam, M. Moeneclaey and H. Bruneel, "The Cramer-Rao bound and ML estimate for data-aided channel estimation in KSP-OFDM," in *Proceedings of the 18th Personal, Indoor and Mobile Radio Communications Symposium, 2007. PIMRC-07*, Athens, Greece, Sept. 2007.

[33] D. Van Welden, H. Steendam and M. Moeneclaey, "Frequency-domain data-aided channel estimation for KSP-OFDM," in *Proc. from 10th International Symposium on Spread Spectrum Techniques and Applications (ISSSTA'08)*, Bologna, Italy, Aug. 2008.

[34] D. Van Welden and H. Steendam, "Near optimal iterative channel estimation for KSP-OFDM," *IEEE Transactions on Signal Processing*, vol. 58, no. 9, pp. 4948–4954, Sept. 2010.

[35] D. Van Welden, H. Steendam and M. Moeneclaey, "Time delay estimation for KSP-OFDM systems in multipath fading channels," in *Proceedings of the 20th Personal, Indoor and Mobile Radio Communications Symposium, 2009. PIMRC-09*, Sept. 2009.

[36] H. Minn, V. K. Bhargava and B. Letaief, "A robust timing and frequency synchronization for OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 2, no. 4, pp. 822–839, July 2003.

[37] H.L. Van Trees, *Detection, Estimation, and Modulation Theory, part I*. Wiley and Sons, Oct. 2001.

[38] A.P. Dempster, N.M. Laird and D.B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society. Series B*, vol. 39, no. 1, pp. 1–38, 1977.

[39] F. Simoens, "Iterative multiple-input multiple-output communication systems," Ph.D. dissertation, Ghent University, Belgium, 2008.

[40] T. Cover and J. Thomas, *Elements of Information Theory*. New York, NY, USA: John Wiley & Sons.

[41] W. E. Ryan, S. Lin, *Channel Codes: Classical and Modern*. Cambridge University Press, 2009.

[42] H. Wymeersch, *Iterative Receiver Design*. New York, NY, USA: Cambridge University Press, 2007.

[43] S. Coleri, M. Ergen, A. Puri and A. Bahai, "Channel estimation techniques based on pilot arrangement in OFDM systems," *IEEE Transactions on Broadcasting*, vol. 48, no. 3, pp. 223–229, Sept. 2002.

[44] F. Tufvesson, T. Maseng, "Pilot assisted channel estimation for OFDM in mobile cellular systems," in *Proc. of IEEE Veh. Tech. Conf.*, Phoenix, U.S.A., May 1997, pp. 1639–1643.

[45] H. Wymeersch, "Software Radio Algorithms for Coded Transmission," Ph.D. dissertation, Ghent University, Belgium, 2005.

[46] C. Herzet, "Code-Aided Synchronization for Digital Burst Communications," Ph.D. dissertation, Université catholique de Louvain, Belgium, 2006.

[47] N. Noels, V. Lottici, A. Dejonghe, H. Steendam, M. Moeneclaey, M. Luise, and L. Vandendorpe, "A theoretical framework for soft-information-based synchronization in iterative (turbo) receivers," *EURASIP Journal on Wireless Communications and Networking*, vol. 2005, no. 2, pp. 117 –129, Apr. 2005.

[48] C. Herzet, N. Noels, V. Lottici, H. Wymeersch, M. Luise, M. Moeneclaey, and L. Vandendorpe, "Code-aided turbo synchronization," *Proceedings of the IEEE*, vol. 95, no. 6, pp. 1255 –1271, June 2007.

[49] E. de Carvalho and D.T.M. Slock, "Maximum-likelihood blind FIR multi-channel estimation with Gaussian prior for the symbols," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 1997. ICASSP-97.*, vol. 5, Apr. 1997, pp. 3593 – 3596.

[50] O. Rousseaux, G. Leus, P. Stoica and M. Moonen, "Gaussian maximum-likelihood channel estimation with short training sequences," *IEEE Transactions on Wireless Communications*, vol. 4, no. 6, pp. 2945 – 2955, Nov. 2005.

[51] R. Cendrillon and M. Moonen, "Efficient equalizers for single and multi-carrier environments with known symbol padding," in *Proceedings of the International Symposium on Signal Processing and its Applications (ISSPA)*, Aug. 2001, pp. 607–610.

[52] H. Steendam, M. Moeneclaey and H. Bruneel, "An ML-based estimate and the Cramer-Rao bound for data-aided channel estimation in KSP-OFDM," *EURASIP Journal on Wireless Communications and Networking, special issue on Multicarrier Systems*, 2008, article ID 186809, 9 pages, 2008. doi:10.1155/2008/186809.

[53] S. Tang, F. Yang, K. Peng, C. Pan, K. Gong, and Z. Yang, "Iterative channel estimation for block transmission with known symbol padding - a new look at TDS-OFDM," in *Global Telecommunications Conference, 2007. GLOBECOM '07. IEEE*, Nov. 2007, pp. 4269 –4273.

[54] F. Yang, J. Wang, J. Wang, J. Song, and Z. Yang, "Novel channel estimation method based on PN sequence reconstruction for Chinese DTTB system," *IEEE Transactions on Consumer Electronics*, vol. 54, no. 4, pp. 1583 –1589, Nov. 2008.

[55] B. Liu, L. Gui, W. Zhang, and B. Song, "On channel estimation method using time domain sequences in ofdm systems," *IEEE Transactions on Broadcasting*, vol. 54, no. 4, pp. 786 –791, Dec. 2008.

[56] P. Stoica and P. Babu, "The Gaussian data assumption leads to the largest Cramér-Rao bound [Lecture Notes]," *IEEE Signal Processing Magazine*, vol. 28, no. 3, pp. 132–133, May 2011.

[57] X. Ma, H. Kobayashi and S.C. Schwartz, "EM-based channel estimation algorithms for OFDM," *Eurasip Journal on Applied Signal Processing*, vol. 2004, no. 10, pp. 1460–1477, 2004.

[58] D. Van Welden, H. Steendam and M. Moeneclaey, "Iterative DA/DD channel estimation for KSP-OFDM," in *Proc. IEEE International Conference on Communications (ICC '08)*, Beijing, China, May 2008, pp. 693–697.

[59] D. Van Welden and H. Steendam, "Iterative EM based channel estimation for KSP-OFDM," in *Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008. IEEE 19th International Symposium on*, Sept. 2008, pp. 1–5.

[60] R. Ravikanth and S.P. Meyn, "Bounds on achievable performance in the identification and adaptive control of time-varying systems," *IEEE Transactions on Automatic Control*, vol. 44, no. 4, pp. 670–682, Apr. 1999.

[61] T. Keller, L. Piazzo, P. Mandarini and L. Hanzo, "Orthogonal frequency division multiplex synchronization techniques for frequency-selective fading channels," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 6, pp. 999–1008, June 2001.

[62] Y. Mostofi and D. C. Cox, "Mathematical analysis of the impact of timing synchronization errors on the performance of an OFDM system," *IEEE Transactions on Communications*, vol. 54, no. 2, pp. 226–230, Feb. 2006.

[63] D. Landström, S.K. Wilson, J.-J. van de Beek; P. Ödling and P.O. Börjesson, "Symbol time offset estimation in coherent OFDM systems," *IEEE Transactions on Communications*, vol. 50, no. 4, pp. 545–549, Apr. 2002.

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

[65] T. Schmidl and D.C. Cox, "Robust frequency and timing synchronization for OFDM," *IEEE Transactions on Communications*, vol. 45, no. 12, pp. 1613–1621, Dec. 1997.

[66] H. Minn, M. Zeng, and V. Bhargava, "On timing offset estimation for ofdm systems," *IEEE Communications Letters*, vol. 4, no. 7, pp. 242–244, July 2000.

[67] J.-J. van de Beek, P. O. Börjesson, M.-L. Boucheret, D. Landström, J. M. Arenas, P. Ödling, S. K. Wilson, "Three non-pilot based time- and frequency estimators for OFDM," *Elsevier Signal Processing*, vol. 80, pp. pp. 1321–1334, 2000.

[68] B. Yang, K. B. Letaief, R. S. Cheng and Z. Cao, "Timing recovery for OFDM transmission," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 11, pp. 2278–2291, Nov. 2000.

[69] L. He, F. Yang, C. Zhang, and Z. Wang, "Synchronization for TDS-OFDM over multipath fading channels," *IEEE Transactions on Consumer Electronics*, vol. 56, no. 4, pp. 2141–2147, Nov. 2010.

[70] G. Liu and S. Zhidkov, "A composite PN-correlation based synchronizer for TDS-OFDM receiver," *IEEE Transactions on Broadcasting*, vol. 56, no. 1, pp. 77–85, Mar. 2010.

[71] D. Van Welden, F. Simoens, H. Steendam, and M. Moeneclaey, "Pilot based time delay estimation for KSP-OFDM systems," in *NEWCOM++ - ACoRN Joint Workshop, Proceedings*, 2009.

[72] ——, "Pilot based time delay estimation for KSP-OFDM systems in a multipath fading environment," in *Proceedings of the 14th International OFDM-workshop*, Sept. 2009, pp. 93–97.

[73] T. Pollet, M. Van Bladel, and M. Moeneclaey, "BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise," *IEEE Transactions on Communications*, vol. 43, no. 234, pp. 191–193, feb/mar/apr 1995.

[74] P. Moose, "A technique for orthogonal frequency division multiplexing frequency offset correction," *IEEE Transactions on Communications*, vol. 42, no. 10, pp. 2908 –2914, Oct. 1994.

[75] J. Zheng and W. Zhu, "An algorithm for calibration of TDS-OFDM carrier frequency offset," *IEEE Transactions on Consumer Electronics*, vol. 55, no. 2, pp. 366–370, May 2009.

[76] D. Van Welden, H. Steendam, and M. Moeneclaey, "Frequency offset estimation for KSP-OFDM," in *Proceedings of the 12th IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), 2011*, June 2011, pp. 271–275.

[77] ——, "Iterative decision-directed joint frequency offset and channel estimation for KSP-OFDM," *submitted to IEEE Transactions on Communications*, 2011.

[78] X. Cai and G. Giannakis, "Error probability minimizing pilots for ofdm with m-psk modulation over rayleigh-fading channels," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 1, pp. 146 – 155, Jan. 2004.

[79] H. Steendam, "The quasi-redundant carrier placement for uw-ofdm," *submitted to VTC fall 2012, Quebec, Canada*.

[80] D. Van Welden, M. Moeneclaey, and H. Steendam, "Parametric versus nonparametric data-aided channel estimation in a multipath fading environment." in *Proceedings of 2006 Symposium on the Communications and Vehicular Technology. SCVT06*, Nov. 2006, pp. 25–28.

[81] D. Van Welden, F. Simoens, H. Steendam, and M. Moeneclaey, "An MSE lower bound for parametric and nonparametric channel estimation," in *Proceedings of the 18th Personal, Indoor and Mobile Radio Communications Symposium, 2007. PIMRC-07*, Athens, Greece, Sept. 2007, pp. 1 –5.

[82] M. Oziewicz, "On application of MUSIC algorithm to time delay estimation in OFDM channels," *IEEE Transactions on Broadcasting*, vol. 51, no. 2, pp. 249 – 255, June 2005.

[83] B. Yang, K. Letaief, R. Cheng, and Z. Cao, "Channel estimation for OFDM transmission in multipath fading channels based on parametric channel modeling," *IEEE Transactions on Communications*, vol. 49, no. 3, pp. 467 –479, Mar. 2001.

[84] M. Huemer, C. Hofbauer, and J. B. Huber, "The potential of unique words in OFDM," in *Proceedings of the 15th International OFDM-workshop*, Sept. 2010, pp. 140–144.

[85] M. Huemer, A. Onic, and C. Hofbauer, "Classical and bayesian linear data estimators for unique word OFDM," *IEEE Transactions on Signal Processing*, vol. 59, no. 12, pp. 6073 –6085, Dec. 2011.

[86] A. Onic and M. Huemer, "Direct vs. two-step approach for unique word generation in UW-OFDM," in *Proceedings of the 15th International OFDM-workshop*, Sept. 2010, pp. 145–149.

[87] D. Van Welden, F. Simoens, H. Steendam, and M. Moeneclaey, "Impact of channel estimation error on the performance of linear FIR equalizers for frequency selective MIMO channels," in *Proceedings of 2005 Symposium on the Communications and Vehicular Technology. SCVT05*, Nov. 2005.

[88] D. Van Welden and H. Steendam, "Interference cancellation of AM narrowband interference signals," in *Proceedings of the 65th IEEE Vehicular Technology Conference, 2007. VTC2007-Spring.*, Apr. 2007, pp. 1617–1621.

[89] F. Simoens, D. Van Welden, H. Wymeersch, and M. Moeneclaey, "Low complexity MIMO detection based on the slowest descent method," *IEEE Communications Letters*, vol. 11, no. 5, pp. 429–431, May 2007.

[90] D. Van Welden and H. Steendam, "Narrowband AM interference cancellation for broadband multicarrier systems," *EURASIP Jurnal On Wireless Communications and Networking*, vol. 2008, 2008.

[91] D. Van Welden, F. Simoens, H. Steendam, and M. Moeneclaey, "A comparison between parametric and nonparametric channel estimation for multipath fading channels," in *4G Mobile and Wireless Communications Technologies*, S. Kyriazakos, I. Soldatos, and G. Karetsos, Eds. River Publishers, 2008, pp. 169–178.

[92] D. Van Welden and H. Steendam, "Clipping versus symbol switching for PAPR reduction in coded OFDM," in *Proceedings of 2008 Symposium on the Communications and Vehicular Technology. SCVT08*, Nov. 2008.

[93] ——, "PAPR reduction by symbol nulling," in *Proceedings of the 69th IEEE Vehicular Technology Conference, 2009. VTC Spring 2009.*, Apr. 2009, pp. 1320–1324.