N° **41841**

UNIVERSITÉ DE LILLE 1 SCIENCES ET TECHNOLOGIES

École Doctorale Sciences Pour l'Ingénieur

THÈSE

pour obtenir le grade de
Docteur de l'Université Lille 1: sciences et technologies

Spécialité: Micro et Nanotechnologies, Acoustique et Télécommunications
préparée à l'Institut d'Electronique de Microélectronique et de Nanotechnologie

présentée et soutenue publiquement le 27 novembre 2015 par:

Farah ZAAROUR

Channel estimation algorithms for OFDM in interference scenarios

Membres du jury

Rapporteurs:	
Pr. **Daniel ROVIRAS**	CNAM Paris
Pr. **Christophe LAOT**	Telecom Bretagne
Examinateur:	
Dr. **Mohammad-Ali KHALIGHI**	Ecole Centrale Marseille
Pr. **Jérôme LOUVEAUX**	Université catholique de Louvain
Directeurs de thèse:	
Pr. **Martine LIENARD**	Université de Lille
Pr. **Iyad DAYOUB**	Université de Valenciennes
Encadrants:	
Dr. **Eric SIMON**	Université de Lille
Dr. **Marie COLIN-ZWINGELSTEIN**	Université de Valenciennes

To my family ... ♡

Acknowledgments

This thesis, as any journey in life, has been full with moments of doubt, anxiety and uncertainty. It has also been full with moments of success, achievement and confidence. During both times, I was lucky to have the support of a lot of people whom I would like to acknowledge. This is to let them know that this work could not have been achieved without them.

First, it is thanks to God, Allah, for the courage, patience and guidance He gave me to begin, persist and finalize this journey.

I am particularly grateful to professor Daniel ROVIRAS and professor Christophe LAOT for accepting to be the reviewers of my manuscript and for the quality of their reviews on my work. I am also grateful to professor Jérôme LOUVEAUX and doctor Mohammad-Ali KHALIGHI for their commitment to be part of my jury committee.

I would like to thank everyone in TELICE, for their warm welcome, friendliness and sense of humor. Thanks to professor Martine LIENARD for hosting me in the lab.

My sincere appreciation goes to doctor Eric SIMON. I would like to thank him not only for his availability and for the particular effort and support he provided me during my thesis, but also for his patience, kindness and for his confidence in me during the most difficult times. Thanks to our discussions, I have learned much more than science during my stay.

I would like to thank Emmanuelle GILLMANN for her unconditional help from my first days in France.

I would also like to thank professor Iyad DAYOUB and doctor Marie COLIN-ZWINGELSTEIN from the university of Valenciennes for the time they devoted to me and for their constructive comments on my work, but most of all for their conviviality.

My gratitude goes to my colleagues and friends in Lebanon and France for their support and for all the times we have spent together. Thanks to Hawraa, Hiba, Dina, Nora, Faten, Walaa, Hassanein, Mortada, Raouf, Shiqi, Huaqiang Shu, Kyoko, Khaled and to all those they know I consider them as friends.

Finally, I would like to thank my family in Lebanon and USA for their support and their participation, each one in his special way to make this possible.

To my father, mother and sisters: Dunia, Leila and Dana, this is for you...

Résumé

La rareté du spectre radio et la demande croissante de bande passante rendent l'optimisation de l'utilisation du spectre essentiel. Tandis qu'une efficacité maximale devrait être atteinte, un niveau minimal d'interférence devrait être maintenu. L'OFDM (*Orthogonal Frequency Division Multiplexing en anglais*) est un schéma de modulation bien connu pour combattre efficacement l'évanouissement multi trajets. OFDM a été retenu comme un schéma de modulation dans plusieurs normes, comme le 3GPP LTE (*Long Term Evolution en anglais*) et un dérivé d'OFDM, le GFDM (*Generalized Frequency Division Multiplexing en anglais*), est un candidat pour le système 5G. L'estimation de canal est une tâche fondamentale dans les systèmes OFDM et elle devient plus difficile en présence d'interférence. Dans cette thèse, notre objectif est de proposer des algorithmes d'estimation de canal pour les systèmes OFDM en présence d'interférence, où les algorithmes classiques d'estimation de canal échouent. En particulier, nous nous concentrons sur les interférences dues à certaines approches utilisées pour optimiser le spectre. Cela nous amène à considérer les deux cas suivants.

1) Tout d'abord, nous considérons l'environnement radio intelligente CR (*Cognitive Radio en anglais*) qui a été proposé pour faire face au problème de rareté du spectre. Les technologies qui emploient l'OFDM comme schéma de modulation (WIMAX, WRAN) peuvent exister dans un scenario de CR. Dans de tels scenarios, un type particulier d'interférence se pose et il est connu comme l'interférence à bande étroite ou NBI (*Narrowband Interference en anglais*). NBI est caractérisée par une puissance élevée qui frappe un petit nombre de sous porteuse dans un système OFDM. Pendant que tous les travaux dans la littérature traitent le cas du canal à variations lentes, nous proposons un nouveau cadre d'estimation de canal pour les canaux à variations rapides con-

taminés par NBI. Cela est accompli avec l'algorithme EM (*Expectation maximization en anglais*) et une expression explicite pour l'estimation de la puissance du bruit est obtenue.

2) Une autre source de consommation de bande passante en OFDM est la présence de pilotes connus insérés dans la trame OFDM pour accomplir l'estimation du canal. Pour tenter de remédier à ce problème, les pilotes superposés SP (*Superimposed Pilots en anglais*) ont été proposés comme substituts aux pilotes classiques. Dans cette thèse, on est intéressé par une nouvelle classe de SP pour OFDM connu comme DNSP (*Data Nulling SP en anglais*). DNSP assure des pilotes sans interférence au détriment d'interférence des donnés. Due à la modernité de DNSP, un récepteur adapté à son design doit être conçu. Les récepteurs turbo sont connus pour être les plus efficaces en supprimant l'interférence. Néanmoins, ils possèdent deux inconvénients majeurs; leur complexité et leur besoin d'un canal précis. La contribution de ce travail est double et elle est orientée vers le traitement des deux inconvénients mentionnés.

Nous proposons d'abord un annuleur d'interférences IC (*Interference canceler en anglais*) basé sur le critère MMSE (*Minimum Mean Square Error en anglais*) à faible complexité pour DNSP. En fait, nous montrons qu'en exploitant la structure spécifique de l'interférence en DNSP, l'inversion de matrice nécessaire dans le cas classique se réduit à l'inversion d'une matrice diagonale. Cependant, la performance de l'IC proposé n'est fiable que quand l'erreur de l'estimation du canal est faible. Donc, dans une autre contribution, nous proposons un IC pour DNSP en tenant compte des erreurs d'estimation du canal.

Enfin l'estimation *robuste* du canal est abordée dans le dernier chapitre.

Abstract

The scarcity of the radio spectrum and the increasing demand on bandwidth makes it vital to optimize the spectrum use. While a maximum efficiency should be attained, a minimal interference level should be maintained. Orthogonal frequency division multiplexing (OFDM) is a well-known modulation scheme reputed to combat multipath fading efficiently. OFDM has been selected as the modulation scheme in several standards such as the 3GPP long term evolution (LTE) and a derivative of OFDM, the generalized frequency division multiplexing (GFDM), is a candidate for 5G systems. In order to guarantee a coherent detection at the receiver, in the absence of the channel knowledge, channel estimation is regarded as a fundamental task in OFDM. It becomes even more challenging in the presence of interference.

In this thesis, our aim is to propose channel estimation algorithms for OFDM systems in the presence of interference, where conventional channel estimators designed for OFDM fail. In particular, we focus on interference that arises as a result of schemes envisioned to optimize the spectrum use. This leads us to consider two main themes; cognitive radio (CR) and superimposed pilots (SP).

First, we consider the CR environment which has been proposed to tackle the problem of spectrum scarcity. Technologies employing OFDM as their modulation scheme such as WIMAX and WRAN, might exist in a CR network. In such scenarios, a particular type of interference arises and is known as the narrowband interference (NBI). NBI is characterized by a high power which strikes a small number of OFDM sub-carriers. While all existing literature addresses slow time-varying channels, we propose a novel channel estimation framework for fast time-varying channels in OFDM with NBI. This is accomplished through an expectation maximization (EM) based algorithm. This formulation allows us to obtain a closed-form expression for the noise power

estimation.

The known pilot sub-carriers inserted within the OFDM frame to accomplish the channel estimation task are another source of bandwidth consumption in OFDM. In an attempt to overcome this drawback, SP have been proposed as substitutes to conventional pilots. In this thesis, we are particularly interested in a very recent class of SP for OFDM, known as the data-nulling SP (DNSP) scheme. DNSP assures interference-free pilots at the expense of data interference. Seen the modernity of DNSP, a suitable receiver has to be designed to cope with its design. Turbo receivers are well known to be the most efficient in canceling interference. However, two main drawbacks of turbo receivers are their high complexity and their need of an accurate channel estimate. The contribution of this work is twofold and is oriented to deal with the two mentioned downsides of turbo receivers.

We first propose a low-complexity soft approximated minimum mean square error (MMSE) interference canceler (IC) particularly for DNSP. In fact, we show that by exploiting the specific interference structure that arises in DNSP, the matrix inversion needed by the approximated MMSE-IC in the classical case reduces to a diagonal matrix inversion. The performance of the proposed IC is reliable when the channel estimation error is small. As another contribution, we extend the design of the approximated IC for DNSP so as to take the channel estimation errors into account. Those two contributions allow us to benefit from the interference cancellation property of turbo receivers, while keeping a low-complexity and dealing with channel estimation errors when present.

Finally, *robust* channel estimation is discussed in the last chapter. In particular, we provide some insights and propositions for its implementation to problems of channel estimation in OFDM.

Contents

List of Figures

List of Tables

List of variables, notations and acronyms

List of variables

N	Number of sub-carriers in an OFDM system
N_g	Cyclic prefix length
$N_T = N + N_g$	Total number of samples in an OFDM block
T_s	Sampling time in an OFDM frame
$T = N_T T_s$	Total duration of an OFDM symbol
K	Total number of OFDM symbols per frame
$\mathbf{x}_k$	kth transmitted OFDM symbol
$\mathbf{y}_k$	kth received OFDM symbol
$\mathbf{H}_k$	Channel matrix associated with the kth OFDM symbol
$\mathbf{n}_k$	Additive white Gaussian noise on the kth OFDM symbol
σ^2	Noise variance on sub-carrier n

σ^2_{TN}	Variance of thermal noise
$\sigma^2_{I,n}$	Variance of unknown interference
N_c	Number of BEM coefficients
$\sigma^2_{\alpha_l}$	Variance of the lth channel tap
$\alpha_{l,k,q}$	l^{th} channel tap sampled at time $kT_t + (q + N_g)T_s$
L	number of channel paths
$\mathbf{B}$	Basis function matrix
$\mathbf{c}_{l,k}$	BEM coefficients
$\boldsymbol{\xi}_{l,k}$	BEM modeling error of the lth path on the kth OFDM symbol
$\mathbf{R}^{(p)}_{\mathbf{c}_l}$	Correlation matrix of BEM coefficients
$\mathbf{A}$ and $\mathbf{U}_l$	AR parameters
$\hat{\mathbf{c}}^{(i)}_{k\|k-1}$	BEM coefficients' prediction
$\mathbf{P}^{(i)}_{k\|k-1}$	Prediction error covariance
$\mathbf{K}_k$	Kalman Gain
$\hat{\mathbf{c}}^{(i)}_{k\|k}$	BEM coefficients' update
$\mathbf{P}^{(i)}_{k\|k}$	Measurement error covariance

$\boldsymbol{\mathcal{A}}$	State Matrix
$\mathbf{u}_k$	State Noise
$f_d T_t$	Normalized Doppler frequency
N_c	Number of BEM coefficients
$\mathbf{h}_k$	Channel impulse response on kth OFDM symbol
$\hat{\mathbf{h}}_k$	Estimate of the channel impulse response on kth OFDM symbol
D	Number of data sub-carriers
P	Number of pilot sub-carriers
L_f	Distance between two adjacent pilots in frequency
$\mathbf{W}$	Hadamard matrix
ρ	Pilot power allocation
$\mathbf{y}_{\mathbf{d},k}$	kth received OFDM symbol on data sub-carriers
$\mathbf{y}_{\mathbf{p},k}$	kth received OFDM symbol on pilot sub-carriers
$\mathbf{n}_{\mathbf{p},k}$	Noise on pilot sub-carriers of the kth OFDM symbol
$\mathbf{n}_{\mathbf{d},k}$	Noise on data sub-carriers of the kth OFDM symbol
$\mathbf{w}_{k,n}$	Channel equalizer for the kth OFDM symbol on the nth sub-carrier
$\tilde{s}_{k,n}$	Equalized symbol for the kth OFDM symbol on the nth sub-carrier
$\overline{\mathbf{s}}'_{k,n}$	Expected value of the data symbols given the LLR

$\mathbf{e}_n$	$N \times 1$ vector with one on its nth entry and zeros elsewhere
$\mathbf{V}_{k,n}$	$N \times N$ diagonal matrix with the variance of the soft estimates on its entries
$\mathbf{W_d}$	$D \times N$ sub-matrix of $\mathbf{W}$
σ_s^2	Variance of data symbols
$\sigma_{\bar{s}}^2$	Variance of estimated data symbols given the LLR
$\mathbf{F}$	$N \times L$ Fourier matrix
$\mathbf{F_p}$	$P \times L$ Fourier matrix
d	Distance between transmitter and receiver
v	Pathloss exponent
τ_m	Delay spread
B_c	Coherence bandwidth
f_d	Doppler frequency
T_c	Coherence time
B	Signal Bandwidth
τ_s	Signal Symbol period
$r(t)$	Continuous received signal in base band
$s(t)$	Continuous transmitted signal in base band
τ_l	Propagation delay associated with the lth path

$h(t, \tau)$	Impulse response of the physical channel
$H(t, f)$	Frequency response of the physical channel at frequency f and time t
$R_{\alpha l}(\Delta t)$	Autocorrelation function of the lth complex gain
$S_{\alpha l}(\Delta f)$	Doppler spectrum of the lth complex gain
$\alpha_l(t)$	Complex gain associated with the lth path
ν_l	Gain of $\alpha_l(t)$
$\psi_l(t)$	Phase of $\alpha_l(t)$
$\mathbf{G}_k$	Global channel matrix in DNSP

List of notations

$\mathbf{X}$	Matrix $\mathbf{X}$ (bold capital letter)
$\mathbf{x}$	Vector $\mathbf{x}$ (bold small letter)
$[\mathbf{X}]_{n,m}$	Element corresponding to the nth row and mth column of matrix $\mathbf{X}$
$(\cdot)^T$	Transpose operator
$(\cdot)^*$	Conjugate operator
$(\cdot)^H$	Hermitian (conjugate transpose) operator
$Tr(\cdot)$	Trace
$E_{x,y}[\cdot]$	Expectation over x and y
$\delta_{k,m}$	Dirac function
$\arg\max_{(\cdot)}$	Maximum argument
$\text{diag}(\mathbf{X})$	Vector with diagonal entries of matrix $\mathbf{X}$
$\text{diag}(\mathbf{x})$	Diagonal matrix with the entries of vector $\mathbf{x}$ on its diagonal
FT	Fourier transform operator
$\circ$	Hadamard multiplication

List of acronyms (in alphabetical order)

ADSL	Asynchronous digital subscriber line
AR	Auto-regressive
AWGN	Additive white Gaussian noise
BCRB	Bayesian Cramer–Rao Bound
BDU	bounded data uncertainty
BEM	Basis Expansion Model
BER	Bit Error Rate
BICM	Bit-Interleaved Coded Modulation
CA	Complex Addition
C-CSP	coded classical superimposed pilots
CE	Channel Estimation
CFO	Carrier frequency offset
CIR	Channel impulse response
CM	Complex Multiplication
CMMOE	Constrained Minimum Mean Output Energy
CP	Cyclic Prefix
CR	Cognitive Radio
CS	Compressive sensing
CSI	Channel State Information

CSP	Classical Superimposed Pilots
D/A (A/D)	Digital to analog (Analog to digital)
DA	Data-Aided
DFT	Discrete Fourier Transform
DNSP	Data Nulling Superimposed Pilots
ECM	Expectation-Conditional Maximization
EM	Expectation Maximization
FCC	Federal Communications Commission
FFT	Fast Fourier Transform
FT	Fast Transform
GFDM	Generalized Frequency Division Multiplexing
IC	Interference Canceler
ICI	Inter-carrier interference
IDFT	Inverse DFT
IFFT	Inverse Fast Fourier Transform
ISI	Inter-symbol interference
ISM	industrial, scientific and medical
LLR	Log likelihood ratio
LMMSE	Linear Minimum Mean Square Error

LOS	Line-of-sight
LS	Least Squares
LTE	Long Term Evolution
MAP	Maximum a posteriori
ML	Maximum Likelihood
MLE	Maximum Likelihood Estimator
MMSE	Minimum Mean Square Error
MSE	Mean Square Error
MUE	Measurement Update Equations
NBI	Narrowband Interference
NLOS	No Line-of-sight
NRNSC	Non-Recursive Non-Systematic Convolutional
OFDM	Orthogonal Frequency Division Multiplexing
P/S (S/P)	Parallel to serial (Serial to parallel)
P-BEM	Polynomial BEM
pdf	probability density function
PEF	Prediction Error Filter
PSAM	Pilot-Symbol Assisted Modulation
PSIC	Parallel Soft Interference Canceler
RA	Real Addition

RLS	Regularized LS
RM	Real Addition
SIR	Signal-to-interference-ratio
SNR	Signal-to-noise-ratio
SOS	Second Order Statistics
SP	Superimposed Pilots
TUE	Time Update Equations
UFMC	Universal Filtered Multi-Carrier
US	Uncorrelated scatters
WLAN	Wireless Local Area Network
WRAN	Wireless Regional Area Network
WSS	Wide-Sense Stationary

Introduction

The management of the spectrum is attributed to regulation authorities which have the responsibility of organizing the frequency bands usage, so that users are guaranteed an acceptable quality of service. In particular, licensed users are not supposed to encounter interference. However, measurements revealed that less than 40 % of the attributed spectrum is being fully used. This leaves an important portion of the spectrum reserved but underutilized. It reveals at the same time the inefficiency of the current regulation policies which offer licensed users the full access to the spectrum at all times. While bandwidth is a naturally scarce resource, and demands on higher data rates is in continuous increase, it becomes vital to make use of the *wasted* spectrum. Thus, *optimizing* the spectral efficiency becomes a priority in order to confront the scarcity of the spectrum on the one hand, and the demand of the constantly emerging services on the other hand. We will hereafter discuss two approaches which will form the road-map to this thesis.

Cognitive radio (CR), first proposed in [Mit93], is based on the idea of spectrum sharing. Its main idea is to allow secondary users to access the licensed spectrum when vacant. This sharing should be well organized (regulated) so that the interference from secondary users on primary users is kept at its minimal levels. This is usually assured by employing spectrum sensing capabilities in the CR networks so as to detect the availability of the spectrum before transmitting, thus avoiding interference.

While CR offers a solution to the spectrum underutilization caused by current regulation policies, there exists another source of bandwidth underutilization which is manifested by the use of known symbols (pilots) for the channel estimation task. This means that a non-negligible portion of the bandwidth is *wasted* rather than being assigned to useful symbols. Nevertheless, in order to coherently acquire the channel at the receiver, some known symbols should

be transmitted. To this end, superimposed pilots (SP) have been proposed to replace classical pilots. In SP schemes, no bandwidth is exclusively dedicated to pilots, but it is rather shared with unknown symbols (data) thus limiting the loss in spectral efficiency.

Orthogonal frequency division multiplexing (OFDM) could be considered a landmark with strong foundations in the physical layer. It has been proposed to overcome the undesirable effect of inter-carrier interference and inter-symbol interference and is perfectly adopted to mulipath channels; a typical scenario in wireless transmission. Furthermore, OFDM is now standardized in several technologies, ranging from wired to wireless communication and extending to television broadcasting transmissions. Consequently, and with reference to the above discussion, systems employing OFDM as their access scheme might possibly share the spectrum with other technologies. The channel estimation task is no exception in OFDM systems, and is typically carried out via pilot-aided algorithms as discussed above.

We now narrow the rest of our discussion to OFDM, which is the modulation of interest in our thesis. In particular, our main concern in this thesis is to provide solutions to channel estimation problems in OFDM in scenarios with interference, where conventional channel estimation algorithms fail. We focus on interference that arises either from spectrum sharing (CR) between users, or from bandwidth sharing (SP) by the same user.

As it could be anticipated, optimizing the spectral efficiency, being it with the employment of CR or with SP does not come without a cost. In fact, different sources of interference should be dealt with at the receiver.

We first consider the CR environment with a fast time-varying OFDM channel and conventional pilots. In CR, the spectrum sensing capabilities we already mentioned can not guarantee an interference-free transmission. In fact, a particular kind of interference arises in OFDM when coexisting with other technologies. It is characterized by a high power over a small number of subcarriers. This is known as the narrowband interference (NBI) and its presence renders conventional channel estimators inefficient. This will be the focus of our work in chapter 2.

Then, we consider a quasi-static OFDM channel with SP and an NBI-free environment. The sharing of the bandwidth in SP schemes could be performed using different approaches. In the classical SP (CSP) scheme, pilots are arithmetically added to data with a different power contribution. This leads to interference on channel estimation from data and on data decoding from pi-

lots. A more recent SP scheme, known as the data-nulling SP (DNSP) scheme, cancels the effect of data on pilots by first precoding the data and then nulling it as some positions for the insertion of pilots. This assures that there is no interference on pilots, yet, interference originates from data on themselves. This will be the focus of our work in chapters 3 and 4.

Thesis Organization:

In **chapter 1**, we provide a brief overview of the multipath propagation channel, the OFDM model in a fast time-varying scenario and a slow time-varying scenario and then briefly discuss conventional channel estimation techniques in OFDM.

In **chapter 2**, the fast time-varing OFDM channel model will be implemented and then adapted to include the NBI model. Extensive work on channel estimation with NBI in OFDM systems exists in literature, but only in the context of low mobility. Thus, we present a novel framework to address the problem of channel estimation in OFDM with NBI in high mobility conditions. In fact, the extension of this problem from low mobility scenarios to high mobility scenarios is not a direct one. Another motivation is the recent proposal dealing with the employment of CR in high speed railway in the CORRIDOR project [Cor]. This project is the first to address this topic in Europe in collaboration with the Société Nationale des Chemins de fer Français (SNCF), which is the French National Railway Company. To tackle this problem, we derive an EM-based channel estimator, in which the Kalman smoother is embedded. We then perform a comparison between the proposed algorithm and two algorithms in the literature. In particular, the comparison is done with an algorithm that takes into account NBI but neglects mobility and another which takes into account mobility but neglects NBI. This comparison demonstrates the need of a channel estimator capable of dealing with NBI and mobility at the same time. We also calculate the additional complexity of the proposed algorithm which is shown to be rather small compared to that of existing algorithms.

In **chapter 3**, the quasi-static OFDM channel model is used and adapted this time to include the DNSP model. DNSP is a recent proposal which emerged in 2014. It was thus essential to design a suitable receiver with an acceptable complexity. To this end, we propose a low-complexity turbo receiver for DNSP in OFDM. We then perform a comprehensive comparison with the CSP in OFDM within the turbo framework. This includes channel estimation, data

decoding, and complexity comparisons. In fact, this study is of important practical interest since it enables the designer to choose the pilot scheme depending on the application and according to the desired trade-off between complexity and performance.

In **chapter 4**, we propose an improved low-complexity receiver for DNSP which takes the channel estimation errors into account. In the proposed receiver of chapter 3, the channel estimate is used as if it were the true channel. This assumption has a low impact when small channel estimation errors are present, but it results in performance degradation otherwise.

Chapter 5 is devoted to the discussion of the possibility of implementing *robust* estimators for the problem of channel estimation in OFDM. The robustness here refers to taking into account uncertainties in the model for a class of robust estimators known as the bounded data uncertainty (BDU) class. In fact, interference in the OFDM system such as the NBI in chapter 2 or the SP in chapter 3, could be viewed as uncertainties. In those two contexts, the conventional channel estimators fail due to the presence of those uncertainties and we consequently had to design more elaborate techniques to deal with them at the receiver. Taking into account those certainties into the design of the estimator presents a new approach to deal with those problems. The content of this chapter is to be viewed as a preliminary study which needs further developments and is provided as a perspective for future research.

Publications:

The results of chapter 2 have been published in IEEE Wireless Communications letters. Chapter 3 resulted in two conference publications, DICTAP 2015 and BMSB 2015, and a paper under review in IET communications. The approach in chapter 4 has been submitted to IEEE Transactions on Wireless Communications for review. Those are listed below:

F. Zaarour and E. Simon, "Fast Time-Varying Channel Estimation for OFDM Systems with Narrowband Interference" *IEEE Wireless Communications Letters*, vol.4 no. 4, pp. 389-392, 2015.

F. Zaarour, E. Simon, M. Zwingelstein-Colin, and I. Dayoub, "A low complexity turbo receiver for data nulling superimposed pilots in OFDM," *Fifth International Conference on Digital Information and Communication Technol-*

ogy and its Applications (DICTAP), pp.32,37, April 29 2015-May 1 2015.

F. Zaarour, E. Simon, M. Zwingelstein-Colin, and I. Dayoub, "Comparison of Superimposed Pilot Schemes in Iterative Receivers for OFDM Systems" *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, June 2015.

F. Zaarour, E. Simon, M. Zwingelstein-Colin, and I. Dayoub, "On a Reduced Complexity Iterative Receiver for Data Nulling Superimposed Pilots in OFDM", *under review in IET communications.*

F. Zaarour, E. Simon, M. Zwingelstein-Colin, and I. Dayoub, "An improved low Complexity Iterative Receiver for Data Nulling Superimposed Pilots in OFDM", *submitted to IEEE Transactions on Wireless Communications.*

Chapter 1

Introduction to OFDM

Contents

1.1 Introduction

Orthogonal frequency division multiplexing (OFDM) is reputed for its ability to mitigate multipath fading and enhance the spectral efficiency due to its orthogonal sub-carriers. This paved the way towards its presence in several standards today, ranging from wireless technologies such as the long term evolution (LTE) in its downlink, to wired technologies such as the asynchronous digital subscriber line (ADSL). However, the route towards standardization took

around 40 years. This dates back to the year 1966 when Cheng [Cha66] first presented the idea of transmitting parallel messages via a linear band-limited channel while maintaining an inter-carrier interference (ICI) free and inter-symbol interference (ISI) free transmission. Nevertheless, the major breakthrough which permitted the practical implementation of OFDM did not occur before the early 70's when Weinstein and Ebert [WE71] introduced the use of the discrete Fourier transform (DFT) as an alternative to oscillators. However, their proposition did not address the orthogonality problem which was solved 10 years later when Peled and Ruiz introduced the idea of the cyclic prefix (CP) [PR80]. Our goal in this chapter is to introduce the notions that will help clearly understand the channel models addressed in later chapters. We first introduce the classical channel propagation model. Then, we present the OFDM system model and distinguish between the fast time-varying and the quasi-static channels. The former will provide a base for the system model for our discussion in chapter 2 while the latter will form the backbone for the system model in chapter 3. This will be followed by a discussion and a brief state of the art of channel estimation techniques in OFDM. This discussion will help us establish the links and understand the motivations for the works presented in later chapters since our ultimate goal is to estimate the channel at the receiver with highest precision and efficiency (spectral) and lowest complexity.

1.2 Propagation Model

1.2.1 Typical propagation model

The basic structure of any communication system is composed of a transmitter, a propagation medium (channel) and a receiver. The propagation medium could be wired or wireless. We will be dealing with the latter in this thesis. In case of a fixed line of sight (LOS) transmission where the receiver has only a single attenuated version of the transmitted signal, then directional antennas can be employed. However, a typical propagation scenario in a wireless medium leads to the reception of multiple copies of the transmitted signal at the receiver with or without a LOS. A propagation scenario with a LOS component follows a Rice distribution while a propagation without a LOS follows a Rayleigh distribution.
The propagation of an electromagnetic wave in a radio channel undergoes diverse physical phenomenon. The three major mechanisms that impact this propagation are the reflection, diffraction and scattering. Reflection occurs when a propagating wave impinges upon a smooth surface with very large dimensions relative to its wavelength. Diffraction is produced when the propa-

gation path between the transmitter and the receiver is obstructed by a dense body with dimensions that are large relative to the signal's wavelength, causing secondary waves to be formed behind the obstructing body. It is often referred to as shadowing because the diffracted field can reach the receiver even when shadowed by an impenetrable obstacle. Finally, scattering is produced when the radio wave falls on a surface whose dimensions are less than of the signal's wavelength, causing the energy to scatter in all directions. Those mechanism

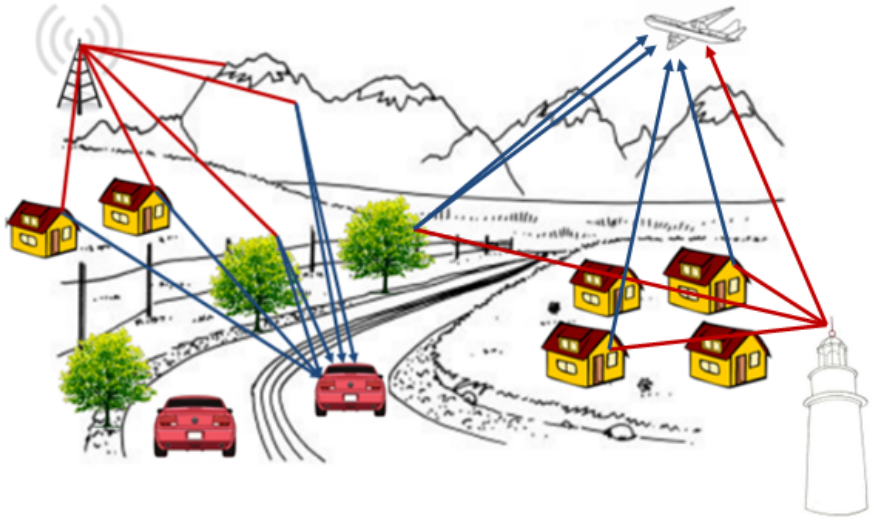


Figure 1.1: Typical Propagation Scenario

lead to the reception of multiple copies of the transmitted signal at the receiver. Thus, the received is made up of the superposition of attenuated and delayed versions of the transmitted signal. A typical propagation scenario is illustrated in Figure 1.1. Those multiple paths characterize what is known as a multipath propagation, and cause fluctuations in the received signal's amplitude, phase and angle of arrival, producing multipath fading. Thus the combination of the primary signal with the delayed copies, leads to a constructive or destructive (fading) interference.

1.2.2 Multipath Fading

There exists two types of fading [B.S00]: large-scale fading and small-scale fading. Large-scale fading represents the average signal power attenuation or

pathloss due to motion over large areas. Hills, forests, and buildings situated between the transmitter and the receiver might initiate this phenomenon. Practically, this fading is characterized by statistics which provides a way of computing an estimate of the path-loss as a function of the distance. Precisely, the pathloss is proportional to $\frac{1}{d^v}$, where d is the distance between the transmitter and the receiver, and v is the pathloss exponent varying from 2 (free-space) to 6 (urban environment). It is normally described in terms of mean path-loss and a log-normally distributed variation about the mean. Small-scale fading refers

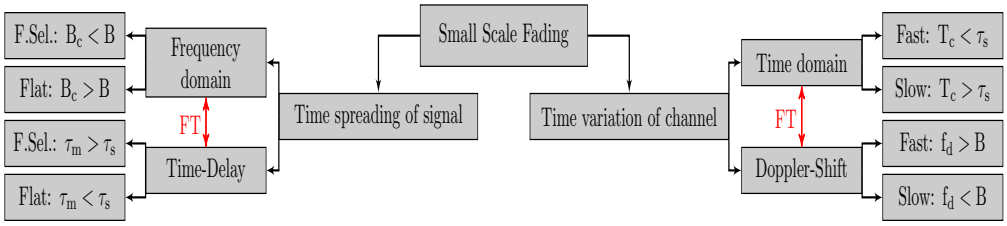


Figure 1.2: Fading channel classification

to the dramatic change in signal amplitude and phase that can be experienced as a result of small changes in the spatial positioning between a receiver and a transmitter. In order to better understand this phenomenon, we introduce the parameters which characterize it. In particular, those are related to the time spreading of the signal and the time variation of the channel. The delay spread τ_m is the difference between the arrival of the first and the last signal. The coherence bandwidth $B_c \approx \frac{1}{\tau_m}$ is the range of frequencies over which the channel is considered flat. Now, the Doppler spread is the spectral broadening caused by the relative motion of the receiver and transmitter, it is equal to the maximum Doppler frequency (f_d). The coherence time ($T_c \approx \frac{1}{f_d}$) is the time over which the channel is considered constant. The variations of those parameters determine the channel classification. In the frequency domain, when $B_c > B$ where B is the signal bandwidth, the channel undergoes flat or frequency non selective fading. Its equivalent in the time-delay domain is observed when the delay spread is much smaller than the signal symbol period τ_s i.e. $\tau_s > \tau_m$. A frequency selective channel is produced when the inverse scenario occurs, either $B > B_c$ in the frequency domain representation or $\tau_m > \tau_s$ in the time-delay representation. Fading can also be slow or fast [1]. A fast fading scenario occurs when the coherence time of the channel is less than the symbol duration

[1] Fast fading is also referred to as time-selective fading.

$T_c < \tau_s$ as seen in the time domain or equivalently when the Doppler spread is greater than the signal bandwidth $f_d > B$ in the frequency domain. Slow fading is produced when $T_c > \tau_s$ or $f_d < B$. Those mechanisms are summarized in Figure 1.2.

1.2.3 Channel Model

The received signal $r(t)$ is composed of the different attenuated and delayed versions of the transmitted signal $s(t)$. It is given in base band as:

$$r(t) = \sum_{l=0}^{L-1} \alpha_l(t) s(t - \tau_l(t)) \tag{1.1}$$

where L is the number of channel paths, $\tau_l(t)$ is the delay of each path with $l = 0, ..., L-1$ and $\alpha_l(t)$ is the amplitude or complex gain associated with the lth path. Thus, the impulse response of the channel as a function of time t and delay τ is given as:

$$h(t, \tau) = \sum_{l=0}^{L-1} \alpha_l(t) \delta(\tau - \tau_l(t)) \tag{1.2}$$

where $\delta(\tau)$ is the Dirac delta function. This formulation allows us to distinguish three kinds of channel relative to the symbol duration τ_s:

- Static: The channel does not vary with time and is referred to as time-invariant
- Quasi-static: The channel is constant within a symbol but varies from one symbol to another
- Time-varying: The channel varies within a symbol as well as from one symbol to another.

Note that in this thesis, we consider that the number of channel paths L and the delays $\tau_l(t)$, $l = 0, ..., L-1$, are constant and known throughout the transmission of a frame of just a few symbols. In fact, the knowledge of those parameters is justified by the possibility to perform an initial estimation at the beginning of the transmission [SRHG12]. In addition, the number of paths could be considered constant due to the fact that the appearance and disappearance of paths are rare since those phenomena depend on the shadowing effect which evolves slowly with respect to the symbol duration. Taking the

previous discussion into account, equation (1.2) could be re-written as:

$$h(t,\tau) = \sum_{l=0}^{L-1} \alpha_l(t)\delta(\tau - \tau_l \cdot T_s). \tag{1.3}$$

Here, τ_l are the delays normalized by the sampling time T_s. It is obvious that in order to resolve the transmitted signal at the receiver, the knowledge of $h(t,\tau)$ is essential. This can be achieved assuming a deterministic or a random channel model.

1.2.3.1 Deterministic Modeling

There exists several ways to this approach. Measurement campaigns aim at measuring the channel impulse response at different points within the measurement perimeter. This, however, does not necessarily provide the channel characteristics such as the delays and paths. Nevertheless, it might be useful for network planning. With the development of the ray tracing technique which is based on the idea of geometric optics, it became feasible to determine the trajectory of the signals. This method can provide the propagation conditions of a user for a given environment. However, it is sensitive to variations in the channel.

1.2.3.2 Random Modeling

The random model of the channel depends directly on the classification presented in figure 1.2. The channel frequency response $H(t,f)$ is directly linked to the impulse response through the Fourier transform (FT) as:

$$\begin{aligned} H(t,f) &= \int_{-\infty}^{+\infty} h(t,\tau)e^{-2j\pi f\tau}d\tau \\ &= \sum_{l=0}^{L-1} \alpha_l(t)e^{-2j\pi f\tau_l T_s} \end{aligned} \tag{1.4}$$

Classically, the impulse response is assumed to be wide-sense stationary (WSS) which indicates that each path $\alpha_l(t)$ is a zero-mean complex Gaussian process; *i.e.* $\mathbb{E}\{\alpha_l(t)\} = 0$, with uncorrelated scatters (US) to indicate that the paths are uncorrelated *i.e.* $\mathbb{E}\{\alpha_{l1}(t)\alpha_{l2}^*(t)\} = 0$. The WSSUS model was first introduced by Bello in [Bel63]. Define $\Delta t = t_1 - t_2$ as the time difference between two realizations of the channel, then the complex amplitudes $\alpha_l(t)$ could be modeled with an autocorrelation function $R_{\alpha_l}(\Delta t)$. Several models could be used to

model the channel, however, we focus here on the Rayleigh distribution since it is the one we will be using throughout this thesis. Thus, in a non-line-of-sight (NLOS) transmission, the signal is supposed to arrive from all different directions at the reception antenna. The channel is seen as the sum of n independent random realizations. This could be expressed as:

$$\alpha_l(t) = \nu_l(t)e^{j\psi_l(t)} = \sum_n \nu_{l,n}e^{j\psi_{l,n}(t)}. \tag{1.5}$$

Applying the central limit theorem, $\alpha_l(t)$ is Gaussian complex with gain $|\alpha_l(t)|= \nu_l$ that follows a Rayleigh distribution given as:

$$p(\nu_l) = \frac{\nu_l^2}{\sigma_{\alpha_l}^2} e^{-\frac{\nu_l^2}{2\sigma_{\alpha_l}^2}}, \tag{1.6}$$

and a uniformly distributed phase $\psi_l(t)$ between 0 and 2π. Its variance, $\sigma_{\alpha_l}^2 = \mathbb{E}\{|\alpha_l(t)|^2\}$. Finally, the Doppler spectrum is obtained from $R_{\alpha l}(\Delta t)$ by the Fourier Transform, *i.e.* $S_{\alpha_l}(f) = \mathrm{FT}(R_{\alpha_l}(\Delta t))$ [Cla68].

1.3 OFDM System Model

1.3.1 Introduction

The basic idea behind OFDM is to transform a frequency selective channel into N parallel sub-channels[2] that undergo flat fading. In fact, one can directly anticipate that the favorable scenario for transmission is to have slow flat fading channel conditions. This is guaranteed when $f_d < B < B_c$. However, this is only possible when $f_d \cdot \tau_m \ll 1$. Then, the symbol duration is chosen such that $\tau_m < \tau_s < T_c$. Thus, in order to achieve high data rates, one would send a large number of narrowband signals over relatively close frequencies. This is the exact rational behind multi-carrier transmission, which consequently facilitates channel equalization at the receiver.

1.3.2 Transmission Chain and mathematical representation

In a classical OFDM system implemented with the DFT, we give the transmission chain in figure 1.3. We will consider frames composed of K OFDM symbols. An OFDM symbol is the ensemble of the N narrowband signals. We consider the $N \times 1$ vector of transmitted symbols on the kth OFDM symbol

[2]The terms sub-channel and sub-carrier are used interchangeably.

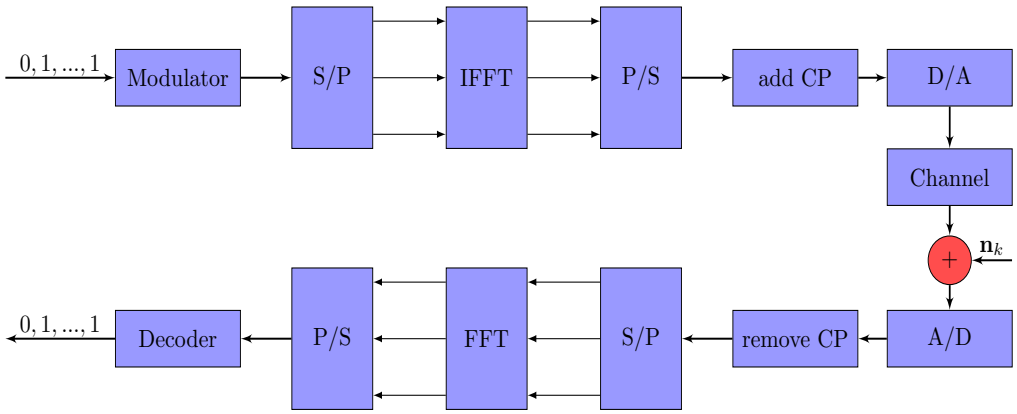


Figure 1.3: OFDM transmission chain

$\mathbf{x}_k = \left[x_{k,0}, ..., x_{k,N-1}\right]^T$ with $\{x_{k,n}\}$ being the power-normalized transmitted symbol on the sub-carrier $n - \frac{N}{2}$. Note that $\mathbf{x}_k$ is obtained after the processing of the binary data which are modulated an then passed through a serial to parallel (S/P) converter. $\mathbf{x}_k$ is then fed to an N-point inverse DFT (IDFT) that transforms the data to the time domain which is again converted from parallel to serial (P/S) . This allows us to insert the CP with length N_g which is typically chosen to be larger than τ_m to avoid ISI. The CP is a repetition of the last part of the symbol at the beginning as depicted in figure 1.4.

Now, the total number of sub-carriers $N_t = N + N_g$ and the total duration of each OFDM symbol is $T_t = N_t T_s$ where T_s is the sampling time. The signal is then passed through a digital to analog (D/A) converter which contains low-pass filters with bandwidth $\frac{1}{T_s}$. Then it is transmitted through the fading channel where an additive white Gaussian noise (AWGN) $\mathbf{n}_k$ of length $N \times 1$ with a covariance matrix defined as $\mathbb{E}\{\mathbf{n}_k \mathbf{n}_k^H\} = \sigma_n^2 \mathbf{I}$ is added. At the receiver, the inverse of the operations performed at the transmitter are done. After the (A/D) converter, the cyclic prefix is removed, the stream is converted (S/P) and an N-DFT operation is performed to obtain the frequency domain received signal $\mathbf{y}_k$ which is given as:

$$\mathbf{y}_k = \mathbf{H}_k \, \mathbf{x}_k + \mathbf{n}_k. \tag{1.7}$$

The binary data is then recovered after channel decoding. The model in (1.7) is valid for any channel variation. However, the $N \times N$ channel matrix is modeled differently. We distinguish between time-varying channels, static and quasi-static channels. The time-varying channel destroys the orthogonality between the sub-carriers and leads to inter-carrier interference (ICI). This is

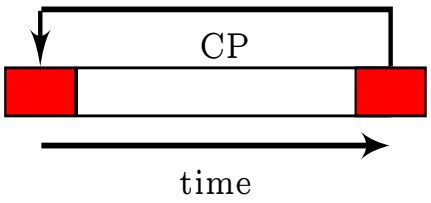


Figure 1.4: Cyclic Prefix insertion in OFDM

manifested in a full channel matrix as in figure 1.5, where the off-diagonal components represent the ICI. This phenomenon becomes harsher for very fast time-varying channels as in figure 1.6. This is usually diagonal in case of a static channel as in 1.7.

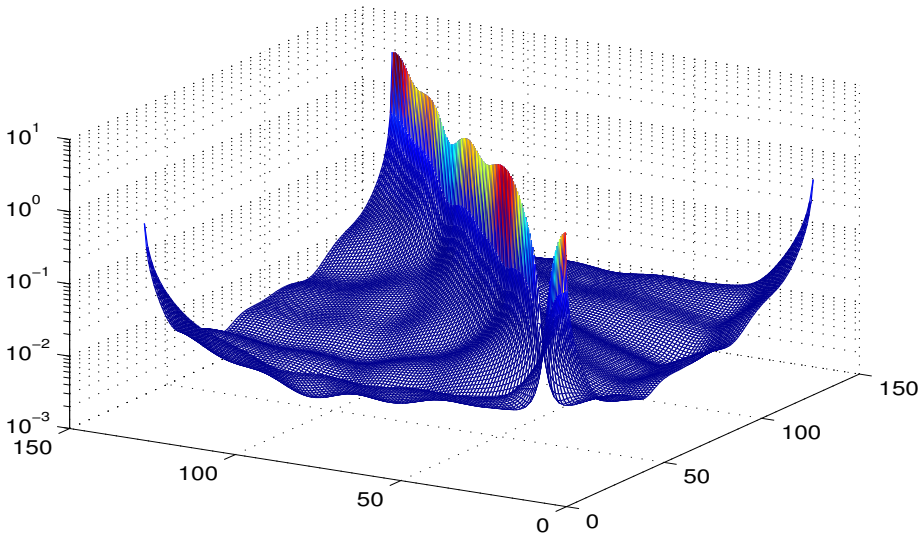


Figure 1.5: Typical fast time-varying channel with $f_d T_t = 0.1$, where the power of channel is concentrated around its diagonal forming a banded diagonal structure

Now, $\mathbf{H}_k$ is given as:

$$[\mathbf{H}_k]_{n,m} = \frac{1}{N} \sum_{l=0}^{L-1} \left[e^{-j2\pi(\frac{m}{N}-\frac{1}{2})\tau_l} \sum_{q=0}^{N-1} \alpha_{l,k,q} e^{j2\pi \frac{m-n}{N} q} \right], \tag{1.8}$$

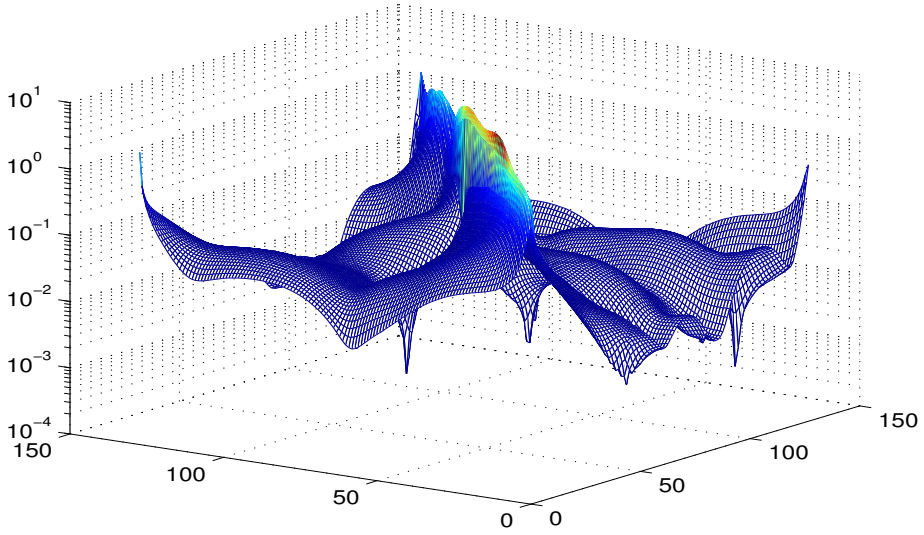


Figure 1.6: Very fast time varying channel for $f_d T_t = 0.5$, where the power of the channel is dispersed all over the sub-carrier indexes

where $\alpha_{l,k,q} = \alpha(kT_t + (q + N_g)T_s)$ is the l^{th} channel tap sampled at time $kT_t + (q + N_g)T_s$. In the case of static and quasi-static channels, $\mathbf{H}_k$ simplifies to:

$$[\mathbf{H}_k]_{n,n} = \sum_{l=0}^{L-1} \left[\alpha_{l,k} e^{-j2\pi(\frac{n}{N}-\frac{1}{2})\tau_l}\right], \tag{1.9}$$

where $\alpha_{l,k} = \alpha(kT_t + (N_g + \frac{N}{2})T_s)$. This formulation is based on the fact that in static and quasi-static channels, the complex amplitudes are constant from one realization of the channel to the other $(\alpha_{l,k})$. This allows us to re-write (1.7) as:

$$\mathbf{y}_k = \text{diag}\{\mathbf{x}_k\}\mathbf{F}\boldsymbol{\alpha}_k + \mathbf{n}_k, \tag{1.10}$$

where $\boldsymbol{\alpha}_k = [\alpha_{0,k}, ..., \alpha_{L-1,k}]$ and $\mathbf{F}$ is the $N \times L$ Fourier matrix with its (n, l)th entry given as $[\mathbf{F}]_{n,l} = e^{-j2\pi(\frac{n}{N}-\frac{1}{2})\tau_l}$.

We finally note that the OFDM channel model presented in this section is in terms of the physical parameters [RHS14a][SK13][SHR$^+$10] (complex gains and delays). This is in order to maintain a continuity with the previous section, so as to clearly link the model to the physical phenomenon. In the rest of this

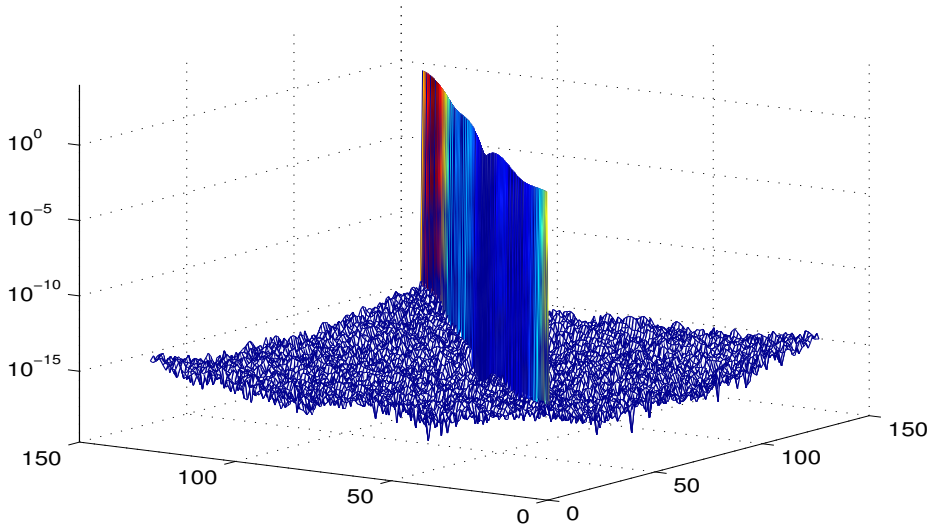


Figure 1.7: Static channel, *i.e.* $f_d T_t = 0$, where the power of channel is concentrated on its diagonal entries

thesis, however, we use a discrete model for the channel. This is simply done by replacing the delays τ_l by l with $l = 0, ..., L-1$.

1.4 Channel Estimation in OFDM

From the above discussion, it became clear that in order to have reliable data detection at the OFDM receiver, the channel should be known, or more realistically, estimated. It is well-known, that for a transmission over a Gaussian channel, at the expense of a 3 dB in the signal-to-noise ratio (SNR), channel estimation can be avoided at the receiver by employing the differential modulation [Pro00]. In this thesis, however, we are interested in coherent modulations and channel estimation algorithms that would resolve the channel at the receiver. Since both the transmitted symbols and the channel are unknown at the receiver, some symbols are usually sacrificed for the insertion of known data. Those are known as pilots. Pilots are then used at the receiver to estimate the channel and then recover the data. The insertion of pilots in OFDM could be either done by adding a preamble at the beginning of the OFDM frame containing only pilots or by the periodic insertion of pilots within the frame.

The first is known as the block type channel estimation and the second as the comb type channel estimation.
Channel estimation is then carried out using the maximum likelihood estimator (MLE), the least squares (LS), the minimum mean square error (MMSE) estimator or the maximum a posteriori (MAP) estimator. Those estimators are based on two different views of the channel impulse response (CIR). While the MLE and LS view the CIR as a deterministic but unknown variable, the MMSE and MAP deal with it as a random variable. A detailed explanation of those estimators is found in [Kay93]. Also, several works in literature dealt with the problem of channel estimation in OFDM for slow time-varying channels [SSR13][SSR15][RHS14b][SRS14] [GHRB12] and fast time-varying channel [HR10][SRHG12][SRH$^+$11].

Classical approaches to channel estimation have two major drawbacks. First, the insertion of pilots decreases the spectral efficiency of the OFDM system by allocating bandwidth to known symbols. Second, they rely on systems with known parameters such as the channel noise variance or system matrices. In case those are not completely known or deviate from their nominal value, those estimators might suffer.

1.5 Conclusion

In this chapter, we have presented the basic OFDM system model for a static, quasi-static and time-varying channels. This system considers multipath fading, which is a typical scenario in an urban medium. The fast time-varying system will be considered in chapter 2 where we address channel estimation in OFDM with narrowband interference (NBI) in a high mobility scenario. The quasi-static channel will be considered in chapter 3 where we address the problem of channel estimation and data detection for data-nulling superimposed pilots (SP). We then briefly discussed conventional channel estimation algorithms in OFDM systems. We will see in chapter 2 and 3 that those algorithms fail in the presence of interference, NBI in chapter 2 and SP in chapter 3.

Chapter 2

EM-based Channel Estimation for OFDM Contaminated by Narrow Band Interference

Contents

2.1 Introduction and Motivation

In the previous chapter, we presented the basic OFDM scheme and the conventional estimation algorithms that allows us to obtain the channel state information (CSI) at the receiver. A main premise for those algorithms is that pilot sub-carriers are interference free so that the channel estimate is reliable. When this is not the case, their performance is expected to suffer.

A report published by the Federal Communications Commission (FCC) [FCC] showed a significant amount of unused radio resources in frequency, time and space. This is due to the current regulatory regime which dedicates frequency bands to licensed users and restricts its use by other users even if it is vacant.

The idea behind cognitive radio (CR) first introduced by Mitola [Mit93] is to use those resources, originally reserved to licensed users, by secondary users while keeping the impact of interference on licensed users at its minimal level. To attain this goal, spectrum sensing capabilities are implemented in CR networks [ALLP12] to assure that radio is free to use. However, interference may still arise. Thus technologies using OFDM as their access scheme and which are likely to be present in a cognitive network, risk interference with other technologies accessing the same band. This interference is known as NBI. The particularity of NBI is that it strikes a small number of sub-carriers in the OFDM symbol with a high power, rendering conventional channel estimation schemes inefficient.

This implies that new channel estimation algorithms should be designed and which are able to cope with NBI. Major difficulties associated with channel estimation in the presence of NBI are the position and the need of the prior knowledge of NBI statistics. Those are practically difficult to obtain.

NBI mitigation methods can be separated into two categories [ZFC04]: digital filtering techniques and interference cancellation techniques. In [Red02], a receiver window which belongs to the first category was adopted in order to reduce noise spreading in multi-carrier systems by lowering the side lobes of the frequency thus requiring that the interference-free CP part of the OFDM symbol is large enough which reduces the spectral efficiency. In [BZ08], a prediction error filter (PEF) which operates by exploiting the flatness of the

OFDM spectrum, was introduced in an attempt to limit spectral leakage of the interference power. However, due to the under utilization of the OFDM sub-carriers (zero frequency, guard bands, flexible sub-carrier allocations schemes), the PEF might suffer from performance degradation.

The interference cancellation techniques, on the other hand, can be also separated into two main approaches. The first is based on Taylor expanded transfer function which maps the disturbance from the interference frequency signal onto OFDM sub-carriers and subtract them from the corresponding received OFDM symbol as in [SNB+04]. The other approach is that which adopts the linear minimum mean square error (LMMSE) criteria. It is important to point out that those methods require the knowledge of the position of NBI within the signal spectrum.

In another approach, the frequency excision method has been proposed in an attempt to avoid the usage of affected frequency bins of OFDM symbols. This approach requires high SNR values and will fail otherwise [VPN11].

In [BZ09], the authors of [BZ08] have proposed the PEF as an erasure insertion mechanism that localizes the erasure to the tones surrounding the interference without affecting other tones. In their two articles, the authors assume a single contaminated tone. In [DV08], a method to predict the error term between the sub-carriers is proposed by assuming that the first sub-carrier is NBI-free. In [HYK+08], a method for detecting and removing jammed pilot tones is exposed yet it is limited to the elimination of only one sub-carrier. The constrained minimum mean output energy (CMMOE) which also requires prior knowledge about the NBI statistics was proposed in [DGPV07]. Second Order Statistics (SOS) of the received data are assumed known at the receiver or estimated from a finite number of data symbols. In [WN05] and [GS02], authors considered spreading the OFDM symbols over sub-carriers using orthogonal codes. However, these methods require modifications to the transmitted OFDM signal which are not supported by current OFDM standards [Cou07].

Channel estimation algorithms have also been proposed in this context. In [GAD11], the authors exploit the inherent sparsity of the NBI signals in the frequency domain and use the compressive sensing (CS) theory to estimate them. In [PL12], the author propose an iterative receiver and use the Expectation-Conditional Maximization (ECM) algorithm to jointly estimate the channel information and detect the data. Recently, in [ZZY13], a robust least square estimation algorithm was presented which requires that the number of pilots be greater than twice the channel order. In [MM09], EM-based joint channel taps and noise power estimation was performed for static channels.

Note that all the aforementioned proposals deal with slow time-varying

channels and no work exists in literature that addresses channel estimation in fast time-varying channels.

In a previous work [SRHG12], an EM approach was developed to estimate fast time-varying OFDM channels in the presence of carrier frequency offset (CFO). The channel was the unwanted parameter and the CFO the parameter to be estimated. This allowed the formulation of the channel estimation with the Kalman smoothing whereas the CFO was obtained through numerical optimization.

In this chapter we propose for the first time a novel estimation framework for fast time-varying OFDM systems contaminated by NBI. This is motivated by the recent proposal [B$^+$14] dealing with CR in high speed railway for which the problem has yet to be addressed. To this end, the approach reported in [SRHG12] is applied to the discussed topic. The original EM formulation is modified such that the noise variances are now the parameters to be estimated and the channel the unwanted parameter. Moreover, in contrast to [SRHG12], this enables the derivation of an analytical formulation for the noise variances estimation We also perform comparisons with existing algorithms that either neglect mobility or NBI. The results confirm that the proposed estimator is well suited for high mobility scenarios. We also demonstrate that this robustness to NBI comes at the cost of a relatively small additional complexity. Furthermore, the robustness to model mismatches is addressed.

Finally, we note that the novel estimation framework proposed for fast time-varying channels with NBI in OFDM, presented in this chapter, has been published in IEEE Wireless Communications Letters [ZS15].

The chapter is organized as follows. In section 2.2, we provide examples of coexisting scenarios in the frequency spectrum. Then, in section 2.3, we review the EM algorithm and its extension to the EM-MAP algorithm. In section 2.4, we detail the system model of our approach and devote section 2.5 for the proposed estimator. We then provide the simulation results and a complexity comparison in sections 2.6 and 2.7 respectively. Section 2.8 concludes this chapter.

2.2 Coexistence in the Frequency Spectrum

As already stated, the spectrum is a scarce resource, despite that, its use is not optimized. Thus, some coexistence scenarios are present in the spectrum. We will briefly expose two coexistence scenarios.

2.2.1 The IEEE 802.11g WLAN and Bluetooth

The unlicensed 2.4 GHz industrial, scientific and medical (ISM) band is attractive for wireless applications. It is both free to use and has good propagation characteristics. It is thus easy to speculate that different wireless technologies will coexist in this band. The IEEE 802.11g standard which is an OFDM-based wireless local area network (WLAN) operates in the 2.4 GHz band along with the Bluetooth standard. Bluetooth (1 MHz) is based on a frequency hoping technology and is viewed by the WLAN (22MHz) as NBI [MM08].

2.2.2 The IEEE 802.11g WRAN in TV White Space

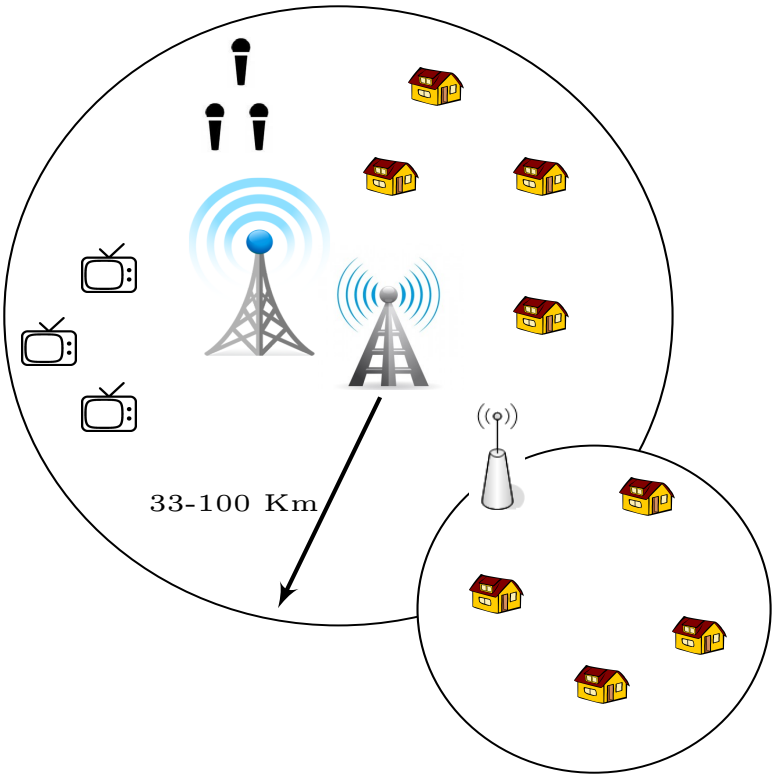


Figure 2.1: WRAN deployment scenario

The IEEE 802.22 wireless regional area network (WRAN) depicted in figure 2.1 is the first wireless standard with cognitive radios. WRAN is a standard for fixed access services in the TV white space. It aims to provide broadband access

to rural areas where the population density does not exceed 60 person/Km^2. The typical coverage area of a WRAN network is 33 Km, while its maximum coverage area could reach up to 100 Km. The goal is thus to develop a standard for a cognitive radio-based PHY/MAC/air interface for use by license-exempt devices on a non-interfering basis in spectrum that is allocated to the TV Broadcast Service. In this band, a CR may use up to 3 consecutive TV channels (18 MHz), whereas police dispatch devices and wireless microphones require approximately 200 KHz of bandwidth and are subsequently considered as NBI primary users in these bands [ASM12].

2.3 Expectation Maximization Algorithm

2.3.1 Introduction

In this section, we give a brief description of the expectation maximization algorithm (EM) and its extension to the EM-MAP algorithm.

The maximum-likelihood (ML) estimate of a parameter Θ is the value that maximizes the likelihood function of Θ for a set of n observed data $\mathcal{X} = [x_1, ..., x_n]$. The likelihood function of Θ given the data $\mathcal{X}$ is given as:

$$p(\mathcal{X}, \Theta) = \prod_{i=1}^{N} p(x_i, \Theta) \tag{2.1}$$

Then the ML estimate is obtained as:

$$\Theta_{ML} = \arg \max_{\Theta} p(\mathcal{X}; \Theta) \tag{2.2}$$

Practically, the maximization is done for $\log p(\mathcal{X}; \Theta)$ since it is easier to deal with [Kay93]. However, when the observed data $\mathcal{X}$ has missing elements or is incomplete, it is no more feasible to obtain the ML estimate and one has to resort to more elaborate techniques. One such technique is the expectation maximization (EM) algorithm. The EM algorithm was first synthesized in the seminal paper [APD77]. Its name corresponds to two iterative steps that make up the algorithm. The expectation (E-step) and the maximization (M-step). It is a broadly applicable approach to the iterative computation ML estimates [MK97]. EM is well-suited for problems with missing data (the data is actually missing) or when the likelihood function is analytically intractable but can be simplified by some assumptions on the missing data. Intuitively, the EM algorithms fills in initial values for the missing data which are the updated by their predicted values using their initial parameter estimation [Bil98].

2.3.2 Mathematical Formulation of EM

The EM algorithm defines two sets of data, the complete and the incomplete data. The incomplete data is the missing observed data $\mathcal{X}$ and the complete data is defined as $\mathcal{Z} = (\mathcal{X}, \mathcal{Y})$ where $\mathcal{Y}$ is the unobserved data. Now, the likelihood function of the complete data $p(\mathcal{Z}, \Theta)$ is a random variable due to the missing information $\mathcal{Y}$. Thus, we calculate the expectation of the log likelihood of the complete data with respect to $\mathcal{Y}$ given the observations $\mathcal{X}$ and the current parameter estimate. This could be formulated as:

$$Q(\Theta, \Theta^{i-1}) = \mathbb{E}[\log p(\mathcal{Z}; \Theta) | \mathcal{X}, \Theta^{i-1}] \tag{2.3}$$

Here, Q is known as the auxiliary function and its formulation makes up the E-step. It thus manufactures data for the complete data problem using the incomplete observed data and the current estimate. It is important to note that the second argument in the Q function is fixed and known at each E-step, while the first argument is the one that conditions the likelihood function.

The M-step simply maximizes the function (Q) calculated in the E-step and is formulated as:

$$\Theta^i = \arg\max_{\Theta} Q(\Theta, \Theta^{i-1}) \tag{2.4}$$

Then the E-step and M-step are repeated iteratively until the value of Θ^i is very close to the value of Θ^{i+1} which is typically measured by their difference compared to a threshold. When the difference is less that this threshold, the algorithm converges. A schematic representation of this process is given in figure 2.2.

2.3.3 Extension to the EM-MAP Algorithm

In certain problems where the likelihood function might have singularities, the EM algorithm might fail. In this case, a simple solution would be to integrate (or impose) some prior information on Θ. This is known as the EM MAP algorithm. The rational behind the EM-MAP is the same as the EM algorithms with slight modifications. First, one has to specify a prior probability density function (pdf) as $p(\Theta)$. In order to follow the same logic as in the previous section, we define the MAP estimate as:

$$\Theta_{MAP} = \arg\max_{\Theta} \{ \log p(\mathcal{X}; \Theta) + \log p(\Theta) \} \tag{2.5}$$

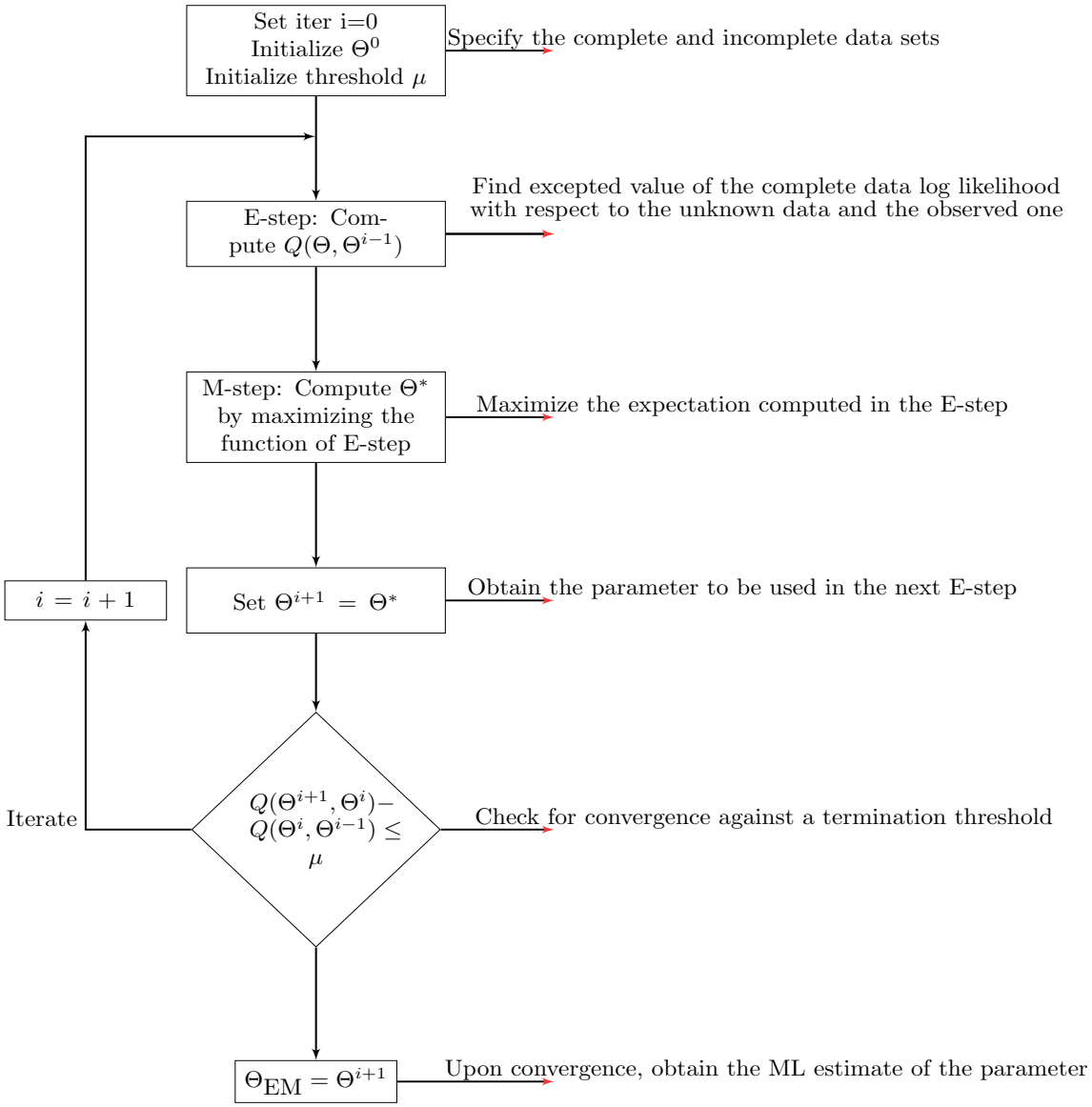


Figure 2.2: EM Algorithm [KML04]

The E-step is similarly performed as in (2.3) with the following modification:

$$Q_{\text{MAP}}(\Theta, \Theta^{i-1}) = Q_{\text{MAP}}(\Theta, \Theta^{i-1}) + \log p(\Theta) \tag{2.6}$$

Now, the M-step is performed as (2.4) but over the modified auxiliary function

as:

$$\Theta^i = \arg \max_{\Theta} Q_{\text{MAP}}(\Theta, \Theta^{i-1}) \tag{2.7}$$

2.4 OFDM System Model with NBI

We first introduce in this section the system design to model time-varying OFDM systems [SRHG12, SRH$^+$11, TCLB07, HR10, HSLR10]. This is identical to the one presented in the previous chapter, except for the additional effect of NBI which has to be taken into account in the system.

For a time-varying OFDM system with N sub-carriers and a cyclic prefix length N_g, each OFDM symbol has a total duration $T_t = N_t T_s$, with T_s being the sampling time and $N_t = N + N_g$. Frames are composed of K OFDM symbols. The transmitted symbols $\mathbf{x}_k = \left[x_{k,0}, ..., x_{k,N-1}\right]^T$ on the kth OFDM symbol is modulated by an N-point inverse fast FT (IFFT). Upon reception, the cyclic prefix is removed and FFT is applied to obtain the frequency domain OFDM received symbol $\mathbf{y}_k$ which is given as:

$$\mathbf{y}_k = \mathbf{H}_k \ \mathbf{x}_k + \mathbf{w}_k, \tag{2.8}$$

where $\mathbf{w}_k$ is a white complex Gaussian noise with zero mean and unknown covariance matrix defined as $\mathbb{E}[\mathbf{w}_k \mathbf{w}_k^H] = \text{diag}(\boldsymbol{\sigma}^2)$ with $\boldsymbol{\sigma}^2 = \left[\sigma_0^2, ..., \sigma_{N-1}^2\right]^T$ where $\sigma_n^2 = \sigma_{TN}^2 + \sigma_{I,n}^2$. This representation is interpreted as the contribution of the thermal noise (σ_{TN}^2) and an unknown interference ($\sigma_{I,n}^2$). In fact, with NBI, the noise is not white. However, since the number of affected sub-carriers is relatively small with respect to the total number of sub-carriers, then for the channel estimation task, it is still acceptable to consider a diagonal covariance matrix which is nearly diagonal [GG03]. In addition, taking this correlation into account might benefit the system by increasing its robustness towards NBI. Thus, as in [MM09], neglecting it could be considered as the worst-case scenario. We show later that although the model itself neglects the correlation between NBI samples, however, the proposed algorithm is able to cope with the actual NBI scenario. Finally, $\mathbf{H}_k$ is the $N \times N$ full channel matrix due to Doppler's shift. The elements of $\mathbf{H}_k$ are given as in (1.8).

Channel taps are assumed to be wide-sense stationary (WSS), narrow-band zero-mean complex Gaussian processes with a variance $\sigma_{\alpha_l}^2$. The conventional Jakes' power spectrum of maximum Doppler frequency f_d is assumed. The average energy of the channel is normalized to one, *i.e.*, $\sum_{l=0}^{L-1} \sigma_{\alpha_l}^2 = 1$.

2.4.1 BEM Channel Model

In high mobility scenarios, the number of parameters to be estimated becomes large due to the rapid variation of the channel. The basis expansion model (BEM) is well known to be an attractive method which allows the approximation of the channel taps as the weighted sum of only a few basis function [TCLB07] [SRH+11] as follows:

$$\boldsymbol{\alpha}_{l,k} = \mathbf{B} \cdot \mathbf{c}_{l,k} + \boldsymbol{\xi}_{l,k}, \tag{2.9}$$

where $\mathbf{B} = [\mathbf{b}_0, ..., \mathbf{b}_{N_c-1}]$ is a $N \times N_c$ matrix that collects the N_c basis functions $\mathbf{b}_d$. Vector $\mathbf{c}_{l,k} = \left[c_{l,k,0}, ..., c_{l,k,N_c-1}\right]^T$ represents the N_c BEM coefficients for the lth channel tap of the kth OFDM symbol, and $\boldsymbol{\xi}_{l,k}$ represents the corresponding BEM modeling error, which can be neglected in case a sufficient number of BEM coefficients is chosen. The optimal BEM coefficients are given by:

$$\mathbf{c}_{l,k} = \left(\mathbf{B}^H\mathbf{B}\right)^{-1}\mathbf{B}^H\boldsymbol{\alpha}_{l,k}. \tag{2.10}$$

Several BEM designs are present in literature, we use the Polynomial BEM (P-BEM). From now on, we can depict the OFDM system model in terms of the BEM and estimate the BEM coefficients rather than the channel matrix entries. Substituting (2.9) in (2.8) yields:

$$\mathbf{y}_k = \boldsymbol{\mathcal{X}}_k \cdot \mathbf{c}_k + \mathbf{w}_k, \tag{2.11}$$

where the $LN_c \times 1$ vector $\mathbf{c}_k$ and the $N \times LN_c$ matrix $\boldsymbol{\mathcal{X}}_k$ are given by:

$$\begin{aligned} \mathbf{c}_k &= \left[\mathbf{c}_{0,k}^T, ..., \mathbf{c}_{L-1,k}^T\right]^T \\ \boldsymbol{\mathcal{X}}_k &= [\mathbf{Z}_{0,k}, ..., \mathbf{Z}_{L-1,k}] \qquad (2.12) \\ \mathbf{Z}_{l,k} &= \left[\mathbf{M}_0 \,\mathrm{diag}\{\mathbf{x}_k\}\, \mathbf{f}_l, ...,\ \mathbf{M}_{N_c-1} \,\mathrm{diag}\{\mathbf{x}_k\}\, \mathbf{f}_l\right], \qquad (2.13) \end{aligned}$$

where vector $\mathbf{f}_l$ is the lth column of the $N \times L$ Fourier matrix $\mathbf{F}$ and $\mathbf{M}_d$ is a $N \times N$ matrix given by:

$$[\mathbf{F}]_{n,l} = e^{-j2\pi(\frac{n}{N}-\frac{1}{2})l}, \tag{2.14}$$

$$[\mathbf{M}_d]_{n,m} = \frac{1}{N}\sum_{q=0}^{N-1}[\mathbf{B}]_{q,d}\; e^{j2\pi\frac{m-n}{N}q}. \tag{2.15}$$

The BEM model can be used to model the low mobility scenario by simply reducing the number of BEM coefficients to one.

2.4.2 The AR Model for $\mathbf{c}_k$

From (2.10), we get that the optimal BEM coefficients $\mathbf{c}_{l,k}$ are correlated complex Gaussian variables with zero-means and correlation matrix given by:

$$\mathbf{R}_{\mathbf{c}_l}^{(p)} = \left(\mathbf{B}^H\mathbf{B}\right)^{-1}\mathbf{B}^H\mathbf{R}_{\boldsymbol{\alpha}_l}^{(p)}\mathbf{B}\left(\mathbf{B}^H\mathbf{B}\right)^{-1}. \tag{2.16}$$

This allows us to represent their dynamics by an auto-regressive (AR) process of order one denoted as $\tilde{\mathbf{c}}_{l,k}$ and represented as follows:

$$\tilde{\mathbf{c}}_{l,k} = \boldsymbol{A}\cdot\tilde{\mathbf{c}}_{l,k-1} + \mathbf{u}_{l,k}, \tag{2.17}$$

where $\mathbf{A}$ is an $N_c \times N_c$ matrix and $\mathbf{u}_{l,k}$ is a $N_c \times 1$ complex Gaussian vector with covariance matrix $\mathbf{U}_l$. The parameters of the AR ($\mathbf{A}$ and $\mathbf{U}_l$) can be computed by the set of the Yule-Walker equations defined as:

$$\mathbf{A} = \mathbf{R}_{\mathbf{c}_l}^{(1)}\left(\mathbf{R}_{\mathbf{c}_l}^{(0)}\right)^{-1}, \qquad \mathbf{U}_l = \mathbf{R}_{\mathbf{c}_l}^{(0)} - \mathbf{A}\mathbf{R}_{\mathbf{c}_l}^{(-1)}, \tag{2.18}$$

yielding the following AR model for $\tilde{\mathbf{c}}$

$$\tilde{\mathbf{c}}_k = \boldsymbol{\mathcal{A}}\cdot\tilde{\mathbf{c}}_{k-1} + \mathbf{u}_k,$$

where $\boldsymbol{\mathcal{A}} = \text{blkdiag}\{\mathbf{A},...,\mathbf{A}\}$ is a $LN_c \times LN_c$ matrix and $\mathbf{u}_k = \left[\mathbf{u}_{0,k}^T,...,\mathbf{u}_{L-1,k}^T\right]^T$ is a $LN_c \times 1$ zero-mean complex Gaussian vector with covariance matrix $\mathbf{U} = \text{blkdiag}\{\mathbf{U}_0,...,\mathbf{U}_{L-1}\}$.

2.5 Proposed Channel Estimator

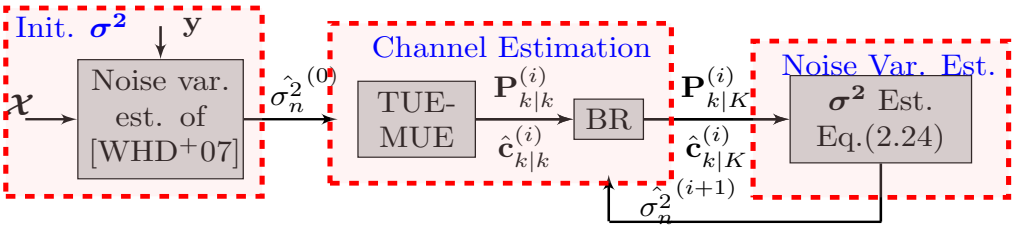


Figure 2.3: Proposed Estimator

2.5.1 EM algorithm

First, we consider a frame of K OFDM received symbols. The objective is to jointly estimate the noise variance vector $\boldsymbol{\sigma^2} = \left[\sigma_0^2, ..., \sigma_{N-1}^2\right]^T$ and the BEM coefficient vector $\mathbf{c}$ of the channel taps. For this, we exploit the functionality of the MAP EM algorithm detailed in section 2.2. First, let $\mathbf{y} = [{\mathbf{y}_0}^T, ..., {\mathbf{y}_{K-1}}^T]^T$ and $\mathbf{c} = [{\mathbf{c}_0}^T, ..., {\mathbf{c}_{K-1}}^T]^T$. We suppose that the transmitted symbols $\mathbf{x}_k$ are known. This configuration is called data aided (DA) and corresponds to the acquisition task. We consider the received data $\mathbf{y}$ as incomplete data, and define the complete data as $\mathbf{z} \overset{\text{def}}{=} \left[\mathbf{y}^T, \mathbf{c}^T\right]^T$. Since the state is described by first order Markov model, the likelihood function of the complete data is given by:

$$p(\mathbf{z}; \boldsymbol{\sigma^2}) = p(\mathbf{c}_0) \prod_{k=1}^{K-1} p(\mathbf{c}_k|\mathbf{c}_{k-1}) \prod_{k=0}^{K-1} p(\mathbf{y}_k|\mathbf{c}_k; \boldsymbol{\sigma}^2).$$

with $\mathbf{C} = \text{diag}(\boldsymbol{\sigma}^2)$, the log-likelihood of the complete data could be expressed as:

$$\ln\left(p(\mathbf{z}; \boldsymbol{\sigma}^2)\right) = -K \ln \sum_{n=0}^{N-1} \sigma_n^2 - \sum_{k=0}^{K-1} \left(\mathbf{y}_k - \boldsymbol{\mathcal{X}}_k \mathbf{c}_k\right)^H \mathbf{C}^{-1} \left(\mathbf{y}_k - \boldsymbol{\mathcal{X}}_k \mathbf{c}_k\right) \quad (2.19)$$

E-step : given the measurements $\mathbf{y}$ and the latest estimate $\hat{\boldsymbol{\sigma^2}}^{(i)}$ from the previous iteration, we calculate:

$$Q(\boldsymbol{\sigma^2}, \hat{\boldsymbol{\sigma^2}}^{(i)}) \overset{\text{def}}{=} \mathbb{E}_{\mathbf{c}|\mathbf{y}, \hat{\boldsymbol{\sigma^2}}^{(i)}} \left[\ln p(\mathbf{z}; \boldsymbol{\sigma^2})\right] \quad (2.20)$$

In order to apply the MAP EM algorithm, we need to utilize an a-priori pdf for the estimated value, $\boldsymbol{\sigma^2}$, $(p(\boldsymbol{\sigma^2}))$. We assume that the entries of $\boldsymbol{\sigma^2}$ are statistically independent distributed according to an inverse-gamma pdf as assumed in [MM09]. It is given by:

$$p(\boldsymbol{\sigma^2}) = \prod_{n=0}^{N-1} \frac{\lambda}{\sigma_n^4} \exp\{-\frac{\lambda}{\sigma_n^2}\} \quad (2.21)$$

where λ is a design parameter which is to be set by simulation in section 2.6.

M-step : this step finds $\hat{\boldsymbol{\sigma^2}}^{(i+1)}$, the value of $\boldsymbol{\sigma^2}$ which maximizes $Q(\boldsymbol{\sigma^2}, \hat{\boldsymbol{\sigma^2}}^{(i)}) + \ln p(\boldsymbol{\sigma^2})$ over all possible values of $\boldsymbol{\sigma^2}$:

$$\hat{\boldsymbol{\sigma^2}}^{(i+1)} = \text{argmax}_{\boldsymbol{\sigma^2}} \left(Q(\boldsymbol{\sigma^2}, \hat{\boldsymbol{\sigma^2}}^{(i)}) + \ln p(\boldsymbol{\sigma^2})\right) \quad (2.22)$$

The iterations of the EM algorithm is terminated when the estimated value of $\boldsymbol{\sigma^2}$ converges towards a stable value. The Q function can be expressed as follows [SRHG12]:

$$Q(\boldsymbol{\sigma^2}, \hat{\boldsymbol{\sigma^2}}^{(i)}) \stackrel{\text{def}}{=} -K \ln \sum_{n=0}^{N-1} \sigma_n^2 - \sum_{k=0}^{K-1} \mathrm{Tr}\{\mathbf{C}^{-1}(\boldsymbol{\mathcal{X}}_k \mathbf{P}_{k|K}^{(i)} \boldsymbol{\mathcal{X}}_k^H + (\mathbf{y}_k - \boldsymbol{\mathcal{X}}_k \hat{\mathbf{c}}_{k|K}^{(i)})(\mathbf{y}_k - \boldsymbol{\mathcal{X}}_k \hat{\mathbf{c}}_{k|K}^{(i)})^H\}, \quad (2.23)$$

This makes it possible to obtain $\hat{\mathbf{c}}_{k|K}^{(i)}$ and $\mathbf{P}_{k|K}^{(i)}$ for all $k = 0, \ldots, K-1$ from the fixed interval Kalman smoother which we detail later. Thus, at each EM iteration, those terms are estimated by using the estimate of $\boldsymbol{\sigma}^2$ at the previous iteration as given below and detailed in Appendix A.2.

$$\hat{\sigma_n^2}^{(i+1)} = \frac{1}{K+2} \sum_{k=0}^{K-1} [\mathbf{M}_k]_{n,n} + \frac{\lambda}{K+2}, \quad (2.24)$$

with

$$\mathbf{M}_k = \boldsymbol{\mathcal{X}}_k \mathbf{P}_{k/K}^{(i)} \boldsymbol{\mathcal{X}}_k^H + (\mathbf{y}_k - \boldsymbol{\mathcal{X}}_k \hat{\mathbf{c}}_{k|K}^{(i)})(\mathbf{y}_k - \boldsymbol{\mathcal{X}}_k \hat{\mathbf{c}}_{k|K}^{(i)})^H. \quad (2.25)$$

2.5.2 Kalman Smoother

The Kalman filter provides the LMMSE for the system state and the MMSE in the case the signal and the noise are jointly Gaussian [Kay93]. From sections 2.4.1 and 2.4.2, we have the set of state space equations as follows:

$$\tilde{\mathbf{c}}_k = \boldsymbol{\mathcal{A}} \cdot \tilde{\mathbf{c}}_{k-1} + \mathbf{u}_k,$$
$$\mathbf{y}_k = \boldsymbol{\mathcal{X}}_k \cdot \mathbf{c}_k + \mathbf{w}_k.$$

The Kalman filter will utilize prior observations to provide the optimal estimate via the set of time update equations (TUE) and measurement update equations (MUE). The TUE project forward in time the current state and error covariance estimate to obtain priori estimates which are input to the MUE, responsible for the feedback and to provide an improved a posteriori estimate. The iterative equations responsible for this process are given by:

Forward recursion:

TUE:

$$\begin{aligned} \hat{\mathbf{c}}_{k|k-1}^{(i)} &= \boldsymbol{\mathcal{A}}\hat{\mathbf{c}}_{k-1|k-1}^{(i)}, \\ \mathbf{P}_{k|k-1}^{(i)} &= \boldsymbol{\mathcal{A}}\mathbf{P}_{k-1|k-1}^{(i)}\boldsymbol{\mathcal{A}}^H + \mathbf{U}. \end{aligned} \tag{2.26}$$

MUE:

$$\begin{aligned} \mathbf{K}_k &= \mathbf{P}_{k|k-1}^{(i)}\boldsymbol{\mathcal{X}}_k^H \left(\boldsymbol{\mathcal{X}}_k\mathbf{P}_{k|k-1}^{(i)}\boldsymbol{\mathcal{X}}_k^H + \mathrm{diag}(\hat{\sigma^2})\right)^{-1}, \\ \hat{\mathbf{c}}_{k|k}^{(i)} &= \hat{\mathbf{c}}_{k|k-1}^{(i)} + \mathbf{K}_k\left(\mathbf{y}_k - \boldsymbol{\mathcal{X}}_k\hat{\mathbf{c}}_{k|k-1}^{(i)}\right), \\ \mathbf{P}_{k|k}^{(i)} &= \mathbf{P}_{k|k-1}^{(i)} - \mathbf{K}_k\boldsymbol{\mathcal{X}}_k\mathbf{P}_{k|k-1}^{(i)}. \end{aligned} \tag{2.27}$$

The integration of backward recursions to enhance the Kalman filter estimate by using all received OFDM symbols and imposing a smoothing effect, results in the Kalman smoother with its equations given below. The schema of the proposed estimator is given in 2.3.

Backward recursion:

$$\begin{aligned} \mathbf{J}_k &= \mathbf{P}_{k-1|k-1}^{(i)}\boldsymbol{\mathcal{A}}^H\mathbf{P}_{k|k-1}^{(i)^{-1}}, \\ \hat{\mathbf{c}}_{k-1|K}^{(i)} &= \hat{\mathbf{c}}_{k-1|k-1}^{(i)} + \mathbf{J}_k\left(\hat{\mathbf{c}}_{k|K}^{(i)} - \hat{\mathbf{c}}_{k|k-1}^{(i)}\right), \\ \mathbf{P}_{k-1|K}^{(i)} &= \mathbf{P}_{k-1|k-1}^{(i)} + \mathbf{J}_k\left(\mathbf{P}_{k|K}^{(i)} - \mathbf{P}_{k|k-1}^{(i)}\right)\mathbf{J}_k^H. \end{aligned} \tag{2.28}$$

2.6 Simulation results

The WRAN has been standardized to operate in the TV bands from 54 to 862 MHz. We use system parameters complying with the WRAN standard and simulate the WRAN channel (Profile A) [H$^+$08]. We set $N = 2084$, $N_g = 368$ and the number of OFDM sub-carriers struck by NBI $N_i = 70$. The discrete-time channel is modeled with a Rayleigh channel with an exponentially decaying power such that: $\sigma_{\alpha_l}^2 = \eta \cdot \exp(-(l+1))$ where η is set so that the power is normalized to unity. In this band, licensed wireless microphones act as NBI to WRAN which are modeled as a Gaussian narrowband process by generating a Gaussian white noise signal followed by a raised cosine filter as in [EF10]. The channel autocorrelation function is assumed to be given by the widely

accepted Jakes' model. Channel estimation is performed by exploiting $K = 2$ training symbols located in the frame preamble carrying pilot symbols on all sub-carrier indexes. The SNR is defined as $10 \log \frac{\sigma_s^2}{\sigma_{TN}^2}$ where σ_s^2 is the power of the received signal, and the signal-to-interference ratio (SIR) is defined as $10 \log \frac{\sigma_s^2}{\sigma_{I,n}^2}$. We perform simulations for values of $f_d T_t$ equal 0.001, 0.01, and 0.05 corresponding to a speed of 7, 70, and 300 Km/h respectively. The number of BEM coefficients N_c used is 1, 2, and 3 corresponding to $f_d T_t$ of 0.001, 0.01, and 0.05 respectively. As indicated in [HR09], the value of N_c sufficient to model the channel taps in the case of high mobility ($f_d T_t = 0.05$) is equal to three and a value of one models the low mobility case. The average mean square error (MSE) defined in (2.29) of the estimate of $\boldsymbol{\alpha}$ denoted as $\hat{\boldsymbol{\alpha}}$ is plotted and the expectation is estimated via 500 Monte Carlo simulations. The precision of the proposed channel estimation algorithm is measured against the Bayesian Cramer–Rao bound (BCRB) with perfect knowledge of $\boldsymbol{\sigma^2}$ as given in [HR09].

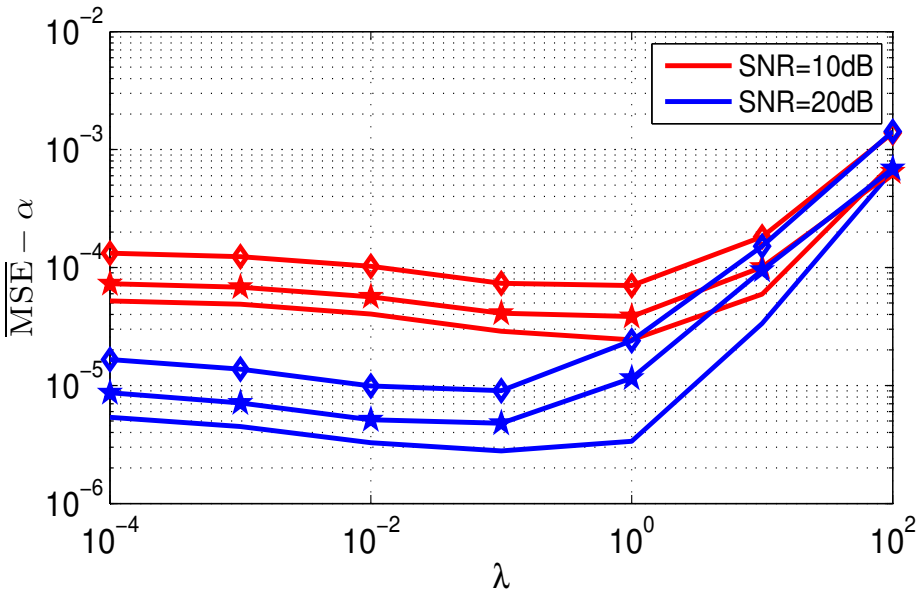


Figure 2.4: MSE performance for $K = 2$, SIR $= 0$ dB, $f_d T_t = 0.05$, $N_c = 3$

$$\overline{\mathrm{MSE}} - \alpha \overset{\mathrm{def}}{=} \frac{1}{KNL} \mathbb{E}\left[(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha})^H (\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}) \right]. \tag{2.29}$$

2.6.1 Algorithm Initialization

To initialize our algorithm, we estimate the scalar σ^2_{TN} by neglecting NBI as in [WHD+07]. We then fix an appropriate value for the variable λ. Figure 2.4 shows the values of λ versus the MSE for the three f_dT_t values for SIR = 0 dB. It could be seen that the appropriate value of λ which is a good trade-off for the ranges of SNR and f_dT_t values is 0.1. Note that this is the same value as in [MM09] for the static case.

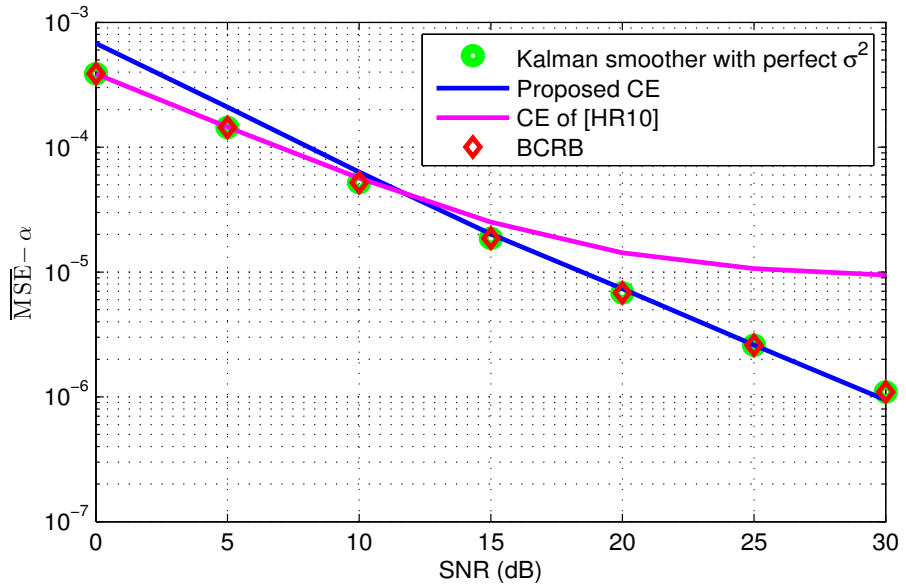


Figure 2.5: MSE performance for $K = 2$, SIR = 0 dB, $f_dT_t = 0.05$, $N_c = 3$

2.6.2 Comparison of Proposed Algorithm with Literature

Since there does not exist an algorithm in literature dealing with channel estimation in high mobility with NBI, we will compare our results to two exiting channel estimators. The first one is a modified version of [HR10] which addresses mobility but ignores NBI. Here, a noise variance estimator which estimates a scalar σ^2_n has been added [WHD+07] and the filtering was substituted with smoothing. The second one is the algorithm proposed in [MM09] which deals with NBI in static channels. We then plot in figure 2.5, the MSE of the proposed channel estimator (CE) and the modified CE of [HR10]. As a lower bound, we plot the BCRB and the Kalman smoother with perfect knowledge of

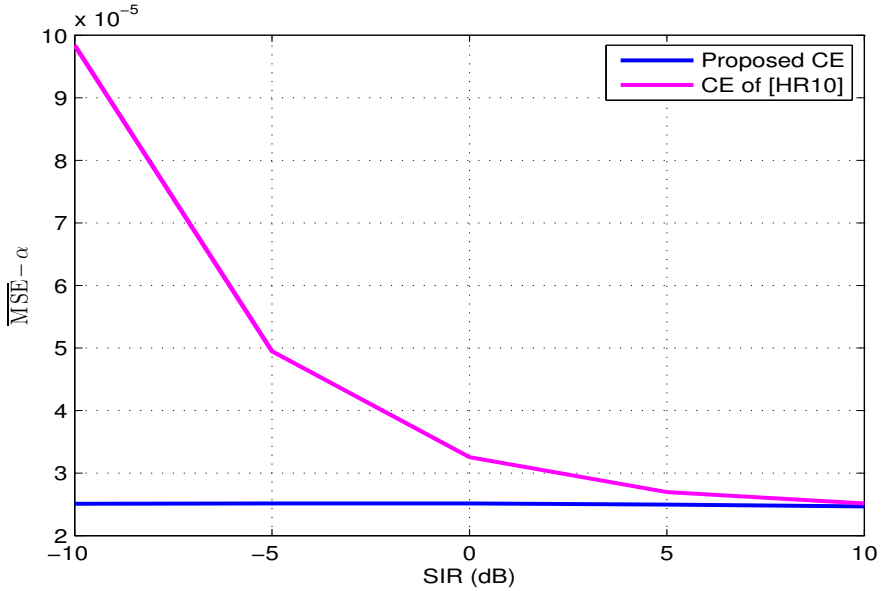


Figure 2.6: MSE performance for $K = 2$, SNR $= 20$ dB, SIR $= 0$ dB

$\boldsymbol{\sigma}^2$, in a high mobility scenario with $f_d T_t = 0.05$, $N_c = 3$ and SIR $= 0$ dB. The plot shows that the Kalman smoother with perfect $\boldsymbol{\sigma}^2$ attains the BCRB for all SNR values and that our algorithm converges towards the BCRB at intermediate SNR values. However, the performance of the modified version of [HR10] degrades with SNR. The algorithm converges after 6 EM iterations. The performance of the proposed CE at low SNR values could be explained by the fact that the estimation of the noise variance at those values could be inaccurate so that neglecting NBI, which is done in [HR10], might be preferable. Note that an identical behavior could be replicated for slow time-varying channels in [MM09]. We then plot in figure 2.6 the MSE performance of the proposed CE and the modified CE of [HR10] versus the SIR. It could be seen that in highly interfered scenarios, the proposed estimator clearly outperforms the estimator of [HR10]. However, as the interference decreases, the MSE of the two estimators converges. To illustrate the need for an algorithm capable of coping with the high mobility scenario, we plot in figure 2.7 the MSE performance of our CE and the CE of [MM09] as a function of $f_d T_t$ for SNR $= 20$ dB and SIR $= 0$ dB. The plot shows that as the speed increases, the CE of [MM09] diverges while our proposed CE maintains a relatively stable performance. Finally, we carry out a bit error rate (BER) study where our proposed CE and the CE of

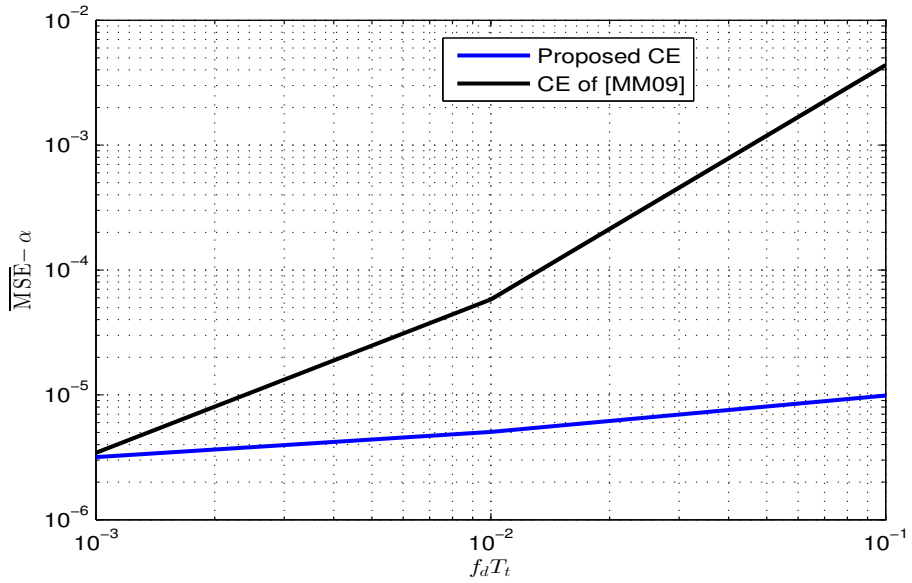


Figure 2.7: MSE performance for $K = 2$, SNR $= 20$ dB, SIR $= 0$ dB

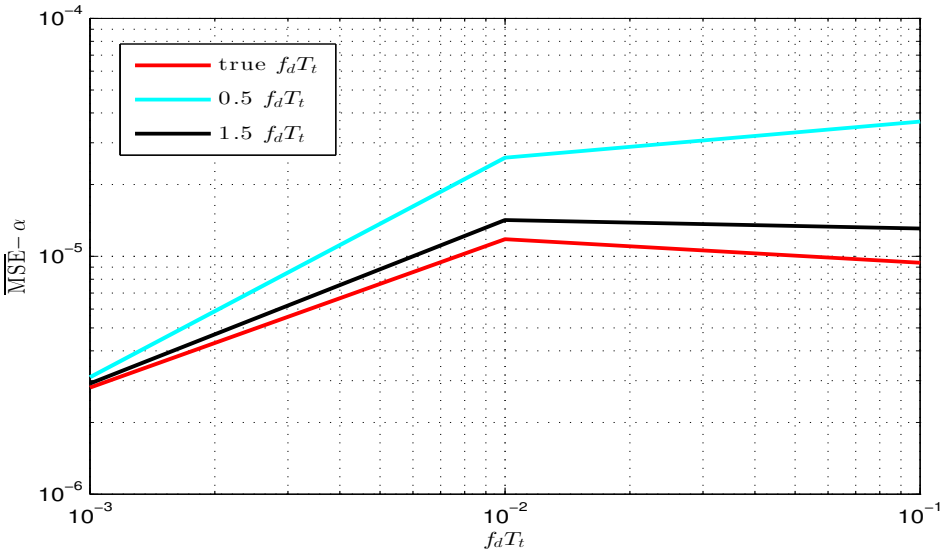


Figure 2.8: MSE performance for $K = 2$, SIR $= 0$ dB, SNR $= 20$ dB, $N_c = 3$

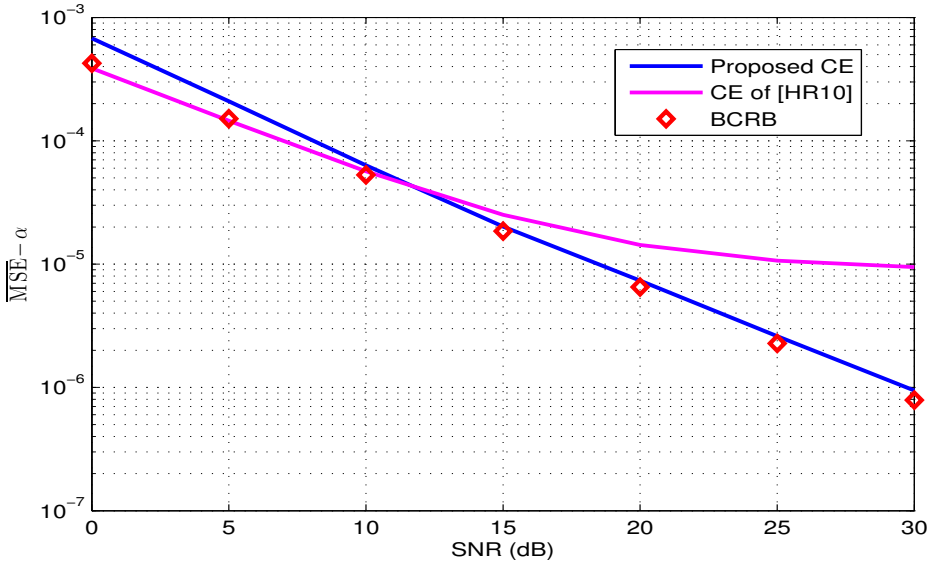


Figure 2.9: MSE performance for $K = 2$, SIR $= 0$ dB, $f_d T_t = 0.05$, $N_c = 3$

the modified version of [HR10] are tested with the data detector of [FRL08], which is designed to handle mobility. In the first iteration, only pilots are used for channel estimation. For subsequent iterations, the soft data estimates from the turbo equalizer and decoder are used as auxiliary pilots. We define L_f, the number of sub-carriers separating two consecutive pilot symbols. Figure 2.10 shows that after only two iterations, for a BER target of 10^{-4}, our algorithm is less than 1 dB away from the CSI. However, the modified version of [HR10] diverges away from the CSI with a loss of more than 3 dB.

2.6.3 Robustness of Proposed Algorithm to Model Mismatch

In an attempt to verify the robustness of the algorithm against model mismatches, we study its performance for a deviation in $f_d T_t$ or in the channel model when the assumed Jakes model is not valid. We plot in figure 2.8 the MSE performance of the algorithm as a function of $f_d T_t$ for a perfect knowledge of $f_d T_t$, for an overestimation and an underestimation of 50% with SNR $= 20$ dB and SIR $= 0$ dB. The plot shows that the algorithm is robust to an overestimation whereas a harsh underestimation (50%) will lead to performance degradation. In figure 2.9, we plot the MSE of the proposed CE when the channel has a flat Doppler spectrum, the CE of [HR10], and the BCRB. The

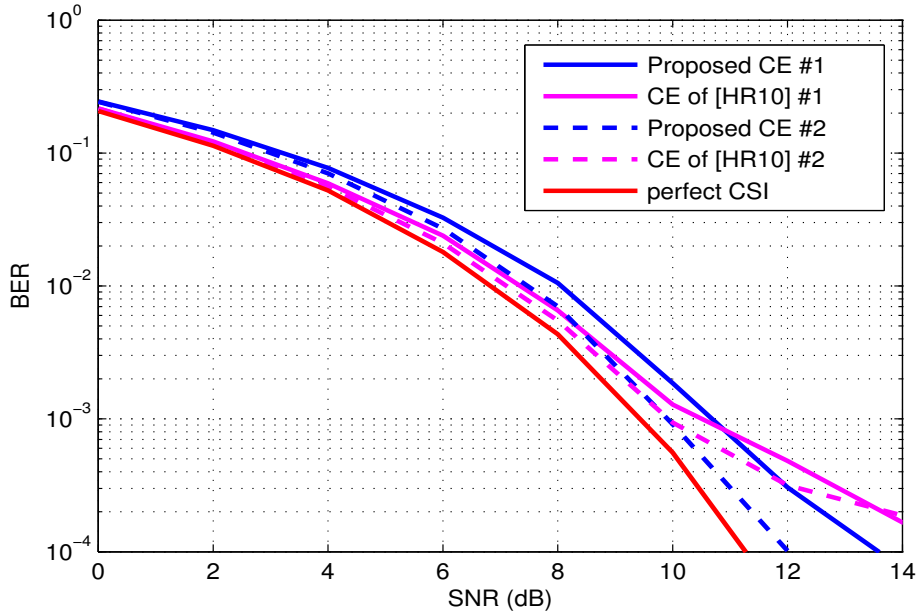


Figure 2.10: BER performance for $K = 4$, $SIR = 0$ dB, $f_d T_t = 0.05$, $N_c = 3$ $L_f = 8$, #:iteration number

plot shows that even with a deviation from the assumed model, the performance of the algorithm is not affected.

2.7 Complexity Comparison with Literature

In order to highlight at what cost the proposed algorithm outperforms the algorithm of [HR10], we calculate the additional complexity required for the proposed algorithm. First, we note that a complex multiplication (CM) is equivalent to 2 real additions (RA) and 4 real multiplications (RM), whereas a complex addition (CA) corresponds to 2 RA. Figure 2.3 in the manuscript shows the three building blocks of the algorithm; the initialization of σ^2, channel estimation and the noise variance estimation. The first two steps (initialization of σ^2 and channel estimation) do not add any complexity to the algorithm since they are equivalently done in [HR10]. The additional complexity is thus due to the noise variance estimation step. For the noise variance estimation, the proposed algorithm needs $NLN_c^2 + NLN_c - N$ complex additions (CA), $2NLN_c + NLN_c^2$ complex multiplications (CM), $2N + 2$ real multiplications

(RM) and $NK + N + 1$ real additions (RA). Since a CM amounts to 4 RM and 2 RA and a CA amounts to 2 RA, then the total additional complexity is $4NLN_c^2 + 6NLN_c + NK - N + 1$ RA and $4NLN_c^2 + 8NLN_c + 2N + 2$ RM. It is noteworthy that the order of magnitude for the calculation of the noise variance is N, which remains small compared to that for calculating the Kalman smoother which is N^3. Thus, this additional complexity has a small effect on the global complexity of the proposed algorithm compared to [HR10]. This is detailed in Appendix A.3.

2.8 Conclusion

In this chapter, a channel estimator for fast time-varying OFDM systems contaminated by NBI has been proposed and analyzed. The proposed estimator is based on the MAP-EM algorithm where we considered the channel as the unwanted parameter and the noise variances as the parameters to be estimated. This formulation allowed us to elegantly integrate the Kalman smoother within the EM algorithm. In addition, a closed-form expression for the noise variance has been obtained. For performance evaluation, we simulated a realistic scenario based on the IEEE 802.22 WRAN standard, which is the first to be standardized with cognitive capabilities. We first showed that for a relatively small additional complexity, the proposed estimator outperforms algorithms designed for fast time-varying OFDM systems and which neglect NBI. In fact, this limited complexity addition is due to the fact that the complexity of the Kalman smoother masks the overall complexity of the algorithm. We then demonstrated that, in contrary to algorithms which have not been designed to handle mobility with NBI, the proposed algorithm is robust to mobility and attains a relatively stable performance up to a speed of 300 Km/h. We also observed that the performance of the algorithm becomes even more interesting, compared to algorithms neglecting NBI, as the interference level in the system increases. Finally, in order to entirely analyze the efficiency of the algorithm, we studied its performance under model mismatches. The results revealed that deviation from the assumed Jake's model does not effect the performance of the algorithm. As for a mismatch in the Doppler frequency, the algorithm is shown to be robust to an overestimation of 50% whereas an equivalent underestimation leads to a performance degradation.

Chapter 3

A Low Complexity Turbo Receiver for Data-Nulling Superimposed Pilots in OFDM

Contents

3.1 Introduction

In the previous chapters, we saw that the acquisition of the CIR requires the insertion of pilot sub-carriers within the OFDM symbol. The ever increasing demand on high data rates implies that the spectrum should be optimized. This led to the birth of superimposed SP.

The classical SP (CSP) scheme [CT05, ANS10, LRF06, OLLM04, NRK11] trades power for bandwidth. In other words, pilots are arithmetically added to data symbols at a fraction of the total transmit power in the time or frequency domain. However, CSP suffers from a degraded channel estimation performance.

To resolve this problem, a new data-nulling SP (DNSP) scheme [DHLG14], inspired by [GMAHS05], has been recently proposed for OFDM systems. In DNSP, the main idea is to precode the data vector before nulling it for pilot insertion. This assures that pilots are interference free and thus a good channel estimation quality is guaranteed. However, the nulling operation leads to interference on data detection. This will eventually necessitate an interference cancellation operation at the receiver. It is interesting to note that the interference that arises in DNSP has a particular structure which we will discuss later.

To deal with this interference, the authors in [DHLG14] proposed a simple iterative reconstruction scheme to improve data detection at the receiver. However, this scheme has limited interference cancellation abilities since it does not profit from the information provided by the decoder to the equalizer even if it exists. This could be understood since it has not been designed in the turbo framework.

Classically, the most efficient interference cancelers are associated with turbo reception. However, their complexity is high and thus their implementation is prohibitive.

Therefore, the question that arises now is how to take advantage of the aforementioned interference structure to design a less complex turbo interference canceler (IC) (also called soft IC). In other words, is it possible to exploit the unitary property of the precoding matrix, present in the interference, to reduce the complexity of the soft IC.

Our aim is thus to propose a suitable receiver for DNSP with an enhanced performance and an acceptable complexity. To do this, we start from the well known highly complex MMSE soft IC. Then, we show that by combining a given approximated version of this MMSE soft IC with the specific structure of the interference, and by using the unitary property of the precoding matrix, the matrix inversion initially needed reduces to the inversion of a diagonal matrix

which is computationally inexpensive. Then, we perform a comparative study for the proposed receiver for DNSP with the CSP with an iterative receiver of the literature [LRF06]. The comparison is done in terms of MSE and BER performance. Finally, for the sake of completeness, a complexity study is performed.

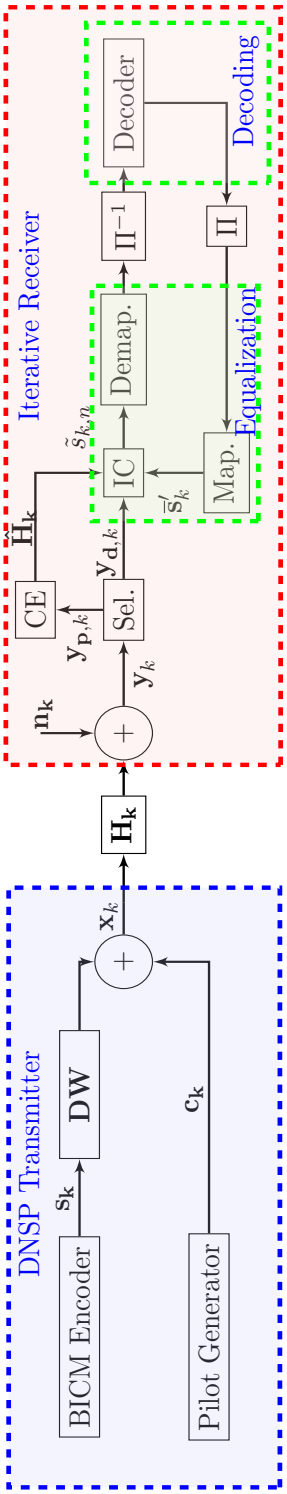
This work has been published in part in [ZSZCD15a] [ZSZCD15b] and the complete analysis has been submitted for review to IET communications.

The chapter is organized as follows. In section 3.2, the system design is given. This is followed by the DNSP transmission model in section 3.3. Then section 3.4 details the proposed iterative receiver for DNSP. We then give the simulation results and discussions in section 3.5. Before concluding the chapter in section 3.7, a detailed complexity calculation is provided.

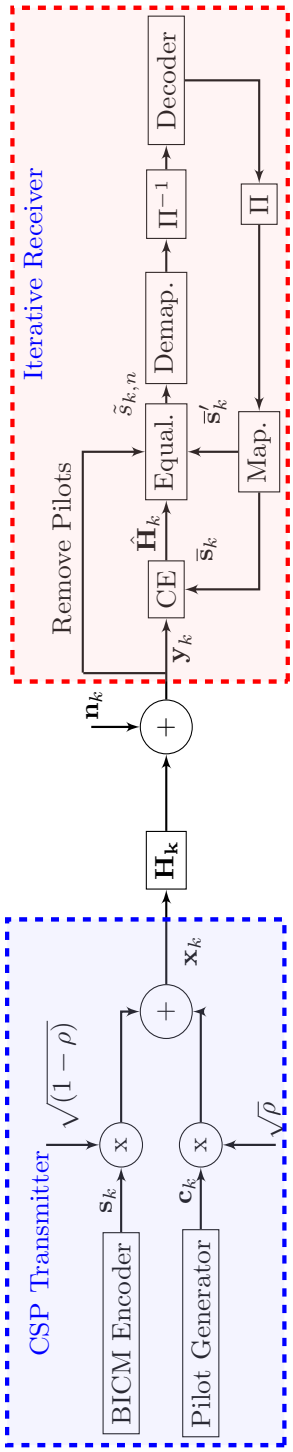
3.2 System Design

3.2.1 Transmitter

The bit-interleaved coded modulation (BICM) scheme, which is a suitable coding scheme for fading channels, is used at the transmitter. It is based on a convolutional code of rate R which is supplied by independent equiprobable binary data. The coded bits are interleaved and mapped onto complex symbols. The transmission chain for the DNSP and CSP scheme is shown in figure 3.1a and 3.1b respectively. A schematic representation of an OFDM symbol which illustrates the pilot and data allocation for the two superimposed pilots schemes (DNSP and CSP) is shown in figure 3.2.



(a) Transmitter-Receiver for the DNSP scheme



(b) Transmitter-Receiver for the DNSP scheme

Figure 3.1: Transmitter-Receiver for SP schemes

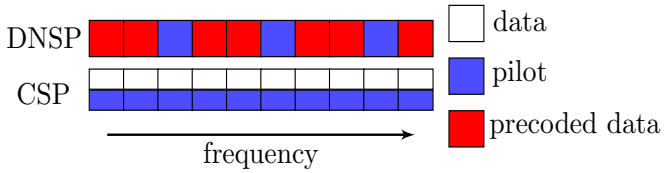


Figure 3.2: Schematic representation of the DNSP and CSP schemes

3.2.2 OFDM System Model

An OFDM system with N sub-carriers and a CP of length N_g is considered. Frames of K OFDM symbols are formed. The total duration of each OFDM symbol is $T_t = N_t T_s$, with T_s being the sampling time and $N_t = N + N_g$. Let $\mathbf{x}_k = \left[x_{k,0}, ..., x_{k,N-1}\right]^T$ be the vector of transmitted symbols on the kth OFDM symbol, $k = 0, ..., K-1$ and $\{x_{k,n}\}$ is the transmitted symbol on the sub-carrier $n - \frac{N}{2}$, $n = 0, ..., N-1$. The kth received OFDM symbol $\mathbf{y}_k$ after removing the CP and applying the fast Fourier transform (FFT) is defined as:

$$\mathbf{y}_k = \text{diag}\{\boldsymbol{\mathcal{H}}_k\}\ \mathbf{x}_k + \mathbf{n}_k, \tag{3.1}$$

where $\mathbf{n}_k$ is a white complex Gaussian noise with zero mean and variance σ_n^2. Note that we use here $\mathbf{H}_k = \text{diag}\{\boldsymbol{\mathcal{H}}_k\}$ to highlight the fact that the channel matrix is diagonal with the entries of $\boldsymbol{\mathcal{H}}_k$ on its main diagonal. $\boldsymbol{\mathcal{H}}_k = [H_{k,0}, ..., H_{k,N-1}]^T$ is the $N \times 1$ vector with $H_{k,n}$ being the frequency-domain channel response at sub-carrier $n - \frac{N}{2}$. Furthermore, $\boldsymbol{\mathcal{H}}_k$ is defined as $\boldsymbol{\mathcal{H}}_k = \mathbf{F}\mathbf{h}_k$ where $\mathbf{F}$ is the $N \times L$ Fourier matrix with its (n,l)th entry given as $[\mathbf{F}]_{n,l} = e^{-j2\pi(\frac{n}{N} - \frac{1}{2})l}$, $n = 0, ..., N-1$, $l = 0, ..., L-1$, and $\mathbf{h}_k$ is the $L \times 1$ vector of the channel impulse response at the kth OFDM symbol where L is the number of channel taps and l indicates the tap number. Note that in this chapter we consider a discrete channel model and use $\mathbf{h}_k$ instead of the physical model with $\boldsymbol{\alpha}_k$. We consider the channel taps are wide-sense stationary (WSS), zero-mean complex Gaussian processes with a variance defined as $\sigma_{h_l}^2$. The average energy of the channel is normalized to one, *i.e.*, $\sum_{l=0}^{L-1} \sigma_{h_l}^2 = 1$. We consider a quasi-static channel where the channel may vary from one OFDM symbol to the other, yet constant within one OFDM symbol. This corresponds to a low mobility scenario. In the next section, we will show how to build $\mathbf{x}_k$ for DNSP.

3.3 DNSP transmission model

The DNSP scheme, defined in [DHLG14], combines the advantage of the pilot-symbol assisted modulation (PSAM) and the CSP schemes. In other words, it is able to overcome the drawback of the deteriorated channel estimation performance in the CSP scheme while keeping its advantage of offering a higher data rate compared to the PSAM scheme. In order to explain how this is done, the data information present on N sub-carriers is precoded so as to have information about all data symbols on each sub-carrier and then certain sub-carriers are nulled for pilot insertion. The spreading operation has the advantage of increasing the transmission diversity and averaging the impairments produced by the nulling operation at the transmitter. This configuration has mainly two consequences, the first is that pilot sub-carriers are interference-free in contrary to pilot sub-carriers in the CSP scheme, and the second is that each data symbol undergoes distortion generated from the other data symbols present on the same sub-carrier. Now define P and D as the number of pilot and data sub-carriers in an OFDM symbol respectively, *i.e.*, $N = D + P$. P pilots are inserted within the OFDM symbol and the pilot indexes are given by $\{pL_f, p = 0 \dots P-1\}$ where L_f is the distance between two adjacent pilots. Define the $N \times N$ unitary precoding matrix $\mathbf{W}$ and the $N \times N$ diagonal matrix $\mathbf{M}$ with ones on data sub-carriers and zeros on pilot sub-carriers. Then $\mathbf{x}_k$ for DNSP is defined as:

$$\mathbf{x}_k = \mathbf{MW}\mathbf{s}_k + \mathbf{c}_k, \tag{3.2}$$

where $\mathbf{s}_k = \left[s_{k,0}, ..., s_{k,N-1}\right]^T$ is the $N \times 1$ data symbols vector on the transmitted kth OFDM symbol where $\mathbb{E}\{|s_{k,n}|^2\} = \sigma_s^2$ and $\mathbf{c}_k$ is the $N \times 1$ pilot vector with pilots on the pilot indexes and zeros elsewhere. Also, $\mathbf{c}_{\mathbf{p},k}$ is the $P \times 1$ pilot sub-vector of $\mathbf{c}_k$ which includes all non-zero pilots $\{c_{p,k}\}$ in $\mathbf{c}_k$, $p = 0, ..., P-1$. Now, $\mathbf{y}_k$ is given as:

$$\mathbf{y}_k = \mathbf{G}_k\mathbf{s}_k + \mathrm{diag}\{\boldsymbol{\mathcal{H}}_k\}\mathbf{c}_k + \mathbf{n}_k. \tag{3.3}$$

where $\mathbf{G}_k$ is defined as:

$$\mathbf{G}_k = \mathrm{diag}\{\boldsymbol{\mathcal{H}}_k\}\mathbf{MW}. \tag{3.4}$$

The formulation of the DNSP scheme in (3.3) renders it clear that the main challenge at the receiver is the data detection process. This is due to the fact that for the detection of each data symbol, the subsequent data symbols found on the same sub-carrier will act as a source of interference on the current symbol. In fact, the interference structure is due to the presence of $\mathbf{W}$, which we exploit later. Thus an IC will be needed at the receiver. This will be

implemented based on the MMSE criterion which will be dealt with in section 3.4. However, the design of the IC assumes the channel knowledge. This is obtained by channel estimation using the interference-free pilot sub-carriers which allows us to use the conventional channel estimators designed for OFDM. This is detailed in the following section.

3.4 Least Squares channel estimation

3.4.1 Obtaining $\hat{\mathbf{h}}_k$

The estimate of $\mathbf{h}_k$, denoted as $\hat{\mathbf{h}}_k$, is obtained in this section with the LS criterion. For this, we introduce $\mathbf{y}_{\mathbf{p},k}$, the $P \times 1$ pilot received vector on the kth OFDM symbol, by selecting the pilot sub-carrier indexes of $\mathbf{y}_k$ and is given as:

$$\mathbf{y}_{\mathbf{p},k} = \text{diag}\{\mathbf{c}_{\mathbf{p},k}\}\mathbf{F}_{\mathbf{p}}\mathbf{h}_k + \mathbf{n}_{\mathbf{p},k}, \tag{3.5}$$

where $\mathbf{F}_{\mathbf{p}}$ is the $P \times L$ Fourier matrix of the pilots with its (p,l)th entry given as $[\mathbf{F}_{\mathbf{p}}]_{p,l} = e^{-j2\pi(\frac{pL_f}{N}-\frac{1}{2})l}$, $p = 0, ..., P-1$, $l = 0, ..., L-1$, and $\mathbf{n}_{\mathbf{p},k}$ is the $P \times 1$ channel noise vector formed by similarly selecting the pilot sub-carrier indexes of $\mathbf{n}_k$. Then $\hat{\mathbf{h}}_k$ is given as [Kay93]:

$$\begin{aligned}\hat{\mathbf{h}}_k &= (\mathbf{F}_{\mathbf{p}}^H\mathbf{F}_{\mathbf{p}})^{-1}\mathbf{F}^H\text{diag}\{\mathbf{c}_{\mathbf{p},k}\}^H\mathbf{y}_{\mathbf{p},k} \\ &= \frac{1}{P}\mathbf{F}_{\mathbf{p}}^H\text{diag}\{\mathbf{c}_{\mathbf{p},k}\}^H\mathbf{y}_{\mathbf{p},k}\end{aligned} \tag{3.6}$$

Note that the pilots are chosen such that $|c_{p,k}|^2 = 1$.

3.4.2 Using $\hat{\mathbf{h}}_k$

As already mentioned and as we will see in the following sections, the IC needs the knowledge of $\mathbf{h}_k$. Thus, in order to be able to calculate the IC equations, we proceed in two steps. First, we assume perfect channel knowledge to obtain the IC equations, which depend on the unknown channel $\mathbf{h}_k$. We then simply replace $\mathbf{h}_k$ by its estimate $\hat{\mathbf{h}}_k$ obtained in (3.6). This approach implies ignoring the channel estimation error and is known as the *mismatched* approach.

Finally, after the data symbols are detected, they could be used to enhance the channel estimation performance in the next iterations.

3.5 Proposed Iterative Receiver for DNSP

After channel estimation is performed, pilot sub-carriers are discarded since they do not provide any information about the data. For this, we introduce $\mathbf{y}_{\mathbf{d},k}$, the $D \times 1$ data received vector on the kth OFDM symbol, by selecting the data sub-carrier indexes of $\mathbf{y}_k$ as:

$$\mathbf{y}_{\mathbf{d},k} = \mathbf{G}_{\mathbf{d},k}\mathbf{s}_k + \mathbf{n}_{\mathbf{d},k} \tag{3.7}$$

where $\mathbf{G}_{\mathbf{d},k}$ is a $D \times N$ full sub-matrix of $\mathbf{G}_k$ after removing the zero rows. It is defined as:

$$\mathbf{G}_{\mathbf{d},k} = \text{diag}\{\boldsymbol{\mathcal{H}}_{\mathbf{d},k}\}\mathbf{W}_{\mathbf{d}}, \tag{3.8}$$

where $\boldsymbol{\mathcal{H}}_{\mathbf{d},k}$ is the $D \times 1$ vector of the data sub-carrier indexes of $\boldsymbol{\mathcal{H}}_k$, $\mathbf{W}_{\mathbf{d}}$ is the $D \times N$ sub-matrix of $\mathbf{W}$ formed by selecting the rows of data sub-carrier indexes. We first review in section 3.5.1 the MMSE soft IC of the literature. Then, we detail our proposed receiver in section 3.5.2. It is well-known that the use of the MMSE soft IC of the literature has a high complexity. We then propose to decrease this complexity by exploiting the structure of the interference which arises in DNSP and implement an approximated MMSE soft IC.

3.5.1 MMSE soft IC of the Literature

The derivation of the MMSE soft IC necessitates the knowledge of the channel. We first provide the formulation of the MMSE soft IC of the literature assuming perfect channel knowledge in (3.11). Then, we formulate it again with the estimated channel in (3.12).

The linear parallel soft interference canceler (PSIC) based on the MMSE criterion is employed in order to avoid the prohibitive complexity of joint equalization and decoding and that of the optimal MAP-based turbo equalizers. The MMSE soft IC structure, introduced in the framework of MMSE equalization in the turbo scheme [LGL01], [GLL97], eliminates the interference that arises during transmission and results in a full channel matrix. The iterative receiver consists of two stages, equalization and channel decoding, see figure 3.1(a). Note that the selection operation (Sel.) in figure 3.1(a) selects either the pilot or the data sub-carriers. The exchange of soft information between the two stages allows us to enhance the decoding process at each iteration. This is assured via the calculation of the log-likelihood ratio (LLR). The equalization stage consists of three main blocks: interference cancellation, mapping and demapping. Upon receiving the transmitted symbol, the IC calculates the

equalized symbol as in figure 3.3:

$$\tilde{s}_{k,n} = \mathbf{w}_{k,n}^H \mathbf{y}_{\mathbf{d},k} - \mathbf{q}_{k,n}^H \overline{\mathbf{s}}'_{k,n}, \tag{3.9}$$

$$\mathbf{q}_{k,n} = \mathbf{G}_{\mathbf{d},k}^H \mathbf{w}_{k,n}, \tag{3.10}$$

where $\overline{\mathbf{s}}'_{k,n} = \left[\overline{s}_{k,0}, ..., \overline{s}_{k,n-1}, 0, \overline{s}_{k,n+1}, ..., \overline{s}_{k,N-1}\right]^T$ is the $N \times 1$ vector of the estimated data symbols and $\overline{s}_{k,n} = \mathbb{E}_{|LLR}\{s_{k,n}\}$, where $\mathbb{E}_{|LLR}\{\cdot\}$ is the expected value given the LLRs. Note that in $\overline{\mathbf{s}}'_k$, the nth entry is nulled in order to avoid using a priori information for the current symbol. The equalized symbols are then supplied to the demapper to calculate the confidence information for each bit. After decoding, the mapper reconstructs the soft symbols and supplies the IC with $\overline{\mathbf{s}}'_{k,n}$ and the variance of the symbol. In order to assure that the equalization and the decoding stages are decorrelated, they are separated by an interleaver or deinteleaver. Now, after defining the $N \times 1$ vector $\mathbf{e}_n = [0, ..., 1, ..., 0]^T$, with 1 on the nth entry and zeros elsewhere, the equalization vector $\mathbf{w}_{k,n}$ is given as:

$$\mathbf{w}_{k,n} = \sigma_s^2 [\mathbf{G}_{\mathbf{d},k} \mathbf{V}_{k,n} \mathbf{G}_{\mathbf{d},k}^H + \sigma_n^2 \mathbf{I}]^{-1} \mathbf{G}_{\mathbf{d},k} \mathbf{e}_n, \tag{3.11}$$

with $\mathbf{V}_{k,n} = \text{diag}[v_{k,0}, ..., v_{k,n-1}, \sigma_s^2, v_{k,n+1}, ..., v_{k,N-1}]$ being the $N \times N$ diagonal matrix where $v_{k,n}$ is defined as $v_{k,n} = \mathbb{E}_{|LLR}\{|(s_{k,n})|^2\} - |\overline{s}_{k,n}|^2$. $v_{k,n}$ represents the variance of the soft estimate $\overline{s}_{k,n}$ and provides confidence information in those estimates. Also, $\mathbf{I}$ is the identity matrix. When perfect a priori information is present, $v_{k,n}$ tends towards 0, whereas it tends towards σ_s^2 when $\overline{s}_{k,n}$ is random and thus no confidence information about $\overline{s}_{k,n}$ is present. Now, the MMSE soft IC with the estimated channel for DNSP is obtained by replacing the channel with its estimate calculated in section 3.4 and $\mathbf{G}_{\mathbf{d},k}$ by its value in (3.8) as:

$$\mathbf{w}_{k,n} = \sigma_s^2 [\text{diag}\{\hat{\boldsymbol{\mathcal{H}}}_{\mathbf{d},k}\} \mathbf{W}_{\mathbf{d}} \mathbf{V}_{k,n} \mathbf{W}_{\mathbf{d}}^H \text{diag}\{\hat{\boldsymbol{\mathcal{H}}}_{\mathbf{d},k}\}^H + \sigma_n^2 \mathbf{I}]^{-1} \text{diag}\{\hat{\boldsymbol{\mathcal{H}}}_{\mathbf{d},k}\} \mathbf{W}_{\mathbf{d}} \mathbf{e}_n. \tag{3.12}$$

The formulation in (3.12) reveals the high complexity of the IC of the literature in terms of latency for the calculation of $\mathbf{V}_{k,n}$ and the number of matrix inversions carried out, which is equivalent to the number of sub-carriers in the OFDM system. This motivates us to seek an alternative with a lower complexity while limiting the performance degradation accompanied by the reduced complexity.

We will refer to this formulation as the *mismatched exact IC*. The *mismatched* refers to the use of the estimated channel instead of the true channel

and the *exact* refers to the use of the variance of each data symbol in matrix $\mathbf{V}_{k,n}$ in the formulation of (3.12).

3.5.2 Proposed Structure

From equation (3.12), we notice that the term $\mathbf{W_d}\mathbf{V}_{k,n}\mathbf{W_d}^H$ embedded within the matrix inversion does not allow us to exploit the property of matrix $\mathbf{W_d}$ ($\mathbf{W_d}\mathbf{W_d}^H = \mathbf{I}$) due to the presence of matrix $\mathbf{V}_{k,n}$. We first point out that $\mathbf{V}_{k,n}$ is a diagonal matrix so in order to profit from this property, we can approximate the entries of $\mathbf{V}_{k,n}$ with a unique representative value on all its diagonal entries. One solution to achieve this is to calculate the variance based on the expected value of $v_{k,n}$ calculated over all LLRs at the output of the channel decoder. In fact, the authors in [TKS02, BID03, TSK02] already resorted to this approximation in order to reduce the complexity at the receiver. This could be formulated as:

$$\mathbb{E}\{v_{k,n}\} = \mathbb{E}\{|s_{k,n}|^2\} - \mathbb{E}\{|\bar{s}_{k,n}|^2\} = \sigma_s^2 - \sigma_{\bar{s}}^2. \tag{3.13}$$

Assuming the ergodicity of $\bar{s}_{k,n}$, $\sigma_{\bar{s}}^2$ can be calculated as:

$$\sigma_{\bar{s}}^2 \approx \frac{1}{KN}\sum_{k=0}^{K-1}\sum_{n=0}^{N-1}|\bar{s}_{k,n}|^2. \tag{3.14}$$

Now, (3.12) can be reformulated as:

$$\begin{aligned}\mathbf{w}_{k,n} = \sigma_s^2[&\mathrm{diag}\{\hat{\boldsymbol{\mathcal{H}}}_{\mathbf{d},k}\}(\sigma_s^2 - \sigma_{\bar{s}}^2)\mathrm{diag}\{\hat{\boldsymbol{\mathcal{H}}}_{\mathbf{d},k}\}^H + \sigma_n^2\mathbf{I} + \\ &\sigma_{\bar{s}}^2\mathrm{diag}\{\hat{\boldsymbol{\mathcal{H}}}_{\mathbf{d},k}\}\mathbf{W_d}\mathbf{e}_n\mathbf{e}_n^T\mathbf{W_d}^H\mathrm{diag}\{\hat{\boldsymbol{\mathcal{H}}}_{\mathbf{d},k}\}^H]^{-1}\mathrm{diag}\{\hat{\boldsymbol{\mathcal{H}}}_{\mathbf{d},k}\}\mathbf{W_d}\mathbf{e}_n.\end{aligned} \tag{3.15}$$

By using Woodbury identity defined as:

$$(\mathbf{A} + \mathbf{u}\mathbf{v}^H)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^H\mathbf{A}^{-1}}{1 + \mathbf{v}^H\mathbf{A}^{-1}\mathbf{u}} \tag{3.16}$$

where $\mathbf{A} = \mathrm{diag}\{\hat{\boldsymbol{\mathcal{H}}}_{\mathbf{d},k}\}(\sigma_s^2 - \sigma_{\bar{s}}^2)\mathrm{diag}\{\hat{\boldsymbol{\mathcal{H}}}_{\mathbf{d},k}\}^H + \sigma_n^2\mathbf{I}$, $\mathbf{u} = \sigma_{\bar{s}}^2\mathrm{diag}\{\hat{\boldsymbol{\mathcal{H}}}_{\mathbf{d},k}\}\mathbf{W_d}\mathbf{e}_n$ and $\mathbf{v}^H = \mathbf{e}_n^T\mathbf{W_d}^H\mathrm{diag}\{\hat{\boldsymbol{\mathcal{H}}}_{\mathbf{d},k}\}^H$, then (3.15) can be written as:

$$\mathbf{w}_{k,n} = \lambda_{k,n}\tilde{\mathbf{w}}_{k,n}, \tag{3.17}$$

where $\tilde{\mathbf{w}}_{k,n}$ and $\lambda_{k,n}$ are defined as:

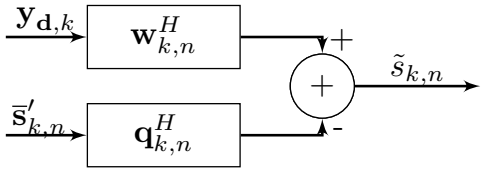


Figure 3.3: Equalizer Block

$$\tilde{\mathbf{w}}_{k,n} = \sigma_s^2 \mathrm{diag}\Big\{\Big[\frac{\hat{H}_{k,0}}{|\hat{H}_{k,0}|^2(\sigma_s^2-\sigma_{\bar{s}}^2)+\sigma_n^2},\ldots,\frac{\hat{H}_{k,D-1}}{|\hat{H}_{k,D-1}|^2(\sigma_s^2-\sigma_{\bar{s}}^2)+\sigma_n^2}\Big]\Big\}\mathbf{W}_{\mathbf{d}}\mathbf{e}_n. \tag{3.18}$$

$$\lambda_{k,n} = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_{\bar{s}}^2 \mathbf{e}_n^T \mathbf{W}_{\mathbf{d}}^H \mathrm{diag}\{\hat{\boldsymbol{\mathcal{H}}}_{\mathbf{d},k}\}^H \tilde{\mathbf{w}}_{k,n}}. \tag{3.19}$$

In fact, the implementation of the approximation in (3.13) allowed the authors of [TKS02, BID03, TSK02] to reduce the number of matrix inversions to a single inversion compared to N inversions originally needed for the *mismatched exact IC*. From equation (3.18), we can see that implementing this approximation for DNSP further reduces this inversion to a diagonal matrix inversion which is computationally inexpensive. We will refer to this formulation as the *mismatched proposed IC.*

3.6 Simulation Results and Discussions

3.6.1 *mismatched exact IC* vs *mismatched proposed IC*

For the simulations, a discrete-time channel with 6 channel taps is considered and modeled as a Rayleigh channel with an exponentially decaying power given by $\eta \cdot \exp(-(l+1))$ where η is set so that the power is normalized to unity. We consider a quasi-static channel and set $N = 128$ sub-carriers, $N_g = 8$ and compose frames of $K = 5$ OFDM symbols. For the DNSP scheme, pilots are inserted at evenly spaced intervals with $L_f = 8$, and data is precoded with the $N \times N$ Walsh Hadamard matrix. Gray mapped QAM modulation is used together with the non-recursive non-systematic convolutional (NRNSC) code $[5, 7]_8$. As a lower bound for the BER plots, we plot the performance of the algorithm with perfect knowledge of the channel CSI. The received SNR for the DNSP scheme is given as $10 \log \frac{D}{N\sigma_n^2}$.

We will first compare the performance of the DNSP with the *mismatched exact IC* and the *mismatched proposed IC*. For that, we plot in figure 3.4,

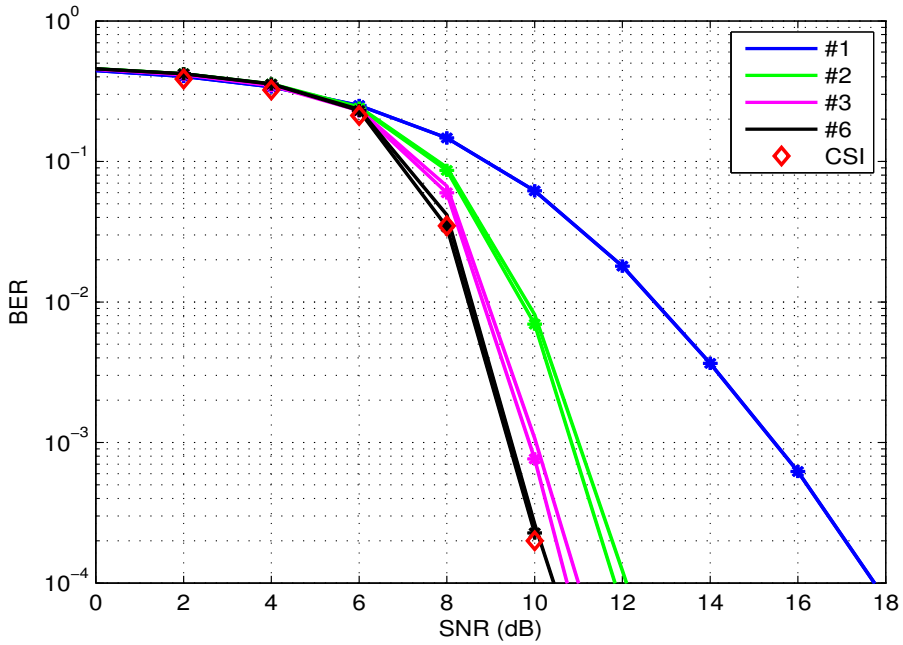


Figure 3.4: BER Performance for 4QAM & coding rate=$\frac{1}{2}$. Solid line: mismatched proposed IC, line with asterisks: mismatched exact IC, #:number of iteration

the BER performance of DNSP with both ICs for a 4QAM modulation with a coding rate of $\frac{1}{2}$. We can see that for early iterations there is a negligible loss in performance of around 0.3 dB at a target BER of 10^{-4}, however, this also fades away at convergence and both ICs perform identically. In figure 3.5, equivalent plots are shown for a 16QAM modulation. We can replicate a close performance as in the case of 4QAM. However, the performance loss due to the approximation in the *mismatched proposed IC* is more important for 16QAM. Nevertheless, at convergence, the loss due to the *mismatched proposed IC* is around 0.8 dB with respect to the *mismatched exact IC* and the CSI. This difference will prove to be negligible compared to the complexity saved by the *mismatched proposed IC* with respect to the *mismatched exact IC*. We thus calculate the complexity of the *mismatched exact IC* and the *mismatched proposed IC* in terms of the required number of real RA and RM. Those are given in table 3.1 and their detailed calculation in Appendix B.2. We can see from table 3.1 that the complexity of the *mismatched exact IC* which is in the

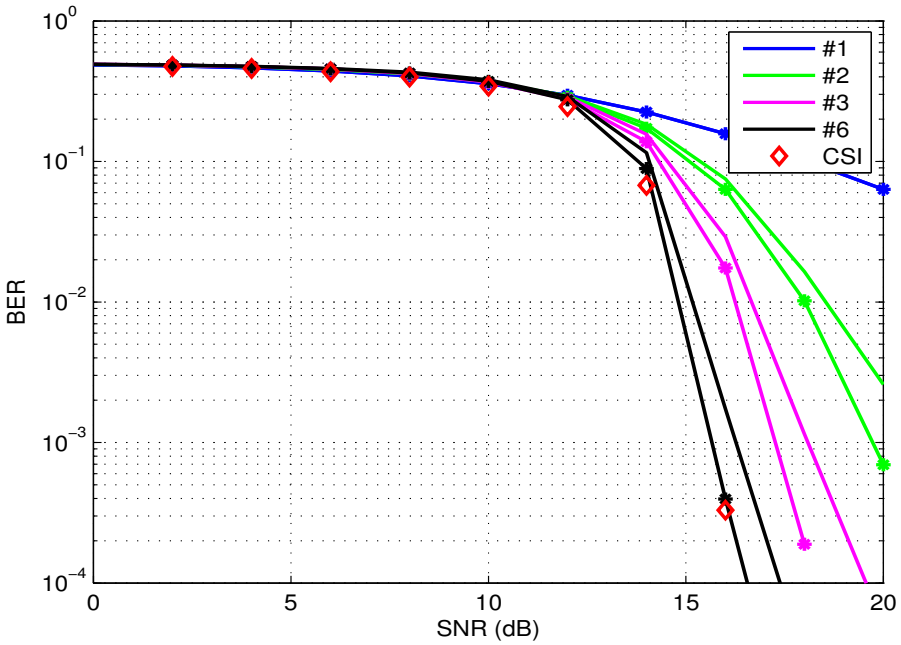


Figure 3.5: BER Performance for 16QAM & coding rate=$\frac{1}{2}$. Solid line: mismatched proposed IC, line with asterisks: mismatched exact IC, #:number of iteration

Table 3.1: Complexity for different IC

Operation	*mismatched exact IC*	*mismatched proposed IC*
RA	$4D^3 + 4ND + 2D^2 + 5D - 3$	$4ND + 5D$
RM	$4D^3 + 6ND + 8D^2 + 3D$	$6ND + 13D + 5$

order of D^3 is much higher than that of the *mismatched proposed IC* which has an order of ND.

3.6.2 DNSP with *mismatched proposed* IC vs CSP in turbo reception framework

For the sake of completeness, we found it interesting to compare the two SP schemes in the turbo framework. As already stated, in CSP, pilots are superimposed on data sub-carriers at a fraction of the transmitted power. To construct the model of the transmitted signal $(\mathbf{x}_k)$ in CSP, let us define the $N \times 1$ pilot sequence for CSP with pilots on all sub-carriers. Now, $\mathbf{x}_k$ is defined as: (see figure 3.1b)

$$\mathbf{x}_k = \sqrt{\rho}\mathbf{p}_k + \sqrt{(1-\rho)}\mathbf{s}_k, \tag{3.20}$$

where ρ is the power allocation ratio. Then, the received OFDM symbol $\mathbf{y}_k$ is given as:

$$\mathbf{y}_k = \boldsymbol{\mathcal{H}}_k(\sqrt{\rho}\mathbf{p}_k + \sqrt{(1-\rho)}\mathbf{s}_k) + \mathbf{n}_k, \tag{3.21}$$

For this pilot scheme, the interference originates from pilots on data and vice versa. This is due to their presence on the same sub-carrier. Here, for data detection, the interference originates from pilot symbols which are known and thus the interference cancellation is much simpler than in the DNSP scheme since only the effect of pilot should be removed. However, for channel estimation, the interference of pilots with data symbols degrades the performance. In the turbo framework, the latter is expected to enhance due to the integration of the soft estimated symbols in the channel estimation process. We will use the framework proposed in [LRF06] for CSP, and will detail it briefly hereafter.

For the first iteration, channel estimation is similarly carried out as in (3.6), where only pilots are used for channel estimation as follows:

$$\hat{\mathbf{h}}_k = (\mathbf{F}^H \mathrm{diag}\{\sqrt{\rho}\mathbf{p}_k\}^H \mathrm{diag}\{\sqrt{\rho}\mathbf{p}_k\}\mathbf{F})^{-1}\mathbf{F}^H \mathrm{diag}\{\sqrt{\rho}\mathbf{p}_k\}^H \mathbf{y}_k. \tag{3.22}$$

For later iterations, $\mathrm{diag}\{\sqrt{\rho}\mathbf{p}_k\}$ is replaced by $\mathrm{diag}\{\sqrt{\rho}\mathbf{p}_k + \sqrt{(1-\rho)}\bar{\mathbf{s}}_k\}$ to enhance the channel estimation performance as already stated. The estimated symbols are also evaluated by soft decisions via the LLR calculation.

Now, we are interested in the relative performance of the two superimposed pilot schemes. The received SNR for the CSP scheme is defined as $10\log\frac{(1-\rho)}{\sigma_n^2}$. We thus set $(1-\rho) = \frac{D}{N}$ in order to guarantee equal data and pilot power allocation for both schemes to perform a fair comparison. We set $\rho = 0.125$ for the CSP scheme which corresponds to a value of $L_f = 8$ for the DNSP scheme. We will first study the channel estimation performance and plot the MSE in figure 3.6 and then the decoding performance and plot the BER performance in figure 3.7. In figure 3.6, we see that the MSE for the DNSP outperforms

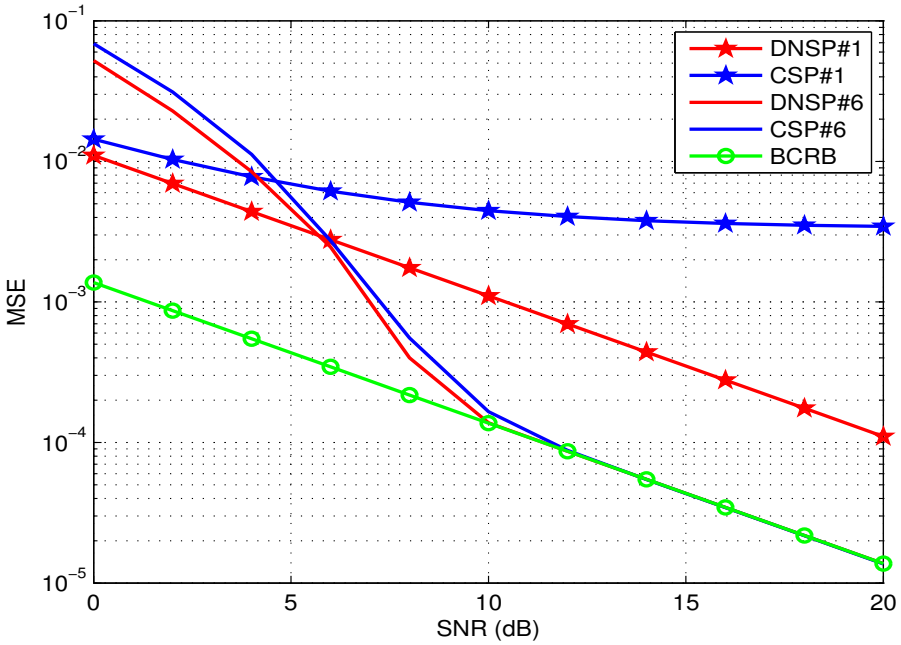


Figure 3.6: MSE Performance for CSP and DNSP , $\rho = 0.125$, $L_f = 8$ #:number of iteration

that of the CSP for the first iteration since there is no interference on channel estimation in this scheme. However, at convergence, both schemes attain the BCRB. We recall here, that this gap between the first iteration for DNSP and the last iteration is due to the use of pilots only in the first and all sub-carriers in the subsequent iterations. In figure 3.7, we see that the BER of the CSP is better than that of the DNSP for the first iteration which could be explained by the fact that for the first iteration, the interference on data for the CSP is removed by excluding the pilots. Whereas for the DNSP scheme, the iterations are needed for interference removal. We set the number of iterations to 6 and we later study the convergence of the different algorithms. Finally, we note from figure 3.7 that at convergence, both schemes attain their CSI curves. However, the CSI curve for the DNSP is lower than that of the CSP. At a target BER of 10^{-4}, there is an SNR loss of more than 3 dB.

In order to explain this difference, we look closely at the inherent structure of the two SP schemes. Although both schemes are guaranteed to have equal pilot and data power, yet the precoding matrix ($\mathbf{W}$), which is a basic element

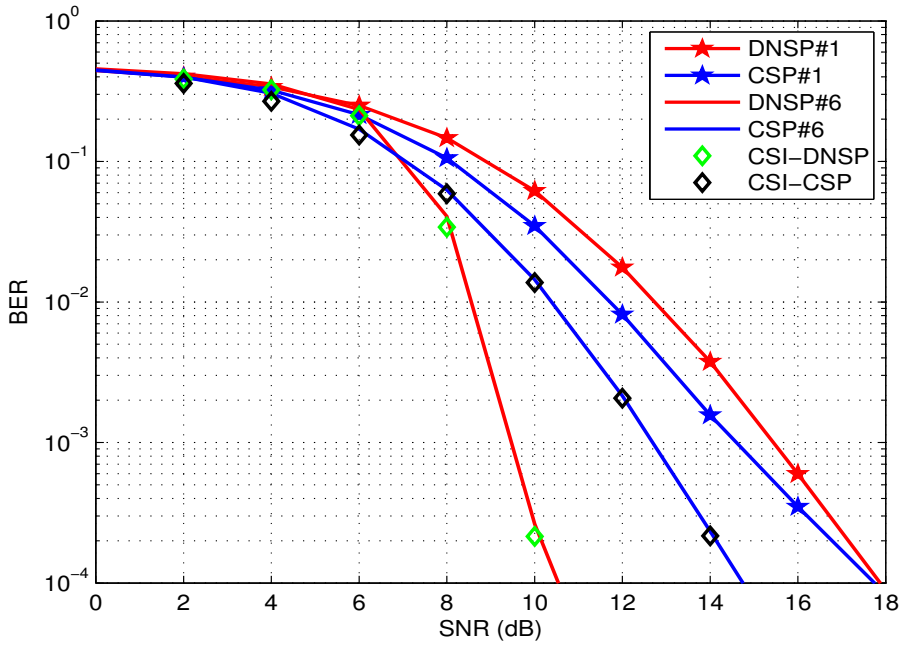


Figure 3.7: BER Performance for CSP and DNSP , $\rho = 0.125$, $L_f = 8$ #:number of iteration

for the construction of the DNSP, increases the transmission diversity, which explains this difference. In order to verify this, we propose to construct a coded CSP (C-CSP) scheme. The idea here is to conserve the same CSP structure but rather than adding the pilots over the data directly, we first precode the data with $\mathbf{W}$ and then add the pilots. Now, we have two sources of interference on data, the pilots and the data themselves due to the precoding operation. Thus, for data detection, two interference cancellation steps should be done, remove the pilots from the received signal and then perform interference cancellation using the *mismatched proposed IC* for the DNSP scheme. Now, the transmitted signal ($\mathbf{x}_k$) is given as:

$$\mathbf{x}_k = \sqrt{\rho}\mathbf{p}_k + \sqrt{(1-\rho)}\mathbf{W}\mathbf{s}_k. \tag{3.23}$$

We now plot the BER of the three SP schemes in figure 3.8. We can see that the performance of the C-CSP becomes identical to that of the DNSP. This proves that the difference in performance loss observed in figure 3.7 is due to the precoding operation in the DNSP scheme.

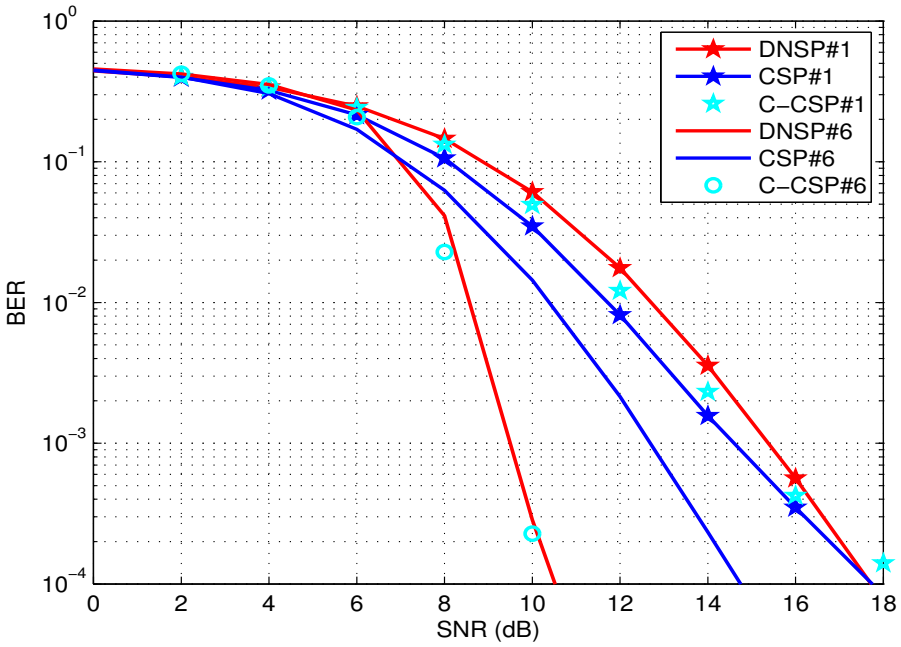


Figure 3.8: BER Performance for CSP, DNSP and C-CSP, $\rho = 0.125$, $L_f = 8$ #:number of iteration

3.6.2.1 Complexity Comparison

We still have to compute the complexity of each SP scheme so that we can have all needed elements to assess their relative performance. We will calculate the global computational complexity (channel estimation and IC) of each scheme. In fact, the channel estimation complexity is equivalent for the three schemes since starting from the second iteration, both pilots and estimated symbols are used for channel estimation for the three schemes. Except for a complexity saving for the first iteration in DNSP which only needs $8LP + P$ RM and $2(P-1) + 4LP$ RA, it is reasonable and fair to consider equivalent complexities since this will be dominated by the complexity of the subsequent iterations.

For the interference cancellation operation, the IC complexity for DNSP is given in table 3.1. For CSP, a classical MMSE equalizer is used and needs $9N$ RM and $6N - 2$ RA. Now, for C-CSP, although the *mismatched proposed IC* is used, however, the interference cancellation is performed over all sub-carriers

compared to only data sub-carriers in DNSP. In addition, we have the pilot removal, an operation which is not needed in the DNSP scheme and which adds $4N$ RM and $3N$ RA to the IC of C-CSP. We detail this comparison in table 3.2. We can observe that the CSP scheme is the least complex at the cost of performance lost. Also, the DNSP is less complex than the C-CSP scheme for the same BER performance.

Table 3.2: Complexity Comparison for different SP schemes

Operation	CSP	DNSP	C-CSP
RA	$6N-2$	$4ND+5D$	$4N^2+8N$
RM	$9N$	$6ND+13D+5$	$6N^2+17N+5$

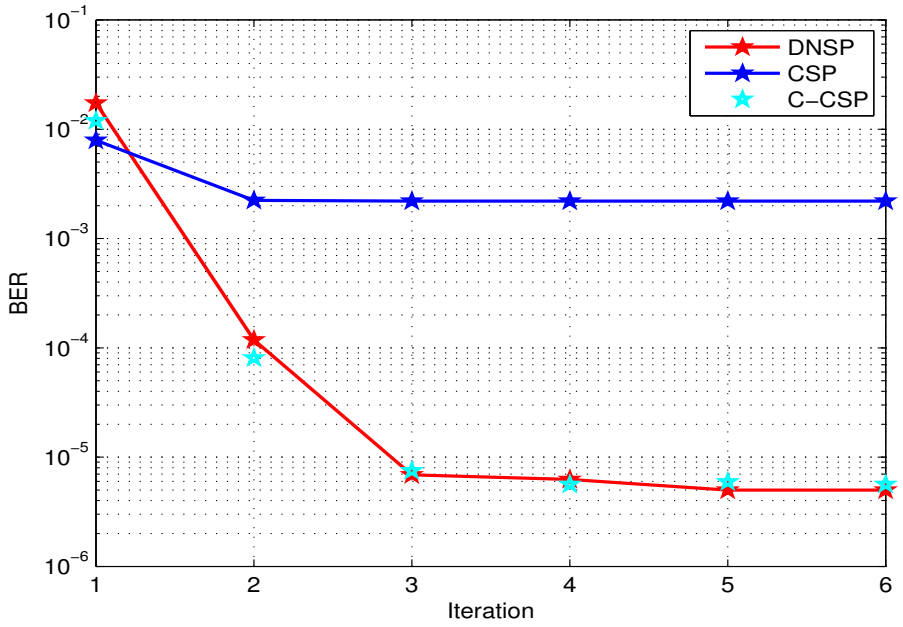


Figure 3.9: BER Performance for CSP, DNSP and C-CSP, $\rho = 0.125$, $L_f = 8$, $SNR = 12$

3.6.2.2 Convergence

Another important performance to study is the convergence of the three algorithms. We plot in figure 3.9 the BER of the CSP, C-CSP and the DNSP with respect to the number of iterations for $SNR = 12$ dB. The figure shows that the CSP converges after two iterations while both the DNSP and the C-CSP need 4 iterations to converge. This is because the DNSP and C-CSP schemes utilize the IC which converges with iterations while the CSP simply removes the known pilots and thus converges faster.

3.7 Conclusion

In this chapter, a well-suited receiver for the recently introduced DNSP scheme in OFDM has been proposed. In the framework of soft interference cancellation, an approximated formulation of an IC with a reduced complexity has been analyzed by exploiting the special interference structure that arises in this scheme, thus avoiding the matrix inversion needed in the classical case. We then perform a comparison between classical superimposed pilots scheme and DNSP in terms of channel estimation and decoding performance. It turns out that for a higher complexity, the DNSP outperforms the CSP scheme a gain of more than 3 dB gain at a BER target of 10^{-4}. In order to explain this difference, we propose to construct a C-CSP scheme which includes the inherent transmission diversity of DNSP through the precoding matrix $\mathbf{W}$. We observe that for a higher complexity, the C-CSP has an equivalent performance to DNSP. Thus, it is more profitable to introduce diversity to superimposed pilots using the DNSP scheme. In other words, we have shown that it is less complex to introduce transmission diversity in SP schemes, by using a precoding matrix, with the DNSP scheme than with the C-CSP scheme.

The answer to which scheme is better could not be easily answered. This is because each application might be concerned with a different performance indicator such as the complexity, the convergence rate, the channel estimation performance and the target BER. Thus, the preference of a scheme to another will highly depend on the application. In an application where a higher BER is tolerated, the CSP scheme might be preferred in return of a faster convergence and reduced complexity. In an application, where a lower BER is required, the DNSP with the *mismatched proposed IC* would be a more suitable choice at the expense of a higher complexity compared to CSP. The C-CSP, however, is the least interesting since its performance could be attained by the DNSP with less complexity.

Chapter 4

A Low Complexity Turbo Receiver for DNSP in OFDM with channel estimation errors

Contents

4.1 Introduction and Motivation

In chapter three, we proposed a low complexity IC for DNSP assuming perfect channel estimation at the receiver. We have seen that this assumption does not have an important effect on the performance of the IC. However, this is only true when the channel estimation error is negligible. To this end, we consider the case of a non-negligible channel estimation error and extend the IC of chapter 3 to take into account this error. An interesting framework to include the channel estimation error in the formulation of the IC has been proposed in [SKD09]. This approach makes use of the statistics of the channel as well as those of the channel estimation errors to integrate their effect into

the IC. This approach is referred to as the improved approach since it tends to improves the performance of the IC by integrating the channel estimation errors. Inspired by this work, we design an improved low complexity interference canceler for DNSP. In fact, in DNSP, the increase in the number of pilots does not necessarily enhance the BER performance as it does in the case of PSAM since the increase in the number of pilots in DNSP leads to the increase of interference on data. This is particularly true for high order modulations. We first derive the *improved exact IC* and the *improved proposed IC* for DNSP in sections 4.2 and 4.3 respectively. In section 4.5, we give the simulation results and discussions and then conclude this chapter in section 4.6.

4.2 *Improved exact IC* for DNSP

Let us first re-write the system model we used in chapter 3:

$$\mathbf{y}_{\mathbf{d},k} = \text{diag}\{\boldsymbol{\mathcal{H}}_{\mathbf{d},k}\}\mathbf{W}_{\mathbf{d}}\mathbf{s}_k + \mathbf{n}_{\mathbf{d},k} \tag{4.1}$$

$$= \mathbf{G}_{\mathbf{d},k}\mathbf{s}_k + \mathbf{n}_{\mathbf{d},k} \tag{4.2}$$

From (3.6), we have the LS channel estimate $\hat{\mathbf{h}}_k$:

$$\hat{\mathbf{h}}_k = \frac{1}{P}\mathbf{F}_{\mathbf{p}}^H \text{diag}\{\mathbf{c}_{\mathbf{p},k}\}^H \mathbf{y}_{\mathbf{p},k}.$$

Now, define the LS channel estimation error as $\Delta\mathbf{h}_k = \mathbf{h}_k - \hat{\mathbf{h}}_k$, then the covariance matrix of the LS channel estimation error $\mathbf{R}_{\Delta\text{h}}$, which will be used for the formulation of the IC in the next section is given below. We have:

$$\hat{\mathbf{h}}_k = \frac{1}{P}\mathbf{F}_{\mathbf{p}}^H \text{diag}\{\mathbf{c}_{\mathbf{p},k}\}^H \text{diag}\{\mathbf{c}_{\mathbf{p},k}\}\mathbf{F}_{\mathbf{p}}\mathbf{h} + \mathbf{n}_{\mathbf{p},k} \tag{4.3}$$

$$= \underbrace{\frac{1}{P}\mathbf{F}_{\mathbf{p}}^H \text{diag}\{\mathbf{c}_{\mathbf{p},k}\}^H \text{diag}\{\mathbf{c}_{\mathbf{p},k}\}\mathbf{F}_{\mathbf{p}}\mathbf{h}}_{\mathbf{h}_k} + \frac{1}{P}\mathbf{F}_{\mathbf{p}}^H \text{diag}\{\mathbf{c}_{\mathbf{p},k}\}^H \mathbf{n}_{\mathbf{p},k}$$

$$= \mathbf{h}_k + \frac{1}{P}\mathbf{F}_{\mathbf{p}}^H \text{diag}\{\mathbf{c}_{\mathbf{p},k}\}^H \mathbf{n}_{\mathbf{p},k}.$$

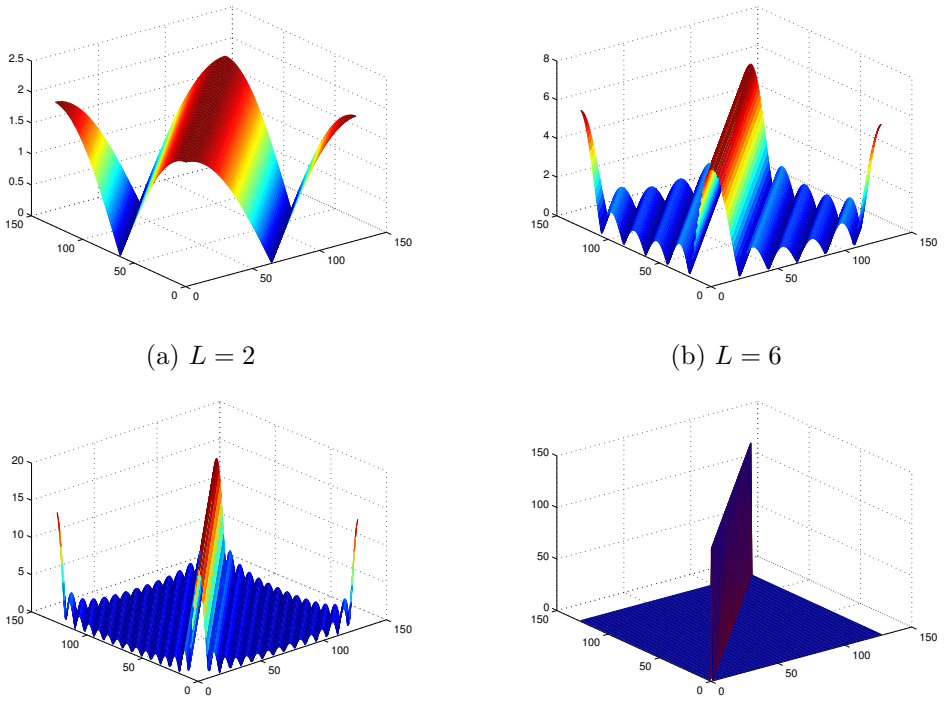
Thus, $\Delta\mathbf{h}_k = -\frac{1}{P}\mathbf{F}_{\mathbf{p}}^H \text{diag}\{\mathbf{c}_{\mathbf{p},k}\}^H \mathbf{n}_{\mathbf{p},k}$ and $\mathbf{R}_{\Delta\text{h}}$is given as:

$$\mathbf{R}_{\Delta\text{h}} = \mathbb{E}\{\Delta\mathbf{h}_k \Delta\mathbf{h}_k^H\}$$

$$= \frac{1}{P^2}\mathbb{E}\{\mathbf{F}_{\mathbf{p}}^H \text{diag}\{\mathbf{c}_{\mathbf{p}}\}^H \mathbf{n}_{\mathbf{p},k}\mathbf{n}_{\mathbf{p},k}^H \text{diag}\{\mathbf{c}_{\mathbf{p}}\}\mathbf{F}_{\mathbf{p}}\}$$

$$= \frac{\sigma_n^2}{P}\mathbf{I} \tag{4.4}$$

Clearly, the estimate of the frequency channel response $\mathbf{H}_k$ is calculated as: $\hat{\mathbf{H}}_k = \mathbf{F}\hat{\mathbf{h}}_k$.



(a) $L = 2$ (b) $L = 6$

(c) $L = 15$ (d) $L = 128$

Figure 4.1: Matrix $\mathbf{F}\mathbf{F}^H$ with different values of L, N=128.

In order to derive the improved IC for DNSP from (3.11), we proceed as in [SKD09]. For this, we need to obtain the posterior distribution of the true channel frequency response ($\mathbf{H}_k$) conditioned on its estimate ($\hat{\mathbf{H}}_k$), $p(\mathbf{H}_k|\hat{\mathbf{H}}_k)$.

For this, we need to find $p(\mathbf{h}_k|\hat{\mathbf{h}}_k)$. First, it is easy to see that $p(\hat{\mathbf{h}}_k|\mathbf{h}_k) = \mathcal{CN}(\mathbf{h}_k, \mathbf{R}_{\Delta \mathrm{h}})$ and $\mathbf{h}_k = \mathcal{CN}(\mathbf{0}, \mathbf{R}_{\mathrm{h}})$. However, to find $p(\mathbf{h}_k|\hat{\mathbf{h}}_k)$, we have to use the following theorem [SKD09]:

Theorem: Let $\mathbf{h}_1$ and $\mathbf{h}_2$ be circularly symmetric complex Gaussian ran-

dom vectors with zero means and known covariance matrices. Then $p(\mathbf{h}_1/\mathbf{h}_2) = \mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $\boldsymbol{\mu} = \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\mathbf{h}_2$ and $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$ where $\boldsymbol{\Sigma}_{ij} = \mathbb{E}[\mathbf{h}_i\mathbf{h}_j^H]$.

Now, $p(\mathbf{h}_k|\hat{\mathbf{h}}_k)$ is given as:

$$p(\mathbf{h}_k|\hat{\mathbf{h}}_k) = \mathcal{CN}(\mathbf{R}_\delta\hat{\mathbf{h}}_k, \mathbf{R}_\delta\mathbf{R}_{\Delta\mathrm{h}}), \tag{4.5}$$

where $\mathbf{R}_\delta = \mathbf{R}_\mathrm{h}(\mathbf{R}_{\Delta\mathrm{h}} + \mathbf{R}_\mathrm{h})^{-1}$ and $\mathbf{R}_\mathrm{h}$ is $L \times L$ diagonal matrix with $\sigma^2_{h_l}$ on its (l, l) th entry.

Then, $p(\boldsymbol{\mathcal{H}}_k|\hat{\boldsymbol{\mathcal{H}}}_k)$ could be expressed as ([BD99], page 56, proposition 5.2):

$$\begin{aligned} p(\boldsymbol{\mathcal{H}}_k|\hat{\boldsymbol{\mathcal{H}}}_k) &= \mathcal{CN}(\mathbf{F}\mathbf{R}_\delta\hat{\mathbf{h}}_k, \mathbf{F}\mathbf{R}_\delta\mathbf{R}_{\Delta\mathrm{h}}\mathbf{F}^H) \\ &= \mathcal{CN}(\mathbf{M}_k, \boldsymbol{\Sigma}_k). \end{aligned} \tag{4.6}$$

Then, exact improved IC for DNSP is thus given as follows and its detailed calculation in Appendix C:

$$\overline{\mathbf{w}}_{k,n} = \sigma_s^2[\mathrm{diag}\{\mathbf{M}_{\mathbf{d},k}\}\mathbf{W}_\mathbf{d}\mathbf{V}_{k,n}\mathbf{W}_\mathbf{d}^H\mathrm{diag}\{\mathbf{M}_{\mathbf{d},k}\}^H + \boldsymbol{\Omega}_k]^{-1}\mathrm{diag}\{\mathbf{M}_{\mathbf{d},k}\}\mathbf{W}_\mathbf{d}\mathbf{e}_n, \tag{4.7}$$

where $\boldsymbol{\Omega}_k = \sigma_n^2\mathbf{I} + \boldsymbol{\Sigma}_k \circ \left(\sum_{n=0}^{N-1} v_{k,n}.\mathbf{W}_n\mathbf{W}_n^H\right)$ with $\mathbf{W}_n$ being the nth column of $\mathbf{W}$ and $\circ$ is the Hadamard multiplication operation. Also, $\mathbf{M}_{\mathbf{d},k}$ is formed by selecting the data sub-carrier entries from $\mathbf{M}_k$. Note that when $\mathbf{R}_{\Delta\mathrm{h}}$ tends to zero, $\mathbf{R}_\delta$ tends to $\mathbf{I}$ and the equation of the exact improved IC in (4.7) boils down to the classical mismatched approach in (3.11).

4.3 *Improved proposed IC* for DNSP

For the same motivation as in chapter 3, we seek a low-complexity improved IC for DNSP. Here, however, replacing $\mathbf{V}_{k,n}$ by the approximation in (3.13) and applying the Woodbury identity does not render a single diagonal matrix inversion as in the case of the mismatched IC. This is due to the presence of the term $\boldsymbol{\Omega}_k$ which, is unlike in [SKD09], is a full matrix and depends on the sub-carrier index n. Matrix $\boldsymbol{\Sigma}_k$, however, has an interesting structure which we can exploit in order to overcome this problem. Now, since $\boldsymbol{\Sigma}_k = \mathbf{F}\mathbf{R}_\delta\mathbf{R}_{\Delta\mathrm{h}}\mathbf{F}^H$, let us start with the term $\mathbf{F}\mathbf{F}^H$. $\mathbf{F}$ is a the $N \times L$ Fourier matrix, *i.e.*, it is formed by choosing the first L columns of the full $N \times N$ Fourier matrix. This selection of the first L columns leads to the loss of the orthogonality of $\mathbf{F}$ and thus $\mathbf{F}\mathbf{F}^H \neq \mathbf{I}$. However, $\mathbf{F}\mathbf{F}^H$ admits a banded diagonal structure

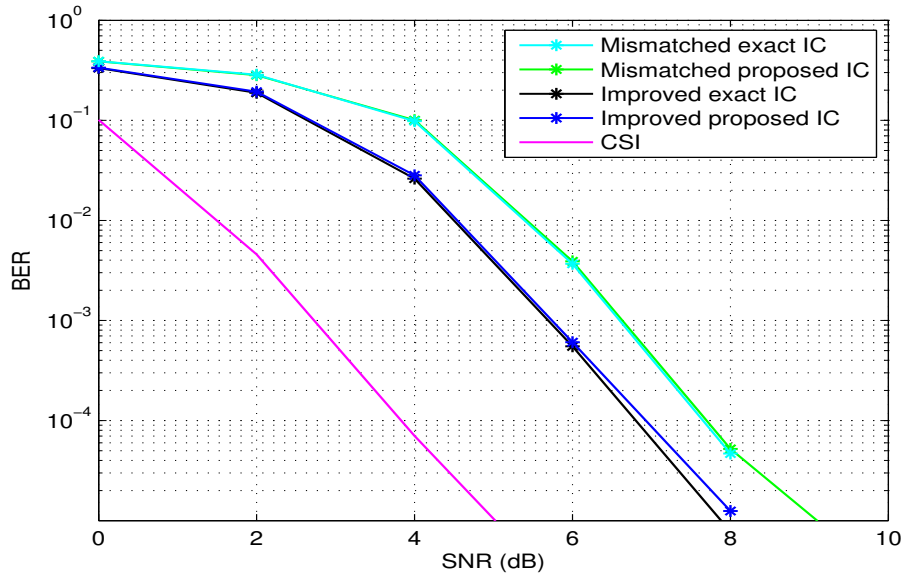


Figure 4.2: BER performance of the sixth iteration for 4QAM, $K = 2$, $L = 15$ dB, $L_f = 8$ for *mismatched exact IC*, *mismatched proposed IC*, *improved exact IC*, *improved proposed IC* and CSI

which decreases with L and tends towards $\mathbf{I}$. We show in figure 4.1 a typical behavior of $\mathbf{F}\mathbf{F}^H$ for different values of L. Now, the lth entry of the $L \times L$ diagonal matrix $\mathbf{R}_\delta \mathbf{R}_{\Delta \mathrm{h}}$ is $\frac{\sigma^2_{h_l}}{\frac{\sigma^2_n}{P}+\sigma^2_{h_l}} \frac{\sigma^2_n}{P}$. The value of σ^2_n decreases with SNR and the entries of $\mathbf{R}_\delta \mathbf{R}_{\Delta \mathrm{h}}$ tends to a unique value proportional to σ^2_n and thus the behavior of $\mathbf{\Sigma}_k$ becomes equivalent to that of $\mathbf{F}\mathbf{F}^H$. Thus, we can see that it is safe to ignore the off-diagonal entries of $\mathbf{\Sigma}_k$ as L increases and the SNR decreases. We will later show by simulations that this assumption does not incur any performance degradation on the proposed IC while allowing us to design a low-complexity IC.

Now, we can proceed as in chapter 3 to derive the approximated improved IC by replacing $\mathbf{V}_{k,n}$ by the approximation in (3.13). Then, the improved approximated IC is given equivalently as in (3.17) as:

$$\overline{\mathbf{w}}_{k,n} = \overline{\lambda}_{k,n} \overline{\tilde{\mathbf{w}}}_{k,n}, \tag{4.8}$$

with:

$$\overline{\tilde{\mathbf{w}}}_{k,n} = \sigma^2_s [\mathbf{M}_{\mathbf{d},k}(\sigma^2_s - \sigma^2_{\tilde{s}})\mathbf{M}^H_{\mathbf{d},k} + (\sigma^2_n + \sigma^2_a)\mathbf{I}]^{-1} \mathbf{M}_{\mathbf{d},k} \mathbf{W}_{\mathbf{d}} \mathbf{e}_n, \tag{4.9}$$

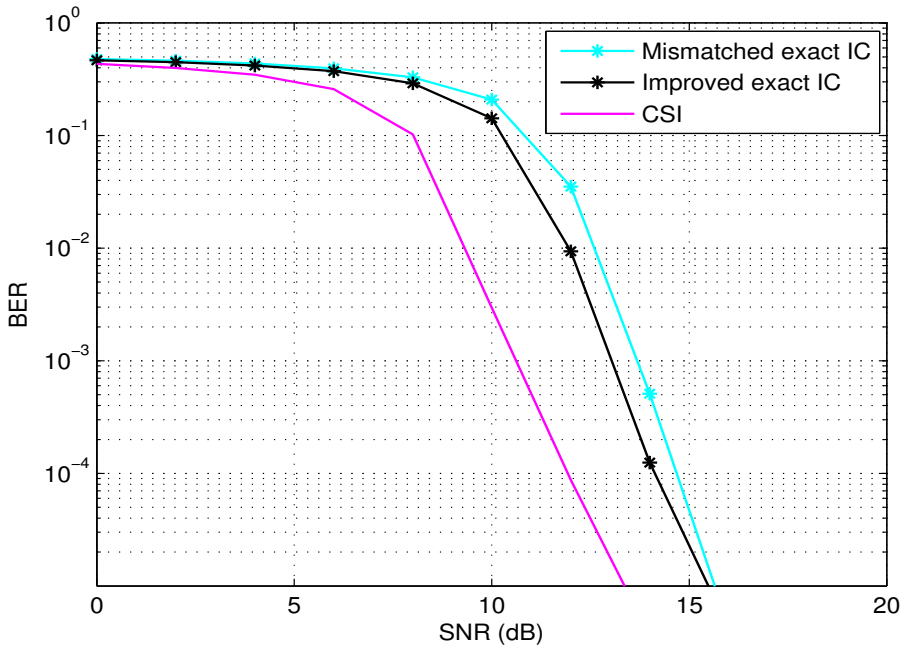


Figure 4.3: BER performance of the sixth iteration for 16QAM, $K = 3$, $L = 15$ dB, $L_f = 8$ for *mismatched exact IC*, *improved exact IC* and CSI

where $\sigma_a^2 = \mathrm{Tr}(\mathbf{R}_\delta \mathbf{R}_{\Delta \mathrm{h}}) \cdot \frac{(\sigma_s^2 - \sigma_{\tilde{s}}^2)N + \sigma_{\tilde{s}}^2}{N}$ and

$$\overline{\lambda}_{k,n} = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_{\tilde{s}}^2 \mathbf{e}_n^T \mathbf{W}_{\mathbf{d}}^H \mathbf{M}_{\mathbf{d},k}^H \overline{\mathbf{w}}_{k,n}}. \tag{4.10}$$

4.4 Improved Approximated IC of [SKD09]

We review in this section an approximated version of the improved IC proposed in [SKD09]. The idea behind this approximation is to assume perfect symbol estimation after the second iteration, *i.e.*, $v_{k,q} = 0$ for all $q \neq n$ and $v_{k,n} = \sigma_s^2$. Let us first define $\mathbf{A} = \mathrm{diag}\{\mathbf{M}_{\mathbf{d},k}\}\mathbf{W}_{\mathbf{d}}$ and $\mathbf{A}_n$ being the nth column of $\mathbf{A}$, then the approximated IC of [SKD09] for DNSP is formulated as:

$$\begin{aligned} \overline{\tilde{\mathbf{w}}}_{k,n} &= \sigma_s^2 [\mathbf{A}_n \sigma_s^2 \mathbf{A}_n^H + (\sigma_n^2 + \sigma_a^2)\mathbf{I}]^{-1} \mathbf{A}_n, \\ &= \frac{\sigma_s^2 \mathbf{A}_n^H}{\sigma_s^2 \mathbf{A}_n^H \mathbf{A}_n + (\sigma_n^2 + \sigma_a^2)} \end{aligned} \tag{4.11}$$

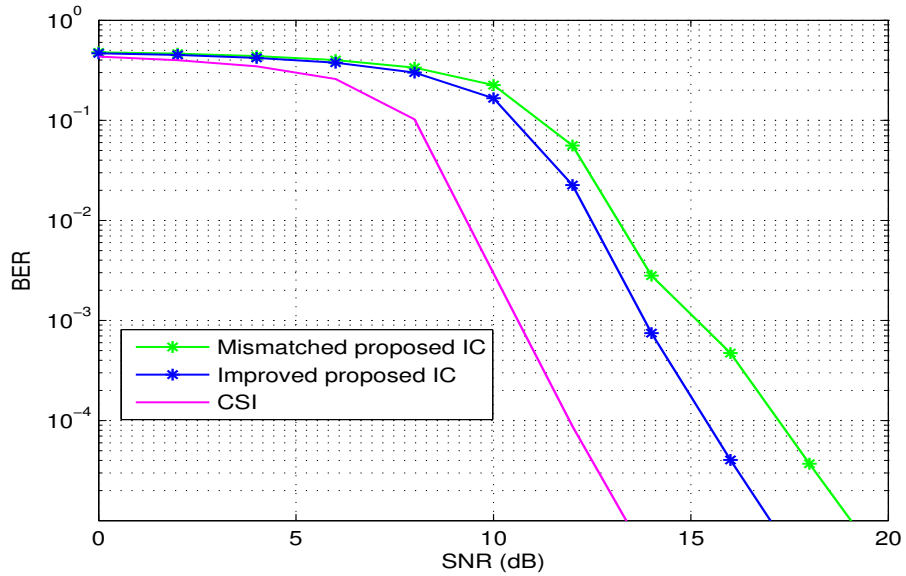


Figure 4.4: BER performance of the sixth iteration for 16QAM, $K = 3$, $L = 15$ dB, $L_f = 8$ for *mismatched proposed IC*, *improved proposed IC* and CSI

and $\sigma_a^2 = \frac{\sigma_s^2 \mathrm{Tr}(\mathbf{R}_\delta \mathbf{R}_{\Delta \mathrm{h}})}{N}$. We will refer to this IC as the *improved IC of [SKD09]*.

4.5 Simulation results and discussions

For the simulations in this chapter, we use the same parameters as in chapter 3 with a quasi-static channel. We first plot in figure 4.2, the BER performance of the sixth iteration, *i.e.*, at convergence, for a 4QAM modulation with $L = 15$, $L_f = 8$ for the *mismatched exact IC*, the *mismatched proposed IC*, the *improved exact IC* and the *improved proposed IC*. Note that the increase in the number of channel paths renders the channel estimation harder. We can see that the loss due to the approximation in both, the improved and the mismatched case, is negligible. This is in accordance with the results in chapter 3 for a 4QAM modulation. The gain from using the *improved* ICs at a target BER of 10^{-5} is around 1 dB. We then plot in figure 4.3, the BER performance of the *mismatched exact IC* and the *improved exact IC* for a 16QAM modulation. Here, at a target BER of 10^{-5}, the loss from the*mismatched exact IC* with respect to the improved optimal IC is negligible. This could be explained that with the increase of SNR, the effect of the channel estimation error decreases

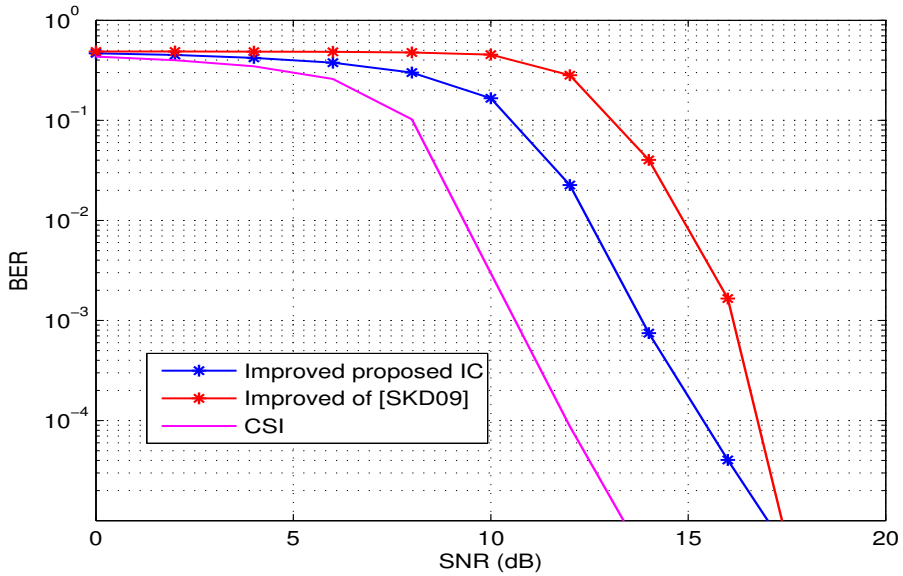


Figure 4.5: BER performance of the sixth iteration for 16QAM, $K = 3$, $L = 15$ dB, $L_f = 8$ for *improved proposed IC*, improved approximated of [SKD09] and CSI

and thus incurs less performance degradation by the *mismatched exact IC*. We then plot in figure 4.4, the BER performance of the *mismatched proposed IC* and the *improved proposed IC*. Here a gain of around 2.3 dB is obtained by using the *improved* version of the proposed IC. In fact, the proposed IC is more sensitive to the effect of channel estimation error. Then, we compare in figure 4.5, the BER performance of the *improved proposed IC* and that of the *improved IC of [SKD09]*. The results show that the gain from the *improved proposed IC* decreases at lower target BERs. This indicates that for applications with a target BER of 10^{-4}, a 1.55 dB gain is obtained by using the *improved proposed IC* compared to the *improved IC of [SKD09]*. This becomes less interesting for a lower target BER. Finally, we show in figure 4.6 that ignoring the non-diagonal elements of matrix $\mathbf{\Sigma}_k$ does not lead to any performance degradation in both the 4QAM and the 16QAM modulations.

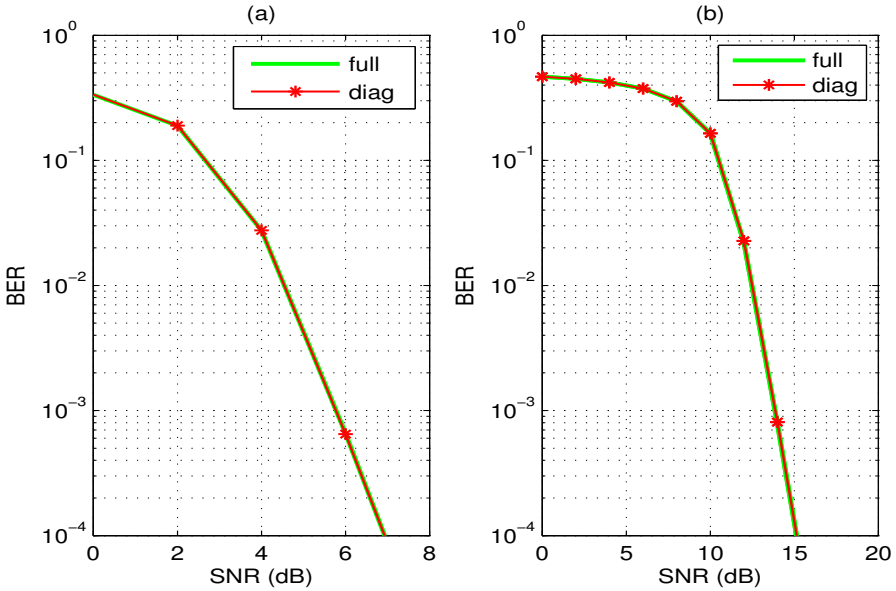


Figure 4.6: BER performance of the sixth iteration for the *improved proposed IC* using $\mathbf{\Sigma}_k$ denoted full and by using only the diagonal elements of $\mathbf{\Sigma}_k$ denoted diag (a) 4QAM (b) 16QAM

4.6 Conclusion

In this chapter, we extended the proposed IC of chapter 3 to include the channel estimation errors. We showed that the gain from using the improved proposed IC is more than 2 dB in a 16QAM modulation compared to 1 dB in a 4QAM modulation for a target BER of 10^{-5}. This is since the proposed IC is more sensitive to channel estimation errors in a 16QAM modulation. Those results lead us to the conclusion that the improved proposed IC is a good compromise between the complexity of the optimal ICs on the one hand and the performance loss due to neglecting the channel estimation errors on the other hand.

Chapter 5

Perspectives: On Robust Estimation with bounded data uncertainties

Contents

5.1 Introduction and Motivation

In the previous chapters we have seen that the presence of *uncertain* information at the OFDM receiver leads to the failure of conventional channel estimation algorithms. In particular, in the presence of NBI in chapter 2, the performance of the Kalman filter deteriorates since the noise variance is *unknown*. The interference in classical superimposed pilots led to the failure of the LS estimator in chapter 2 since a portion of the power had been attributed to the *unknown* data. This necessitated the design of more elaborate techniques to overcome those problems. In fact, this behavior of the conventional channel estimators is explained by the fact that, by design, they require the knowledge of certain underlying parameters. This knowledge varies from one estimator to another and it depends on its nature. The LS estimator for example requires the knowledge of the observation matrix. The Kalman filter necessitates, in

addition to the observation matrix, the formulation of a state equation with a known state matrix and the presence of the state and observation noise characteristics.

In this chapter, we investigate a different approach than those presented in the previous chapters, which aims at dealing with the interference as model uncertainties in the estimator's design. Our aim is to shed the light on the potentials and the underlying challenges of a bounded data uncertainty (BDU) class of robust estimators [SN99][Say01][SC02]. We also give insights about their applications to problems in OFDM with interference. We will refer to the BDU class as the *robust* approach. We note that the *robust* approach is a min-max approach where one aims at minimizing the cost function for the worst possible case.

This chapter is divided as follows. In sections 5.2 and 5.3, we briefly review the work by Sayed on the *robust* approach for the regularized LS in 5.2 and the Kalman filter in 5.3. In section 5.4, we provide our understanding on this approach and raise some questions and cast doubts on some assumptions related to the *robust* approach. Our goal is to try to get answers and give perspectives for future research directions. Finally, we conclude this chapter in section 5.5.

5.2 The regularized LS and its *robust* counterpart

The problem of ill-conditioned matrices which may arise in LS could be overcome by using the regularized LS (RLS). To understand the motivation behind this approach, let us consider the following linear model:

$$\mathbf{b} = \mathbf{A}\mathbf{x} + \mathbf{v} \tag{5.1}$$

where $\mathbf{b}$ is the $m \times 1$ observation vector, $\mathbf{A}$, the $m \times n$ observation matrix, $\mathbf{x}$, the $n \times 1$ vector and $\mathbf{v}$, the $m \times 1$ noise vector. Note that in this chapter, we use the notations in [SN99]. The cost function of the RLS is formulated as:

$$J(\mathbf{x}) = \mathbf{x}^H\mathbf{Q}\mathbf{x} + (\mathbf{A}\mathbf{x} - \mathbf{b})^H\mathbf{R}(\mathbf{A}\mathbf{x} - \mathbf{b}), \tag{5.2}$$

where $\mathbf{Q}$ and $\mathbf{R}$ are symmetric known matrices chosen by the user for regularization purposes. They might represent, for example, the covariance matrices of $\mathbf{x}$ and $\mathbf{v}$ respectively. The solution to this problem is obtained by minimizing the cost function in (5.2) and is given as:

$$\hat{\mathbf{x}} = [\mathbf{Q} + \mathbf{A}^H\mathbf{R}\mathbf{A}]^{-1}\mathbf{A}^H\mathbf{R}\mathbf{b}. \tag{5.3}$$

Note that for $\mathbf{Q} = \mathbf{0}$ and $\mathbf{R} = \mathbf{I}$, the formulation in (5.3) boils down to the conventional LS solution. This equation clearly reveals that the solution necessitates the knowledge of $\mathbf{A}$, $\mathbf{b}$.

Now, as per our discussion in section 5.1, only a partial knowledge of $\mathbf{A}$ and $\mathbf{b}$ is guaranteed. In this case, it is clear to see that if the deviation of $\mathbf{A}$ and $\mathbf{b}$ from their actual values is large, the solution $\hat{\mathbf{x}}$ will lead to a high error variance since it will be far from $\mathbf{x}$. This motivated the introduction of a new cost function to include the possible uncertainties [Say01][SN99] as:

$$J(\mathbf{x}, \mathbf{u}) = \mathbf{x}^H \mathbf{Q} \mathbf{x} + (\mathbf{A}\mathbf{x} - \mathbf{b} + \mathbf{T}\mathbf{u})^H \mathbf{R} (\mathbf{A}\mathbf{x} - \mathbf{b} + \mathbf{T}\mathbf{u}). \tag{5.4}$$

Now, the formulation of the cost function in (5.4) depends on two unknowns; $\mathbf{x}$ which is to be estimated and a new variable $\mathbf{u}$ which is the unknown perturbation vector. Comparing (5.2) and (5.4), we can see that the possible uncertainties in $\mathbf{A}$ and $\mathbf{b}$ are modeled in the term $\mathbf{T}\mathbf{u}$. The known matrix $\mathbf{T}$ allows the designer to choose a certain range space for the perturbation $\mathbf{u}$. While $\mathbf{u}$ is unknown, yet a bound on its norm is assumed to be known, *i.e.* $\mathbf{u} \leq \phi(x)$. Then the min-max *robust* solution of $\hat{\mathbf{x}}$ is given as:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \max_{\|\mathbf{u}\| \leq \phi(x)} J(\mathbf{x}, \mathbf{u}). \tag{5.5}$$

The formulation in (5.5) could be viewed as a two-player game where the designer tries to obtain the minimum $\mathbf{x}$ for the worst possible case of $\mathbf{u}$.

In order to be able to concretely understand the formulation in (5.5), we will consider a special case of uncertainties referred to as the structured uncertainties. If we define $\delta\mathbf{A}$ and $\delta\mathbf{b}$ as the unknown portions of $\mathbf{A}$ and $\mathbf{b}$, then in the structured uncertainties case, it is assumed that $\delta\mathbf{A}$ and $\delta\mathbf{b}$ satisfy the following model:

$$[\delta\mathbf{A} \ \delta\mathbf{b}] = \mathbf{T}\Delta[\mathbf{E}_a \ \mathbf{E}_b], \tag{5.6}$$

where $\mathbf{E}_a$ and $\mathbf{E}_b$ are known values which depend on the given problem, and Δ is any contraction matrix, *i.e.*$\|\Delta\| \leq 1$, in which the uncertainties are modeled. In other words, for a given problem, if a designer is able to model the uncertainties in $\mathbf{A}$ and $\mathbf{b}$ as the model in (5.6), then he should be able to formulate a *robust* solution to the problem.

Now, due to the unknown portions of $\mathbf{A}$ and $\mathbf{b}$, we can replace $\mathbf{A}$ and $\mathbf{b}$ in (5.2) by $\mathbf{A} + \delta\mathbf{A}$ and $\mathbf{b} + \delta\mathbf{b}$ respectively. Then, (5.4) could be re-written as:

$$J(\mathbf{x}, \mathbf{u}) = \mathbf{x}^H \mathbf{Q} \mathbf{x} + (\mathbf{A}\mathbf{x} - \mathbf{b} + (\delta\mathbf{A}\mathbf{x} - \delta\mathbf{b}))^H \mathbf{R} (\mathbf{A}\mathbf{x} - \mathbf{b} + (\delta\mathbf{A}\mathbf{x} - \delta\mathbf{b})), \tag{5.7}$$

and with $\mathbf{Tu} = (\delta\mathbf{Ax} - \delta\mathbf{b}) = \mathbf{T}(\Delta\mathbf{E}_a\mathbf{x} - \mathbf{E}_b)$, then $\mathbf{u}$ is defined as $\mathbf{u} = \Delta(\mathbf{E}_a\mathbf{x} - \mathbf{E}_b)$.

Then, after solving (5.5) which as we will see, conserves the structure of its conventional counterpart, where only the weighing matrices ($\mathbf{W}$ and $\mathbf{R}$) are replaced by their *corrected* versions ($\hat{\mathbf{W}}$ and $\hat{\mathbf{R}}$), the *robust* solution to (5.7) is given as [SN99][Say01]:

$$\hat{\mathbf{x}} = [\hat{\mathbf{Q}} + \mathbf{A}^H\hat{\mathbf{R}}\mathbf{A}]^{-1}\mathbf{A}^H\hat{\mathbf{Q}}\mathbf{b}. \tag{5.8}$$

where $\hat{\mathbf{Q}}$ and $\hat{\mathbf{R}}$ are given as:

$$\hat{\mathbf{Q}} = \mathbf{Q} + \hat{\lambda}\mathbf{E}_a^H\mathbf{E}_a; \qquad \hat{\mathbf{R}} = \mathbf{R} + \mathbf{RT}(\hat{\lambda}\mathbf{I} - \mathbf{T}^H\mathbf{RT})^{\dagger}\mathbf{T}^H\mathbf{R}. \tag{5.9}$$

We can see from (5.9) that in order to obtain the *corrected* weighing matrices ($\hat{\mathbf{Q}}$ and $\hat{\mathbf{R}}$), a scalar $\hat{\lambda}$ is to be determined. This is obtained through a numerical optimization problem as the minimum of a certain non-linear function $G(\lambda)$ as:

$$\hat{\lambda} = \arg \min_{\lambda \geq \|\mathbf{T}^H\mathbf{RT}\|} G(\lambda), \tag{5.10}$$

with $G(\lambda)$ defined as:

$$G(\lambda) = \|\mathbf{x}(\lambda)\|^2_{\mathbf{Q}(\lambda)} + \lambda\|\mathbf{E}_a\mathbf{x}(\lambda) - \mathbf{E}_b\|^2 + \|\mathbf{Ax}(\lambda) - \mathbf{b}\|^2_{\mathbf{R}(\lambda)} \tag{5.11}$$

with:

$$\mathbf{R}(\lambda) = \mathbf{R} + \mathbf{RT}(\lambda\mathbf{I} - \mathbf{T}^H\mathbf{RT})^{\dagger}\mathbf{T}^H\mathbf{R} \tag{5.12}$$

$$\mathbf{Q}(\lambda) = \mathbf{Q} + \lambda\mathbf{E}_a^H\mathbf{E}_a \tag{5.13}$$

$$\mathbf{x}(\lambda) = [\mathbf{Q}(\lambda) + \mathbf{A}^H\mathbf{R}(\lambda)\mathbf{A}]^{-1}[\mathbf{A}^H\mathbf{R}(\lambda)\mathbf{b} + \lambda\mathbf{E}_a^H\mathbf{E}_b]. \tag{5.14}$$

5.3 The Kalman filter and its *robust* counterpart

When the parameter to be estimated evolves with time according to a dynamical model, it becomes more reasonable to resort to the Kalman filter rather than to the RLS to perform estimation. The implementation of the Kalman filter requires, in addition to the observation equation, a state equation. The

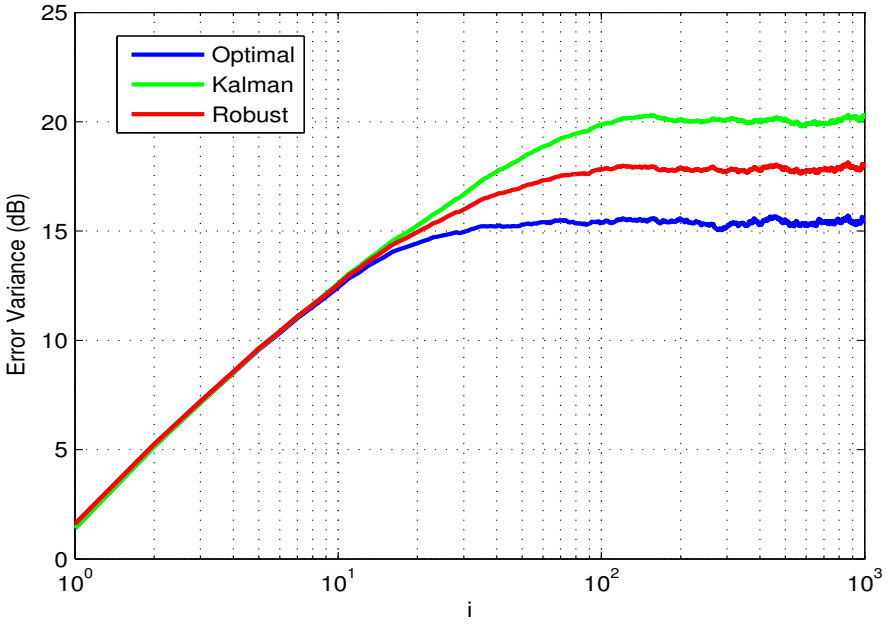


Figure 5.1: Error variance of three filters with Δ_i selected uniformly within the interval [-1,1], T(1)=0.0198, F(1,2)=0.0196

set of those two equations is known as the state space equations and is given as:

$$\mathbf{x}_{i+1} = \mathbf{F}\mathbf{x}_i + \mathbf{w}_i \tag{5.15}$$

$$\mathbf{b}_i = \mathbf{A}\mathbf{x}_i + \mathbf{v}_i \tag{5.16}$$

where (5.15) is the state equation and (5.16) is the observation equation with $\mathbf{F}$ being the state matrix, $\mathbf{G}$, the covariance matrix of the state noise $\mathbf{w}_i$, $\mathbf{A}$, the observation matrix, $\mathbf{R}$, the variance of the observation noise $\mathbf{v}_i$ and $\mathbf{Q}$ the covariance matrix of $\mathbf{x}_i$.

The Kalman filter estimates a process by using a form of feedback control. The filter estimates the process state at some time and then obtains feedback in the form of noisy measurements. This is done via the measurement and time update equations. In fact, those equations admit a deterministic interpretation as a solution to a RLS problem [Say01]. In other words, the measurement update and the time update equations of the Kalman filter could be derived by

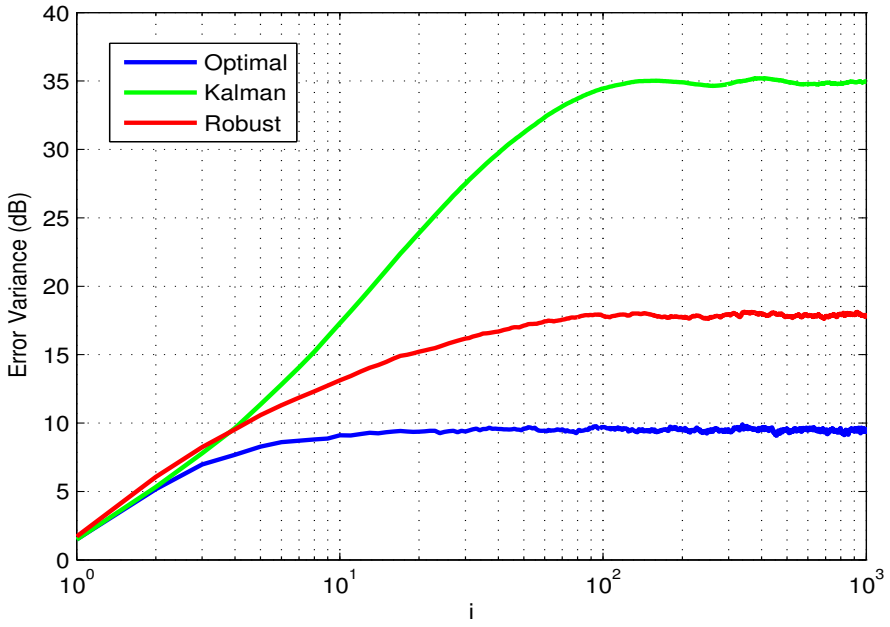


Figure 5.2: Error variance of three filters with Δ_i selected uniformly within the interval [-1,1], T(1)=0.198, F(1,2)=0.0196

posing a regularized least-squares problem and imposing some identifications. We give the details of this derivation in Appendix D. In fact, this allowed the author in [Say01] to extend the *robust* solution presented in section 5.2 for RLS to the *robust* Kalamn filter. The detailed formulations of the different forms of the *robust* Kalman filter could be found in [Say01].

In this section, our goal is to show the relative performance of the *robust* Kalman filter and its classical counterpart. We do that with a numerical example used in [Say01] to motivate the interest of the *robust* Kalman filter. This numerical example is also widely use in robust problems [PS99]. We will first recall the results provided by Sayed where he showed that by considering a structured uncertainty in the state matrix constructed as in (5.6), the steady-state performance of the *robust* Kalman filter in terms of the error variance is better than that of the Kalman filter.

Consider now the following state-space model with uncertainties:

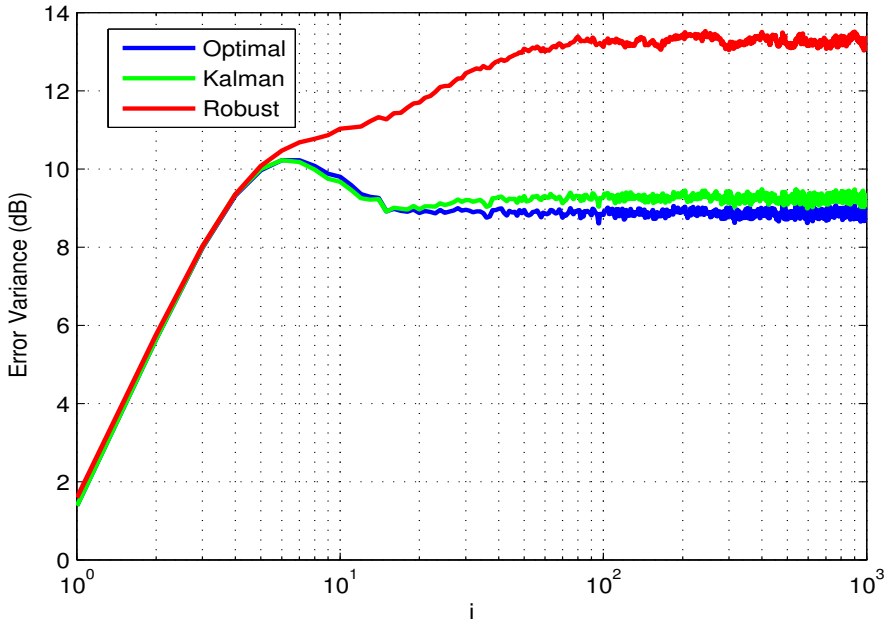


Figure 5.3: Error variance of three filters with Δ_i selected uniformly within the interval [-1,1], T(1)=0.0198, F(1,2)=0.3912

$$\mathbf{x}_{i+1} = (\mathbf{F} + \mathbf{T}\Delta_i\mathbf{E}_f)\mathbf{x}_i + \mathbf{w}_i \tag{5.17}$$

$$\mathbf{b}_i = \mathbf{A}\mathbf{x}_i + \mathbf{v}_i \tag{5.18}$$

Here, $\mathbf{E}_f$ and $\mathbf{T}$ represent the known parameters of the uncertainty given in (5.6). Notice also the dependence of Δ_i on the time index i. Let us now consider the numerical example of [Say01]:

$$\mathbf{F} = \begin{bmatrix} 0.9802 & 0.0196 \\ 0 & 0.9802 \end{bmatrix}; \qquad \mathbf{G} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix};$$

$$\mathbf{A} = \begin{bmatrix} 1 & -1 \end{bmatrix}; \qquad \mathbf{T} = \begin{bmatrix} 0.0198 \\ 0 \end{bmatrix}; \mathbf{R} = 1$$

$$\mathbf{E}_f = \begin{bmatrix} 0 & 5 \end{bmatrix}; \qquad \mathbf{Q} = \begin{bmatrix} 1.9608 & 0.0195 \\ 0.0195 & 1.9605 \end{bmatrix};$$

With this model, we have $\mathbf{F}$ at our disposal, whereas the actual state matrix ($\mathbf{F}_a = (\mathbf{F} + \mathbf{T}\Delta_i\mathbf{E}_f)$) is:

$$\mathbf{F}_a = \begin{bmatrix} 0.9802 & 0.0196 + 0.099\Delta_i \\ 0 & 0.9802 \end{bmatrix};$$

We plot the error variance of three filters; the Kalman filter with the knowledge of $\mathbf{F}$ which is referred to as Kalman in the figures, the *robust* Kalman filter referred to as Robust, and the optimal Kalman filter referred to as Optimal. More precisely, the Kalman filter uses the underlying model as if it were the actual one, the *robust* Kalman filter takes the uncertainties into account while designing the cost function of the estimator and the Optimal Kalman filter has the knowledge of the true matrix $\mathbf{F}_a$. First, we plot in figure 5.1, the error variance performance of the three filters with Δ_i selected uniformly within the interval [-1,1] and with the given parameters. We can see that the error variance of the *robust* filter is lower than that of the Kalman filter and closer to the performance of the optimal filter. The performance loss is around 2.5 dB at steady-state. The degradation in the performance of the Kalman filter is explained by the fact that it has been formulated with an erroneous model. Next, we keep the same parameters as for figure 5.1 and only change $\mathbf{T}(1) = 0.198$. $\mathbf{T}$ here represents the uncertainty in the model and increasing its value might be viewed as increasing the uncertainty in the system. We regenerate the same plots of figure 5.1 in figure 5.2. The results show that the performance degradation of the Kalman filter is more important than in the previous configuration and reaches around 17 dB at steady-state. This indicates that as the uncertainty in the model increases, the performance degradation of the Kalman filter with respect to both the optimal and the robust filter increases. Those two figures have been exposed in [Say01] to show the interest of the *robust* approach. While those results might be very promising, more recent work on this topic [XM09], showed that through this same numerical example, changing a parameter in the system matrix, renders the performance of the *robust* filter worse than the Kalman filter. We regenerate this example in figure 5.3 where the initial parameters are kept the same while replacing the (1,2)th entry of $\mathbf{F}$ by 0.3912. $\mathbf{F}$ represents the certainty in the system and increasing its value means increasing the certainty in the system. The figure shows that the performance of the *robust* Kalman filter deteriorates considerably and that of the Kalman filter approaches the performance of the optimal filter.

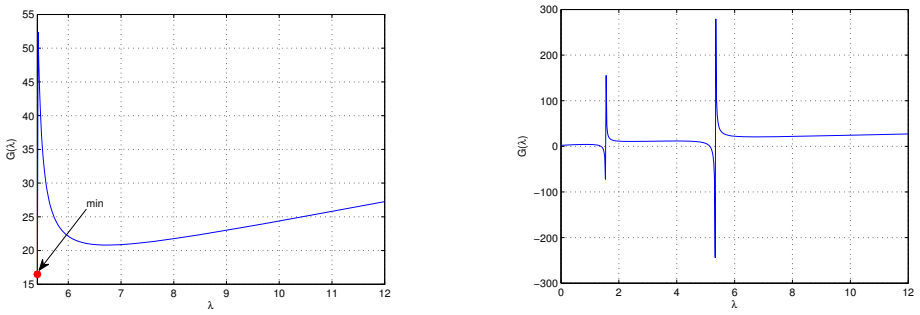
Those results are explained by the authors in [XM09] as the results of the conservative nature of the *robust* approach. In other words, they explain that

this change in the state matrix decreased the uncertainty and under such conditions, the *robust* approach becomes pessimistic and performs worse than the conventional filter. However, the authors do not precise how the *robust* approach is conservative.

We finally note that for obtaining parameter $\hat{\lambda}$, the authors in [Say01] assume that the function $G(\lambda)$ tends to reach amplitudes close to its minimum value at arguments of λ close to its lower bound ($\lambda_l = \|\mathbf{T}^H\mathbf{R}\mathbf{T}\|$). They consequently propose a practical approximation for $\hat{\lambda}$ as $\hat{\lambda} = (1+\alpha)\lambda_l$.

5.4 Ambiguities associated with the *robust* approach

We reserve this section for some ambiguous points associated with the *robust* approach. First, the results in section 5.3 suggests that this approach is efficient only under certain uncertainty conditions. In particular, the authors in [XM09] argue that the *robust* approach outperforms the conventional approach only when a high uncertainty is present in the system. They further relate this fact to the conservative nature of the *robust* approach.



(a) $G(\lambda)$ in interval $[\lambda_l, 12]$, (b) $G(\lambda)$ in interval $[0, 12]$

Figure 5.4: $G(\lambda)$

It would be interesting to investigate a *theoretical framework* that enables the designer to determine the uncertainty for a given problem.

The second point which is not thoroughly justified for the implementation of the *robust* approach is the choice of the parameter λ. We will first consider a numerical example provided by the authors in [SC02] and plot the function

$G(\lambda)$.

Given the following parameters:

$$\mathbf{A} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ -2 & 0 \\ 0 & 2 \end{bmatrix}; \qquad \mathbf{b} = \begin{bmatrix} -1 \\ 0 \\ 1 \\ -2 \end{bmatrix}; \qquad \mathbf{Q} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix};$$

$$\mathbf{E}_a = \begin{bmatrix} 1 & 0 \\ 0 & -2 \\ 1 & 1 \\ 0 & 0 \end{bmatrix}; \qquad \mathbf{R} = 1; \qquad \mathbf{E}_b = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}; \qquad \mathbf{T} = \begin{bmatrix} -1 & 0 \\ 1 & 2 \\ 0 & 0 \\ 1 & -1 \end{bmatrix};$$

we have $\lambda_l = \|\mathbf{T}^H\mathbf{R}\mathbf{T}\| = 5.4142$. We plot the G function for this problem for the interval $[\lambda_l, 12]$ in figure 5.4a and for the interval $[0, 12]$ in figure 5.4b. We notice that the minimum of $G(\lambda)$ within the interval $[\lambda_l, \infty[$ is λ_l itself. We notice that an important point (the lower bound of the interval) has been neglected by the authors. Although, the authors show in [Say01] that the approximation of this parameter has no effect on the performance of the algorithm, nevertheless, if we closely look into (5.8), we can see that in case the true λ lies on the lower bound, then the *robust* solution boils down to the conventional RLS solution with a new regularization parameter $\hat{\mathbf{Q}}$.

It would be also interesting to investigate this point in more details to understand its effect on the solution to the robust problem.

5.5 Conclusion

In this chapter, we reviewed a min-max *robust* approach which takes into account the model uncertainties in the design of the estimator. The attractive aspect of this approach is that it conserves the structure of its conventional counterparts. We have then recalled the state of the art and presented a numerical example which has been used by the authors in literature to motivate for this approach. This example shows the interesting potential of the *robust* approach, but also its *conservative* nature. We have raised some concerns regarding this approach which are open questions that provide an interesting perspective in this field. One very interesting framework to apply the *robust* approach to is the turbo reception framework. In fact, the *uncertainties* present

in turbo estimation due to the symbol estimation error is unknown but could be *bounded*. This can be done by making use of the variance of the symbol estimation error provided at each turbo iteration.

General conclusion and perspectives

A high quality channel estimation at the OFDM receiver should be maintained throughout the communication interval, so as to guarantee an acceptable system performance with the ultimate goal of coherent data detection. While designing channel estimation algorithms, a reasonable balance should be established between the system performance and the receiver's complexity on the one hand, as well as a trade-off is to be considered between the quality of the channel estimate and the spectral efficiency of the system on the other hand. In other words, a compromise should be done so as to maximize the spectral efficiency and reduce the complexity while maintaining an acceptable channel estimation performance. Those three criteria form the road map of our thesis.

Channel estimation in OFDM is often accomplished by inserting known pilot symbols within the OFDM frame. This is known as the pilot-symbol assisted modulation (PSAM). In order to be able to apply conventional channel estimation algorithms (LS, MMSE, ...) to those systems, a minimum number of pilots should be inserted with a minimal interference level. The first constraint which is the number of pilots, means that valuable bandwidth which could have been attributed to data is to be sacrificed. Consequently, this decreases the data rate which is in an ever-increasing demand. The second constraint of interference-free pilots can not be always met and interference could originate from different sources. A solution to the first problem has been envisioned. It aims at replacing conventional pilots (PSAM scheme) by superimposed pilots (SP). SP do not require a dedicated bandwidth, however, their presence leads to the rise in interference. In particular, we distinguish between two variants of SP; the classical SP (CSP) and the data-nulling SP (DNSP). In CSP, the interference arises from data on pilots in channel estimation and from pilots on data in data detection. In DNSP, interference on data originates from nulling

data sub-carriers after precoding for the insertion of pilot sub-carriers. Thus, we can see that by addressing the first problem, the second problem is worsened.

In a different context, where the initial goal was to optimize the spectral efficiency, the cognitive radio (CR) environment imposes the co-existence of technologies in the same frequency band. This may generate interference and even with PSAM, conventional channel estimation algorithms fail.

Thus, we dedicate this thesis to the design of channel estimation algorithms for OFDM systems in the presence of interference.

We organize this work in five chapters. Chapter one is dedicated to a brief introduction of the radio channel and the basic OFDM system. We distinguish between two types of channels in OFDM; the quasi-static channel and the fast time-varying channel. It is well known that in the case of high mobility, *i.e.* fast time-varying channel, the orthogonality between the sub-carriers of the OFDM system and thus the channel estimation approaches designed for quasi-static systems differ from those designed for fast time-varying systems. Further, we differentiate between two kinds of interference; narrowband interference (NBI) which exists in cognitive radio (CR) networks and interference from SP as discussed above.

In chapter two, we consider the problem of channel estimation in fast time-varying channels for OFDM systems contaminated by NBI. NBI strikes a small number of sub-carriers with a high power and renders conventional channel estimation algorithms inefficient. For the first time, we propose a channel estimation framework in this context. This has been done via the MAP EM algorithm. The problem is formulated so as to elegantly integrate the Kalman smoother for the estimation of the channel taps. We then simulate the proposed algorithm in a real world scenario and show its robustness to mobility in terms of MSE and BER performance, in contrast to existing algorithms which have not been designed to handle mobility. In addition, this is obtained at the cost of a reasonable additional complexity.

In chapters three and four, we consider the problem of channel estimation and data detection in DNSP. DNSP is a recently proposed scheme for OFDM and the design of a dedicated receiver is yet to be addressed. Thus we propose a receiver design for DNSP based on turbo reception. Turbo reception is well-known to be the most efficient in interference cancellation. However, this

efficiency comes at the cost of high complexity and the necessity of a precise channel estimate.

In chapter three, we deal with the complexity issue and propose a low-complexity receiver for DNSP by exploiting the special interference structure in this scheme. The receiver is based on the MMSE soft interference canceler (IC). Then, we initiate a performance comparison between DNSP with the proposed receiver and the CSP scheme in an iterative receiver. We first show that an important gain in terms of BER is achieved by DNSP compared to classical superimposed pilots (CSP) at the cost of a higher complexity. This is explained by the increased transmission diversity due to the precoding, which is part of the DNSP design. We then show that a coded-CSP scheme is able to achieve the same BER performance as DNSP but this time at the cost of a higher complexity. From this study we conclude that increasing the transmission diversity in SP schemes is more profitable in terms of complexity via the DNSP scheme.

The performance of the proposed IC for DNSP in chapter three is only reliable when the channel estimation error is small. Although DNSP offers by design interference-free pilots, yet in some scenarios the channel estimate might not be reliable. We thus design in chapter four an enhanced IC for DNSP which takes the channel estimation errors into account. The results show that the enhanced IC is particularly interesting in scenarios with a high channel estimation error. In addition, the gain in BER is more interesting in 16QAM which could reach up to 2.3 dB compared to 1 dB in 4QAM at a target BER of 10^{-5} since the IC is more sensitive to channel estimation errors for higher order modulations. However, we should point out that in DNSP and unlike in PSAM, the increase in the number of pilots does not necessarily improve the BER since a higher number of pilots leads to more interference.

In the last chapter, we shed the light on a min-max *robust* approach. We first expose the *robust* counterparts of the regularized least-squares and the Kalman filter. We then present some interrogations about this approach and present a perspective for its possible application.

Finally, the work presented in this thesis has an interesting potential for future research:

- Perspectives on chapter two:

  In chapter two, we used a Gaussian model for NBI. It would be interesting to obtain more accurate models for NBI through measurement

campaigns. This will help give a clearer idea about the effect of this model on the performance of the algorithm. In fact, a new channel sounder has been designed at TELICE and which is planned to be used in future measurement campaigns.

Despite the advantages of the OFDM system, which led to its standardization in 4G systems, its documented drawbacks should be addressed while choosing an appropriate modulation for 5G systems. The universal filtered multi-carrier (UFMC) and the generalized frequency division multiplexing (GFDM) are the two serious candidates for 5G systems. UFMC is very close in nature to OFDM and is referred to as the universal filtered OFDM (UF-OFDM) as well. Unlike filtered OFDM which applies a filter to each sub-carrier and filter bank based multi-carrier (FBMC) which applies a filter to the entire band, the UFMC applies filters on a per band basis thus preserving the advantages of OFDM while reducing the effect of high spectral lobes in OFDM. This property could be very interesting in rendering the system more robust to interference. GFDM is also a generalized form of OFDM. However, unlike OFDM which uses one cyclic prefix (CP) per symbol, it uses only one CP for a group of symbols. Recent studies point out that GFDM is well suited to cognitive radio (CR), since the rectangular pulse used in OFDM is replaced by a choice of pulse shaping. This lowers the side lobes in the frequency domain. As a perspective, the study of the impact of NBI in UFMC and GFDM systems which are, by design, robust to interference, could further give insights about the advantage of one system to the other from an interference point of view.

- Perspectives on chapters three and four:

  In chapters 2 and three of this thesis, we saw the different variants of superimposed pilot (SP) schemes proposed for OFDM so as to replace the bandwidth-consuming conventional pilots. The channel estimation task and pilot placement schemes are also important aspects to be studied in UFMC and GFDM systems. UFMC, as already stated, preserves the OFDM structure and thus the extension of the channel estimation and the pilot schemes to UFMC should be direct forward. However, in GFDM, since the data now is transmitted in the time domain and then frequency domain equalization (FDE) is performed at the receiver, the transmitted data are mixed up, the insertion of conventional pilots as in OFDM becomes inefficient. The study of particular pilot schemes for GFDM is an interesting challenge and an open perspective. The work

presented in chapters 3 and 4 of this thesis could be the starting point of this work.

- Perspectives on chapter five:

  The further research on *robust* channel estimators to problems of channel estimation with interference in OFDM is an open topic. An interesting area of research could be to consider uncertainties in the covariance matrix of the channel noise rather than in the system matrices. This could be an interesting application to the environments with NBI. We note here that we have already considered the application of the H_∞ filter to the NBI problem. However, the results were disappointing although the H_∞ filter is designed without restrictions on the distribution or the knowledge of the system's noise. Another challenging aspect is to provide a theoretical framework which explains the underlying functionality of the robust approaches.

Appendix A

Calculations related to chapter two

A.1 Computation of the Q Function

In order to obtain the expression for $Q(\boldsymbol{\sigma^2}, \hat{\boldsymbol{\sigma^2}}^{(i)})$, we take the expectation with respect to $\mathbf{c}$ conditioned on $\mathbf{y}$, given the current parameter estimate $\hat{\boldsymbol{\sigma^2}}^{(i)}$. We have from (2.19):

$$Q(\boldsymbol{\sigma^2}, \hat{\boldsymbol{\sigma^2}}^{(i)}) = -K \ln \prod_n \sigma_n^2 + \sum_{k=0}^{K-1} \mathrm{Tr} \left\{ \mathbb{E}_{\mathbf{c}|\mathbf{y}, \hat{\boldsymbol{\sigma^2}}^{(i)}} \left[(\mathbf{y}_k - \boldsymbol{\mathcal{X}}_k \mathbf{c}_k)^H \mathbf{C}^{-1} (\mathbf{y}_k - \boldsymbol{\mathcal{X}}_k \mathbf{c}_k) \right] \right\} \tag{A.1}$$

From (A.1), we obtain:

$$Q(\boldsymbol{\sigma^2}, \hat{\boldsymbol{\sigma^2}}^{(i)}) = \mathrm{Tr} \left\{ \mathbf{C}^{-1} (\mathbf{y}_k \mathbf{y}_k^H + \boldsymbol{\mathcal{X}}_k \mathbf{S}_{k|K}^{(i)} \boldsymbol{\mathcal{X}}_k^H - \mathbf{y}_k \hat{\mathbf{c}}_{k|K}^{(i)H} \boldsymbol{\mathcal{X}}_k^H - \boldsymbol{\mathcal{X}}_k \hat{\mathbf{c}}_{k|K}^{(i)} \mathbf{y}_k^H) \right\} \tag{A.2}$$

where $\mathbf{S}_{k|K}^{(i)} \overset{\text{def}}{=} \mathbb{E}_{\mathbf{c}|\mathbf{y}, \hat{\boldsymbol{\sigma^2}}^{(i)}} \left[\mathbf{c}_k \mathbf{c}_k^H \right]$ and $\hat{\mathbf{c}}_{k|K}^{(i)} \overset{\text{def}}{=} \mathbb{E}_{\mathbf{c}|\mathbf{y}, \hat{\boldsymbol{\sigma^2}}^{(i)}} [\mathbf{c}_k]$. Let us define:

$$\begin{aligned} \mathbf{P}_{k|K}^{(i)} &\overset{\text{def}}{=} \mathbb{E}_{\mathbf{c}|\mathbf{y}, \hat{\boldsymbol{\sigma^2}}^{(i)}} \left[(\mathbf{c}_k - \hat{\mathbf{c}}_{k|K}^{(i)}) (\mathbf{c}_k - \hat{\mathbf{c}}_{k|K}^{(i)})^H \right] \\ &= \mathbf{S}_{k|K}^{(i)} - \hat{\mathbf{c}}_{k|K}^{(i)} \hat{\mathbf{c}}_{k|K}^{(i)H} \end{aligned} \tag{A.3}$$

Then by substituting (D.23) into A.2, we obtain (2.23).

A.2 Calculation of $\hat{\sigma_n^2}^{(i+1)}$

We provide here the calculation of $\hat{\sigma_n^2}^{(i+1)}$ in equation (2.24).

E-step:

$$\begin{aligned}
\hat{\sigma_n^2}^{(i+1)} &= \arg\max_{\sigma_n^2}\{\, Q(\boldsymbol{\sigma^2}, \hat{\boldsymbol{\sigma^2}}^{(i)}) + \ln p(\boldsymbol{\sigma^2}) \\
&= -K\ln\sum_{n=0}^{N-1}\sigma_n^2 - \sum_{k=0}^{K-1}\mathrm{Tr}\big\{\mathbf{C}^{-1}(\boldsymbol{\mathcal{X}}_k\mathbf{P}_{k|K}^{(i)}\boldsymbol{\mathcal{X}}_k^H + (\mathbf{y}_k - \boldsymbol{\mathcal{X}}_k\hat{\mathbf{c}}_{k|K}^{(i)})(\mathbf{y}_k - \boldsymbol{\mathcal{X}}_k\hat{\mathbf{c}}_{k|K}^{(i)})^H\big\} \\
&+ \ln[\frac{\lambda}{\sigma_n^4}\exp\{\frac{-\lambda}{\sigma_n^2}\}] \\
&= -K\ln\sum_{n=0}^{N-1}\sigma_n^2 - \sum_{k=0}^{K-1}\mathrm{Tr}\big\{\mathbf{C}^{-1}(\boldsymbol{\mathcal{X}}_k\mathbf{P}_{k|K}^{(i)}\boldsymbol{\mathcal{X}}_k^H + (\mathbf{y}_k - \boldsymbol{\mathcal{X}}_k\hat{\mathbf{c}}_{k|K}^{(i)})(\mathbf{y}_k - \boldsymbol{\mathcal{X}}_k\hat{\mathbf{c}}_{k|K}^{(i)})^H\big\} \\
&+ \ln\lambda - 2\ln\sigma_n^2 - \frac{\lambda}{\sigma_n^2}
\end{aligned}$$

M-step:

Now, we derive with respect σ_n^2 and equate to zero.

Define $\mathbf{M}_k = (\boldsymbol{\mathcal{X}}_k\mathbf{P}_{k|K}^{(i)}\boldsymbol{\mathcal{X}}_k^H + (\mathbf{y}_k - \boldsymbol{\mathcal{X}}_k\hat{\mathbf{c}}_{k|K}^{(i)})(\mathbf{y}_k - \boldsymbol{\mathcal{X}}_k\hat{\mathbf{c}}_{k|K}^{(i)})^H$, we have:

$$\begin{aligned}
\frac{-K}{\sigma_n^2} - \sum_{k=0}^{k=K-1} -\frac{1}{\sigma_n^4}[\mathbf{M}_k]_{n,n} - \frac{2}{\sigma_n^2} + \frac{\lambda}{\sigma_n^4} &= 0 \\
-\frac{K}{-\sigma_n^2} - \frac{2}{\sigma_n^2} &= -\sum_{k=0}^{K-1}\frac{1}{\sigma_n^4}[\mathbf{M}_k]_{[}n,n] - \frac{\lambda}{\sigma_n^4} \\
K+2 &= \sigma_n^2\left[\sum_{k=0}^{K-1}\frac{1}{\sigma_n^4}[\mathbf{M}_k]_{n,n} + \frac{\lambda}{\sigma_n^4}\right] \\
K+2 &= \sum_{k=0}^{K-1}\frac{1}{\sigma_n^2}[\mathbf{M}_k]_{n,n} + \frac{\lambda}{\sigma_n^2}
\end{aligned}$$

Then,

$$\hat{\sigma_n^2}^{(i+1)} = \frac{1}{K+2}\sum_{k=0}^{K-1}[\mathbf{M}_k]_{n,n} + \frac{\lambda}{K+2} \tag{A.4}$$

A.3 Additional complexity calculation for proposed algorithm with respect to algorithm in [HR10]

In order to calculate the additional complexity required by the proposed algorithm, we have to calculate:

$$\hat{\sigma_n^2}^{(i+1)} = \underbrace{\frac{1}{K+2}}_{O1} \underbrace{\sum_{k=0}^{K-1}}_{O2} \underbrace{[\mathbf{M}_k]_{n,n}}_{O3} + \underbrace{\frac{\lambda}{K+2}}_{O4}. \tag{A.5}$$

$$\text{and } \mathbf{M}_k = \underbrace{\underbrace{\boldsymbol{\mathcal{X}}_k \mathbf{P}_{k/K}^{(i)} \boldsymbol{\mathcal{X}}_k^H}_{O3c} + \underbrace{\underbrace{(\mathbf{y}_k - \boldsymbol{\mathcal{X}}_k \hat{\mathbf{c}}_{k|K}^{(i)})}_{O3a} (\mathbf{y}_k - \boldsymbol{\mathcal{X}}_k \hat{\mathbf{c}}_{k|K}^{(i)})^H}_{O3b}}_{O3d}.$$

For calculating $O1$, 1 RM and 1 RA is required. $O4$ requires only 1 RM since the addition has been already done in $O1$. Now, for the calculation of $O3$ which we divide into sub-operations, calculating $\mathcal{X}_k \hat{\mathbf{c}}_{k|K}^{(i)}$ in $O3a$ needs NLN_c CM and $N(LN_c - 1)$ CA and then N CA for the subtraction from $\mathbf{y}_k$. Operation $O3b$ results in an $N \times N$ Hermitian matrix. Thus the real diagonal entries will need 2 RM and 1 RA each, and the non-diagonal entries will require $N^2 - N$ CM. However, we notice that for the calculation of (A.5), only the diagonal entries of $\mathbf{M}_k$ are required and thus it is enough to compute only those entries. Thus the complexity of this step is reduced to $2N$ RM and N RA. For operation

Table A.1: Additional Complexity of the proposed algorithm

Operation	RA	RM	CA	CM
$O1$	1	1	-	-
$O2$	$N(K-1)$	-	-	-
$O3a$	-	-	NLN_c	NLN_c
$O3b$	N	$2N$	-	-
$O3c$	-	-	$NLN_c^2 - N$	$NLN_c^2 + NLN_c$
$O3d$	N	-	-	-
$O4$	–	1	-	-
Total	$NK + N + 1$	$2N + 2$	$NLN_c^2 + NLN_c - N$	$NLN_c^2 + 2NLN_c$

$O3c$, $NLN_c^2 + NLN_c$ CM and $NLN_c(N_c - 1) + N(LN_c - 1)$ CA are needed to compute the diagonal entries of $O3c$. $O3d$ will thus require N RA since the addition of the diagonal entries of $O3b$ and $O3c$ needs a single RA per sub-carrier. Then to finalize this calculation, operation $O2$ requires $(K-1)$ RA for each sub-carrier and thus a total of $N(K-1)$ RA. This is summarized in table A.1.

Appendix B

Calculations related to chapter three

B.1 Demonstration of equation (3.17) from (3.11)

$$\mathbf{w}_{k,n} = \sigma_s^2[\mathbf{G}_{\mathbf{d},k}\mathbf{V}_{k,n}\mathbf{G}_{\mathbf{d},k}^H + \sigma_n^2\mathbf{I}]^{-1}\mathbf{G}_{\mathbf{d},k}\mathbf{e}_n, \tag{B.1}$$

Now using the approximation (3.13), $\mathbf{V}_{k,n}$ becomes:

$$\mathbf{V}_{k,n} = \begin{bmatrix} \sigma_s^2 - \sigma_{\bar{s}}^2 & 0 & \dots & \dots & 0 \\ \vdots & \ddots & & & \\ \vdots & & \sigma_s^2 & & \\ \vdots & & & \ddots & \\ 0 & & & & \sigma_s^2 - \sigma_{\bar{s}}^2 \end{bmatrix}$$

Then (B.1) can be re-written as:

$$\mathbf{w}_{k,n} = \sigma_s^2(\mathbf{G}_{\mathbf{d},k}[(\sigma_s^2 - \sigma_{\bar{s}}^2)\ \mathbf{I} + \sigma_s^2\mathbf{e}_n\mathbf{e}_n^T]\mathbf{G}_{\mathbf{d},k}^H + \sigma_n^2\mathbf{I})^{-1}\mathbf{G}_{\mathbf{d},k}\mathbf{e}_n. \tag{B.2}$$

By using Woodbury identity defined as:

$$(\mathbf{A} + \mathbf{u}\mathbf{v}^H)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^H\mathbf{A}^{-1}}{1 + \mathbf{v}^H\mathbf{A}^{-1}\mathbf{u}} \tag{B.3}$$

where $\mathbf{A} = \text{diag}\{\hat{\boldsymbol{\mathcal{H}}}_{\mathbf{d},k}\}(\sigma_s^2 - \sigma_{\bar{s}}^2)\text{diag}\{\hat{\boldsymbol{\mathcal{H}}}_{\mathbf{d},k}\}^H + \sigma_n^2\mathbf{I}$, $\mathbf{u} = \sigma_{\bar{s}}^2\text{diag}\{\hat{\boldsymbol{\mathcal{H}}}_{\mathbf{d},k}\}\mathbf{W}_{\mathbf{d}}\mathbf{e}_n$ and $\mathbf{v}^H = \mathbf{e}_n^T\mathbf{W}_{\mathbf{d}}^H\text{diag}\{\hat{\boldsymbol{\mathcal{H}}}_{\mathbf{d},k}\}^H$, we have:

$$\mathbf{w}_{k,n} = \sigma_s^2\left[\mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}uv^H\mathbf{A}^{-1}}{1 + v^H\mathbf{A}^{-1}u}\right]\mathbf{G}_{\mathbf{d},k}\mathbf{e}_n. \tag{B.4}$$

Since $\mathbf{A}$ is hermitian and $u = \sigma_{\bar{s}}^2 v$, then (B.4) can be written as:

$$\mathbf{w}_{k,n} = \sigma_s^2 \left[\frac{\mathbf{A}^{-1}\mathbf{G}_{\mathbf{d},k}\mathbf{e}_n}{1 + \sigma_{\bar{s}}^2 \mathbf{e}_n^T \mathbf{G}_{\mathbf{d},k}^H \mathbf{A}^{-1} \mathbf{G}_{\mathbf{de}_n}} \right] \mathbf{G}_{\mathbf{d},k}\mathbf{e}_n. \tag{B.5}$$

Now, defining $\tilde{\mathbf{w}}_{k,n} = \sigma_s^2 \mathbf{A}^{-1}\mathbf{G}_{\mathbf{d},k}\mathbf{e}_n$, then (B.5) becomes:

$$\frac{\mathbf{w}_{k,n} = \sigma_s^2 \tilde{\mathbf{w}}_{k,n}}{\sigma_s^2 + \sigma_{\bar{s}}^2 \mathbf{e}_n^T \mathbf{G}_{\mathbf{d},k}^H \tilde{\mathbf{w}}_{k,n}}, \tag{B.6}$$

which could be written as:

$$\mathbf{w}_{k,n} = \lambda_{k,n} \tilde{\mathbf{w}}_{k,n}, \tag{B.7}$$

where $\lambda_{k,n}$

$$\lambda_{k,n} = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_{\bar{s}}^2 \mathbf{e}_n^T \mathbf{W}_{\mathbf{d}}^H \mathrm{diag}\{\hat{\boldsymbol{\mathcal{H}}}_{\mathbf{d},k}\}^H \tilde{\mathbf{w}}_{k,n}}. \tag{B.8}$$

The only matrix inversion required now is $\mathbf{A}^{-1}$ and which is a diagonal matrix inversion due to the property of $\mathbf{W}$ ($\mathbf{W}_{\mathbf{d}}\mathbf{W}_{\mathbf{d}}^H = \mathbf{I}$).

B.2 Complexity Computation of the *mismatched exact IC* and the *mismatched proposed IC*

We detail here the given complexities of the three ICs given in Table 3.1. We first note that a CM is equivalent to 4 RM and 2 RA and a CA is equivalent to 2 RA.

B.2.1 Complexity of *mismatched exact IC*

Define $O1 = \mathrm{diag}\{\hat{\boldsymbol{\mathcal{H}}}_{\mathbf{d},k}\}\mathbf{W}_{\mathbf{d}}$, then the equalization vector for the *optimal IC* is given as:

$$\mathbf{w}_{k,n} = \overbrace{\sigma_s^2 \underbrace{\underbrace{[\underbrace{O1\mathbf{V}_{k,n}O1^H}_{O2} + \sigma_n^2\mathbf{I}]}_{O3}{}^{-1} O1\mathbf{e}_n}_{O4}}^{O5} \tag{B.9}$$

To accomplish operation O1, we need $2ND$ RM. Its Hermitian has no additional complexity. O2 results in a $D \times D$ matrix. The calculation of its diagonal

entries require 3 RM and 1 RA each, thus a total of $3D$ RM + $(2D-1)$ RA, its D^2-D non diagonal entries need 1 CM, 2 RM and $(D-1)$ CA each, thus a total of (D^2-D)(1 CM + 2 RM + $(D-1)$ CA). O3 and O4 need D RA and D^3 CM respectively. Operation $O5$ could be written as: $\sigma_s^2 O4 O1As$ where $O1As = O1Ae_{nO1As}$. O1As is a selection operation and does not add to the complexity. O4 O1As results in a $D \times 1$ vector which needs D CM and (D-1) CA per entry and thus a total of $D(D$ CM + $(D-1)$ CA), finally the multiplication with σ_s^2 needs $2D$ RM.

Now the filtering operation is given as follows:

$$\tilde{s}_{k,n} = \overbrace{\underbrace{\mathbf{w}_{k,n}{}^H (\mathbf{y}_{\mathbf{d},k} - \underbrace{O1B\bar{\mathbf{s}}_k}_{O6A})}_{O6B}}^{O6C} \tag{B.10}$$

$O6A$ results in a $D \times 1$ vector and each entry needs N CM and $(N-1)$ CA, thus a total of $D(N$ CM + $(N-1)$ CA), D CA for $O6B$. $O6C$ results in a scalar with D CM + $(D-1)$ CA.

B.2.2 Complexity of *mismatched proposed IC*

The equalization vector for the *proposed IC* is given as:

$$\tilde{\mathbf{w}}_{k,n} = \underbrace{\sigma_s^2 \overbrace{[\underbrace{\underbrace{\mathrm{diag}\{\hat{\boldsymbol{\mathcal{H}}}_{\mathbf{d},k}\}(\sigma_s^2 - \sigma_{\bar{s}}^2)\mathrm{diag}\{\hat{\boldsymbol{\mathcal{H}}}_{\mathbf{d},k}\}^H}_{O2} + \sigma_n^2 \mathbf{I}}_{O3}]^{-1}}^{O4} O1\mathbf{e}_n}_{O6} \tag{B.11}$$

First, $O1 = \mathrm{diag}\{\hat{\boldsymbol{\mathcal{H}}}_{\mathbf{d},k}\}\mathbf{W}_{\mathbf{d}}$ needs $2ND$ RM. To accomplish operation $O2$ which results in a $D \times D$ real diagonal matrix, we need 3 RM and 1 RA for each entry, for a total of D(3 RM + 1 RA). $O3$ is the addition of 2 real diagonal matrices and thus needs D RA. Finally, operation $O6$ could be written as: σ_s^2 O4 O1 e_n, O1 e_n is a selection operation. O4 O1 e_n is the multiplication of a real diagonal matrix and a complex vector which results in a $D \times 1$ vector and needs 2 RM per entry so $2D$ RM in total. The last operation which is the multiplication by σ_s^2 also requires $2D$ RM.

$$O7 = \lambda_k = \overbrace{\frac{\sigma_s^2}{\underbrace{\sigma_s^2 + \underbrace{\sigma_{\bar{s}}^2 \underbrace{\mathbf{e}_n^T \mathbf{W}_{\mathbf{d}}^H \mathrm{diag}\{\hat{\boldsymbol{\mathcal{H}}}_{k\mathbf{d},k}\}^H \tilde{\mathbf{w}}_{k,n}}_{O7A}}_{O7B}}_{O7C}}^{O7D} \tag{B.12}$$

$O7A$ is the selection of the (n,n) entry of an $N \times N$ matrix and needs 3 RM + 1 RA, $O7B$ needs 1 RM, $O7C$ 1 RA and $O7D$ 1 RM.

Now the filtering operation is given as:

$$\tilde{s}_{k,n} = \overbrace{\underbrace{\lambda_k \tilde{\mathbf{w}}_{k,n}}_{O8C} \overbrace{(\mathbf{y}_{\mathbf{d}} - \underbrace{\mathrm{diag}\{\hat{\boldsymbol{\mathcal{H}}}_{\mathbf{d},k}\}\mathbf{W}_{\mathbf{d}}\bar{\mathbf{s}}_k}_{O8A})}^{O8B}}^{O8D}$$

$O8A$ results in a $D \times 1$ vector which needs $D(N$ CM $+ (N-1)$ CA), $O8B$ needs D CA, $O8C$ needs $2D$ RM, and $O8D$ results in a scalar that necessitates D CM $+ (D-1)$ CA.

Appendix C

Calculations related to chapter four

We demonstrate here the derivation of the improved IC for DNSP.

The channel impulse response $\mathbf{h}_k$, the covariance matrix $\mathbf{R}_h$ of $\mathbf{h}_k$ and the covariance matrix of the LS channel estimation error $\mathbf{R}_{\Delta \mathrm{h}}$ ar given as:

$$\mathbf{h}_k = \begin{bmatrix} h_0 \\ \vdots \\ h_{L-1} \end{bmatrix}; \qquad \hat{\mathbf{h}}_k = \begin{bmatrix} \hat{h}_0 \\ \vdots \\ \hat{h}_{L-1} \end{bmatrix}$$

$$\mathbf{R}_h \begin{bmatrix} \sigma^2_{h_0} & & \\ & \ddots & \\ & & \sigma^2_{h_{L-1}} \end{bmatrix}; \qquad \mathbf{R}_{\Delta \mathrm{h}} = \begin{bmatrix} \frac{\sigma_n^2}{P} & & \\ & \ddots & \\ & & \frac{\sigma_n^2}{P} \end{bmatrix}$$

Then:

$$p(\mathbf{h}_k|\hat{\mathbf{h}}_k) = \mathcal{CN}(\mathbf{R}_\delta \hat{\mathbf{h}}_k, \mathbf{R}_\delta \mathbf{R}_{\Delta \mathrm{h}}), \tag{C.1}$$

where $\mathbf{R}_\delta = \mathbf{R}_{\mathrm{h}}(\mathbf{R}_{\Delta \mathrm{h}} + \mathbf{R}_{\mathrm{h}})^{-1}$ [SKD09].

Now, $p(\mathbf{H}_k|\hat{\mathbf{H}}_k)$ could be expressed as ([BD99], page 56, proposition 5.2):

$$\begin{aligned} p(\mathbf{H}_k|\hat{\mathbf{H}}_k) &= \mathcal{CN}(\mathbf{F}\mathbf{R}_\delta \hat{\mathbf{h}}_k, \mathbf{F}\mathbf{R}_\delta \mathbf{R}_{\Delta \mathrm{h}} \mathbf{F}^H) \\ &= \mathcal{CN}(\mathbf{M}_k, \mathbf{\Sigma}_k). \end{aligned} \tag{C.2}$$

. Let us recall the equation of the exact mismatched IC for DNSP as:

$$\mathbf{w}_{k,n} = \sigma_s^2 \underbrace{[\mathbf{G}_{\mathbf{d},k} \mathbf{V}_{k,n} \mathbf{G}^H_{\mathbf{d},k} + \sigma_n^2 \mathbf{I}]}_{\mathbf{R}_\mathbf{y}}{}^{-1} \underbrace{\mathbf{G}_{\mathbf{d},k} \mathbf{e}_n}_{\mathbf{R}_{\mathbf{y},s}}. \tag{C.3}$$

Now, the improved approximated IC is calculated as:

$$\overline{\mathbf{w}}_{k,n} = \sigma_s^2 \overline{\mathbf{R}}_{\mathbf{y}}^{-1} \overline{\mathbf{R}}_{\mathbf{y},s}, \tag{C.4}$$

where $\overline{\mathbf{R}}_{\mathbf{y}} = \mathbb{E}_{\mathbf{H}_k|\hat{\mathbf{H}}_k}\{\mathbf{R}_{\mathbf{y}}\}$ and $\overline{\mathbf{R}}_{\mathbf{y},s} = \mathbb{E}_{\mathbf{H}_k|\hat{\mathbf{H}}_k}\{\mathbf{R}_{\mathbf{y},s}\}$.

Now, we need to calculate $\mathbb{E}_{\mathbf{H}_k|\hat{\mathbf{H}}_k}\{\mathbf{G}_{\mathbf{d},k}\mathbf{V}_{k,n}\mathbf{G}_{\mathbf{d},k}^H\}$ and $\mathbb{E}_{\mathbf{H}_k|\hat{\mathbf{H}}_k}\{\mathbf{G}_{\mathbf{d},k}\mathbf{e}_n\}$.

$$\mathbb{E}_{\mathbf{H}_k|\hat{\mathbf{H}}_k}\{\mathbf{G}_{\mathbf{d},k}\mathbf{V}_{k,n}\mathbf{G}_{\mathbf{d},k}^H\} = \mathbb{E}_{\mathbf{H}_k|\hat{\mathbf{H}}_k}\left\{\sum_{n=0}^{N-1} v_n \mathbf{g}_n \mathbf{g}_n^H\right\} \tag{C.5}$$

$$= \sum_{n=0}^{N-1}\left\{\mathbb{E}_{\mathbf{H}_k|\hat{\mathbf{H}}_k} v_n \mathbf{g}_n \mathbf{g}_n^H\right\} \tag{C.6}$$

where $\mathbf{g}_n$ is the nth column of matrix $\mathbf{G}_{\mathbf{d},k}$ where we drop the index k to simplify notations.

$$\mathbb{E}_{\mathbf{H}_k|\hat{\mathbf{H}}_k}\{v_n \mathbf{g}_n \mathbf{g}_n^H\} = v_n \mathbb{E}_{\mathbf{H}_k|\hat{\mathbf{H}}_k}\{\mathbf{g}_n \mathbf{g}_n^H\} \tag{C.7}$$

and

$$\mathbb{E}_{\mathbf{H}_k|\hat{\mathbf{H}}_k}\{\mathbf{g}_n \mathbf{g}_n^H\} = \mathbf{A}_n \mathbf{A}_n^H + (\mathbf{\Sigma}_k \circ \mathbf{W}_n). \tag{C.8}$$

where $\mathbf{A} = \text{diag}\{\mathbf{M}_{\mathbf{d},k}\}\mathbf{W}_{\mathbf{d}}$, $\mathbf{W}_n$ is the nth column of $\mathbf{W}_{\mathbf{d}}$ and $\circ$ is the Hadamard multiplication operation.

Then,

$$\mathbb{E}_{\mathbf{H}_k|\hat{\mathbf{H}}_k}\{\mathbf{G}_{\mathbf{d},k}\mathbf{V}_{k,n}\mathbf{G}_{\mathbf{d},k}^H\} = \mathbf{A}\mathbf{V}_{k,n}\mathbf{A}^H + \mathbf{\Sigma}_k \circ \left(\sum_{n=0}^{N-1} v_n \mathbf{W}_n \mathbf{W}_n^H\right). \tag{C.9}$$

And,

$$\mathbb{E}_{\mathbf{H}_k|\hat{\mathbf{H}}_k}\{\mathbf{G}_{\mathbf{d},k}\mathbf{e}_n\} = \mathbf{A}\mathbf{e}_n. \tag{C.10}$$

Finally, replacing (C.9) and (C.10) into (C.4), we obtain the formulation of the exact improved IC for DNSP.

We still have the approximated improved IC for DNSP which is derived following the same procedure as in Appendix B.

Appendix D

Calculations related to chapter 5

Consider the following state-space model at instant i:

$$\mathbf{x}_{i+1} = \mathbf{F}_i \mathbf{x}_i + \mathbf{G}_i \mathbf{w}_i \tag{D.1}$$

$$\mathbf{b}_i = \mathbf{A}_i \mathbf{x}_i + \mathbf{v}_i \tag{D.2}$$

We also define the following matrix identities which will be useful in our derivation:

$$(\mathbf{A} + \mathbf{BDC})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{D}^{-1} + \mathbf{CA}^{-1}\mathbf{B})^{-1}\mathbf{CA}^{-1} \tag{D.3}$$

$$\mathbf{DC}(\mathbf{A} + \mathbf{BDC})^{-1} = (\mathbf{D}^{-1} + \mathbf{CA}^{-1}\mathbf{B})^{-1}\mathbf{CA}^{-1} \tag{D.4}$$

$$(\mathbf{A} + \mathbf{UCV})^{-1} = \mathbf{A}^{-1}\mathbf{U}(\mathbf{C}^{-1} + \mathbf{VA}^{-1}\mathbf{U})^{-1}\mathbf{VA}^{-1} \tag{D.5}$$

Then after receiving a new measurement $\mathbf{b}_{i+1}$, $\mathbf{x}_i$ is estimated by solving the following equation by minimizing with respect to $\mathbf{x}_i$ and $\mathbf{x}_{i+1}$:

$$\mathbb{J} = \min_{\mathbf{x}_i, \mathbf{w}_i} [\|\mathbf{x}_i - \hat{\mathbf{x}}_{i/i}\|^2_{\mathbf{P}_{i/i}{}^{-1}} + \|\mathbf{w}_i\|^2_{\mathbf{Q}_i{}^{-1}} + \|\mathbf{b}_{i+1} - \mathbf{A}_{i+1}\mathbf{x}_{i+1}\|^2_{\mathbf{R}_{i+1}{}^{-1}}]. \tag{D.6}$$

We have that the differential of a quadratic form is given by:

$$d(\mathbf{x}^T\mathbf{Q}^{-1}\mathbf{x}) = d(\mathbf{x}^T)\mathbf{Q}^{-1}\mathbf{x} + \mathbf{x}^T\mathbf{Q}^{-1}d(\mathbf{x}) = 2d(\mathbf{x}^T)\mathbf{Q}^{-1}\mathbf{x} \tag{D.7}$$

Then, derivatives of $\mathbb{J}$ w.r.t $\mathbf{x}_i$ and $\mathbf{x}_{i+1}$ are respectively given as:

$$\frac{d\mathbb{J}}{\mathbf{x}_i} = -\mathbf{P}^{-1}_{i/i}(\hat{\mathbf{x}}_{i/i} - \mathbf{x}_i) - \mathbf{F}_i^T\mathbf{Q}_i^{-1}(\mathbf{x}_{i+1} - \mathbf{F}_i\mathbf{x}_i), \tag{D.8}$$

$$\frac{d\mathbb{J}}{\mathbf{x}_{i+1}} = \mathbf{Q}_i^{-1}(\mathbf{x}_{i+1} - \mathbf{F}_i\mathbf{x}_i) - \mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}(\mathbf{b}_{i+1} - \mathbf{A}_{i+1}\mathbf{x}_{i+1}) \tag{D.9}$$

Now, **b**y equating (D.8) and (D.9) to zero, we o**b**tain the following system of equations:

$$(\mathbf{P}_{i/i}^{-1} + \mathbf{F}_i^T\mathbf{Q}_i^{-1}\mathbf{F}_i)\mathbf{x}_i = \mathbf{P}_{i/i}^{-1}\hat{x}_{i/i} + \mathbf{F}_i^T\mathbf{Q}_i^{-1}\mathbf{x}_{i+1}, \tag{D.10}$$

$$(\mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}\mathbf{A}_{i+1} + \mathbf{Q}_i^{-1})\mathbf{x}_{i+1} = \mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}\mathbf{b}_{i+1} + \mathbf{Q}_i^{-1}\mathbf{F}_i\mathbf{x}_i. \tag{D.11}$$

Su**b**stituting the value of $\mathbf{x}_i$ from equation (D.10) into (D.11), we o**b**tain:

$$\left[\mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}\mathbf{A}_{i+1} + \mathbf{Q}_i^{-1} - \mathbf{Q}_i^{-1}\mathbf{F}_i(\mathbf{F}_i^T\mathbf{Q}_i^{-1}\mathbf{F}_i + \mathbf{P}_{i/i}^{-1})^{-1}\mathbf{F}_i^T\mathbf{Q}_i^{-1}\right]\mathbf{x}_{i+1} = \mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}\mathbf{b}_{i+1} + \mathbf{Q}_i^{-1}\mathbf{F}_i(\mathbf{F}_i^T\mathbf{Q}_i^{-1}\mathbf{F}_i + \mathbf{P}_{i/i}^{-1})^{-1}\mathbf{P}_{i/i}^{-1}\hat{\mathbf{x}}_{i/i} \tag{D.12}$$

Now, by using the (D.3) with identifications, $\mathbf{A} = \mathbf{Q}_i$, $\mathbf{B} = \mathbf{F}_i$, $\mathbf{C} = \mathbf{F}_i^T$ and $\mathbf{D} = \mathbf{P}_{i/i}$, we get $\mathbf{Q}_i^{-1} - \mathbf{Q}_i^{-1}\mathbf{F}_i(\mathbf{F}_i\mathbf{Q}_i^{-1}\mathbf{F}_i + \mathbf{P}_{i/i}^{-1})^{-1}\mathbf{F}_i^T\mathbf{Q}_i^{-1} = (\mathbf{Q}_i + \mathbf{F}_i + \mathbf{P}_{i/i}\mathbf{F}_i^T)$. After defining $\mathbf{P}_{i+1} = \mathbf{Q}_i + \mathbf{F}_i\mathbf{P}_{i/i}\mathbf{F}_i^T$, (D.12) could be re-written as:

$$[\mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}\mathbf{A}_{i+1} + \mathbf{P}_{i+1}]\mathbf{x}_{i+1} = \mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}\mathbf{b}_{i+1} + \mathbf{Q}_i^{-1}\mathbf{F}_i(\mathbf{P}_{i/i}^{-1} + \mathbf{F}_i^T\mathbf{Q}_i^{-1}\mathbf{F}_i)^{-1}\mathbf{P}_{i/i}^{-1}\hat{\mathbf{x}}_{i/i}. \tag{D.13}$$

Then, by using (D.4) with identifications, $\mathbf{A} = \mathbf{P}_{i/i}^{-1}$, $\mathbf{B} = \mathbf{F}_i^T$, $\mathbf{C} = \mathbf{F}_i$ and $\mathbf{D} = \mathbf{Q}_i^{-1}$ we get $\mathbf{Q}_i^{-1}\mathbf{F}_i(\mathbf{P}_{i/i}^{-1} + \mathbf{F}_i^T\mathbf{Q}_i^{-1}\mathbf{F}_i)^{-1} = (\mathbf{Q}_i^{-1} + \mathbf{F}_i\mathbf{P}_{i/i}\mathbf{F}_i^T)^{-1}\mathbf{F}_i\mathbf{P}_{i/i}$. Then (D.13) is re-written as:

$$[\mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}\mathbf{A}_{i+1} + \mathbf{P}_{i+1}^{-1}]\mathbf{x}_{i+1} = \mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}\mathbf{b}_{i+1} + \mathbf{P}_{i+1}^{-1}\mathbf{F}_i\hat{\mathbf{x}}_{i/i}. \tag{D.14}$$

Now, $\hat{\mathbf{x}}_{i+1/i+1}$ is written as:

$$\hat{\mathbf{x}}_{i+1/i+1} = [\mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}\mathbf{A}_{i+1} + \mathbf{P}_{i+1}^{-1}]^{-1}\mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}\mathbf{b}_{i+1} + [\mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}\mathbf{A}_{i+1} + \mathbf{P}_{i+1}^{-1}]\mathbf{P}_{i+1}^{-1}\mathbf{F}_i\hat{\mathbf{x}}_{i/i}. \tag{D.15}$$

Now using the matrix identity (D.4) with identifications, $\mathbf{A} = \mathbf{R}_{i+1}$, $\mathbf{B} = \mathbf{A}_{i+1}$, $\mathbf{C} = \mathbf{A}_{i+1}^T$ and $\mathbf{D} = \mathbf{P}_{i+1}$, we have $[\mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}\mathbf{P}_{i+1}^{-1}]^{-1}\mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1} = \mathbf{R}_{i+1}\mathbf{R}_{i+1}[\mathbf{R}_{i+1} + \mathbf{A}_{i+1}\mathbf{P}_{i+1}\mathbf{A}_{i+1}^T]$. After defining $\mathbf{G}_{i+1} = \mathbf{P}_{i+1}\mathbf{A}_{i+1}^T\underbrace{[\mathbf{R}_{i+1} + \mathbf{A}_{i+1}\mathbf{P}_{i+1}\mathbf{A}_{i+1}^T]^{-1}}_{\mathbf{R}_{e,i+1}}$, (D.15) **b**ecomes:

$$\hat{\mathbf{x}}_{i+1/i+1} = \mathbf{G}_{i+1}\mathbf{b}_{i+1} + [\mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}\mathbf{A}_{i+1}]\mathbf{P}_{i+1}^{-1}\mathbf{F}_i\hat{\mathbf{x}}_{i/i}. \tag{D.16}$$

We use (D.3) again with identifications $\mathbf{A} = \mathbf{P}_{i+1}^T$, $\mathbf{B} = \mathbf{A}_{i+1}^T$, $\mathbf{C} = \mathbf{A}_{i+1}$ and $\mathbf{D} = \mathbf{R}_{i+1}^{-1}$, we have:

$$[\mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}\mathbf{A}_{i+1}+\mathbf{P}_{i+1}^{-1}] = \mathbf{P}_{i+1}-\mathbf{P}_{i+1}\mathbf{A}_{i+1}^T[\mathbf{R}_{i+1}^{-1}+\mathbf{A}_{i+1}\mathbf{P}_{i+1}^{-1}\mathbf{A}_{i+1}^T]\mathbf{A}_{i+1}\mathbf{P}_{i+1}$$
$$= \mathbf{P}_{i+1}-\mathbf{G}_{i+1}\mathbf{A}_{i+1}\mathbf{P}_{i+1} = (\mathbf{I}-\mathbf{P}_{i+1}\mathbf{A}_{i+1})\mathbf{P}_{i+1}. \tag{D.17}$$

Then (D.16) is re-written as:

$$\hat{\mathbf{x}}_{i+1/i+1} = \mathbf{G}_{i+1}\mathbf{b}_{i+1}+(\mathbf{I}-\mathbf{G}_{i+1}\mathbf{A}_{i+1})\mathbf{F}_i\hat{\mathbf{x}}_{i/i} \tag{D.18}$$

Now using (D.5), we have:

$$\mathbf{P}_{i/i} = \mathbf{P}_i-\mathbf{P}_i\mathbf{A}_i^T(\mathbf{R}_i+\mathbf{A}_i\mathbf{P}_i\mathbf{A}_i^T)^{-1}\mathbf{A}_i\mathbf{P}_i \tag{D.19}$$
$$= (\mathbf{P}_i^{-1}+\mathbf{A}_i^T\mathbf{R}_i^{-1}\mathbf{A}_i)^{-1} \tag{D.20}$$

Now:

$$\mathbf{P}_{i+1}\mathbf{A}_{i+1}^T\mathbf{R}_{e,i+1}^{-1} = \mathbf{P}_{i+1}\mathbf{A}_{i+1}^T(\mathbf{R}_{i+1}+\mathbf{A}_{i+1}\mathbf{P}_{i+1}\mathbf{A}_{i+1}^T)^{-1} \tag{D.21}$$
$$\begin{aligned}
&= \mathbf{P}_{i+1}\mathbf{A}_{i+1}^T[\mathbf{R}_{i+1}^{-1}-\mathbf{R}_{i+1}^{-1}\mathbf{A}_{i+1}(\mathbf{P}_{i+1}^{-1}+\mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}\mathbf{A}_{i+1})^{-1}\mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}]\\
&= \mathbf{P}_{i+1}[\mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}-\mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}\mathbf{A}_{i+1}(\mathbf{P}_{i+1}^{-1}+\mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}\mathbf{A}_{i+1})^{-1}\mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}]\\
&= \mathbf{P}_{i+1}[\mathbf{I}-\mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}\mathbf{A}_{i+1}(\mathbf{P}_{i+1}^{-1}+\mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}\mathbf{A}_{i+1})^{-1}]\mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}\\
&= \mathbf{P}_{i+1}[\mathbf{I}-\mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}\mathbf{A}_{i+1}(\mathbf{P}_{i+1}+(\mathbf{A}_{i+1}^T)^{-1}\mathbf{R}_{i+1}\mathbf{A}_{i+1}^{-1})]\mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}\\
&= \mathbf{P}_{i+1}[\mathbf{I}-\mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}\mathbf{A}_{i+1}\mathbf{P}_{i+1/i+1}]\mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}\\
&= \mathbf{P}_{i+1}[\mathbf{I}-(\mathbf{P}_{i+1/i+1}^{-1}-\mathbf{P}_{i+1}^{-1})\mathbf{P}_{i/i}]\mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}\\
&= \mathbf{P}_{i+1/i+1}\mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}
\end{aligned} \tag{D.22}$$

So finally, we have

$$\hat{\mathbf{x}}_{i+1/i+1} = \mathbf{F}_i\hat{\mathbf{x}}_{i/i}+\mathbf{P}_{i+1/i+1}\mathbf{A}_{i+1}^T\mathbf{R}_{i+1}^{-1}e_{i+1}. \tag{D.23}$$

Equation (D.23) corresponds to the M**U**E of the Kalman filter.

Bibliography

[ALLP12] E. Axell, G. Leus, E.G. Larsson, and H.V. Poor. Spectrum sensing for cognitive radio : State-of-the-art and recent advances. *IEEE Signal Processing Magazine*, 29(3):101–116, May 2012.

[ANS10] Z. Andalibi, H.H. Nguyen, and J.E. Salt. Channel estimation in bit-interleaved coded modulation with iterative decoding. *IET Communications*, 4(17):2095–2103, November 2010.

[APD77] D. B. Rubin A. P. Dempster, N. M. Laird. Maximum likelihood from incomplete data via the EM algorithm. *Royal Statistical Soc.*, 39(1):1–38, 1977.

[ASM12] E.H.M. Alian, H.E. Saffar, and P. Mitran. Cross-Band Interference Reduction Trade-Offs in SISO and MISO OFDM-based Cognitive Radios. *IEEE Transactions on Wireless Communications*, 11(7):2436–2445, 2012.

[B+14] M Berbineau et al. Cognitive radio for high speed railway through dynamic and opportunistic spectrum reuse. *TRA 2014, Paris, France*, April 2014.

[BD99] M. Bilodeau and D.Brenner. *Theory of Multivariate Statistics*. New York: Springer-Verlag, 1999.

[Bel63] P. Bello. Characterization of randomly time-variant linear channels. *IEEE Transactions on Communications Systems*, 11(4):360–393, December 1963.

[BID03] Raphael LE BIDAN. *Turbo-equalization for bandwidth-efficient digital communications over frequency-selective channels*. PhD

thesis, l'Institut National des Sciences Appliquées de Rennes, 2003.

[Bil98] J. A. Bilmes. *A Gentle Tutorial of the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models*. Berkeley CA, 1998.

[B.S00] B.SKLAR. *Digital Communications: Fundamentals and Applications*. Prentice Hall, 2000.

[BZ08] A. Batra and J.R. Zeidler. Narrowband interference mitigation in OFDM systems. In *IEEE Military Communications Conference, 2008. MILCOM 2008*, pages 1–7, 2008.

[BZ09] A. Batra and J.R. Zeidler. Narrowband interference mitigation in BICM OFDM systems. In *IEEE International Conference on Acoustics, Speech and Signal Processing, 2009. ICASSP 2009*, pages 2605–2608, 2009.

[Cha66] Robert W. Chang. Synthesis of band-limited orthogonal signals for multichannel data transmission. *The Bell System Technical Journal*, 45(10):1775–1796, Dec 1966.

[Cla68] R.H. Clarke. A statistical theory of mobile-radio reception. *Bell System Technical Journal, The*, 47(6):957–1000, July 1968.

[Cor] *corridor.ifsttar.fr.*

[Cou07] Alan James Coulson. Narrowband interference suppression for OFDM systems, 2007.

[CT05] Tao Cui and C. Tellambura. Pilot symbols for channel estimation in OFDM systems. In *GLOBECOM '05*, volume 4, pages 5 pp.–2233, Dec 2005.

[DGPV07] D. Darsena, G. Gelli, L. Paura, and F. Verde. A constrained maximum-sinr nbi-resistant receiver for OFDM systems. *IEEE Transactions on Signal Processing*, 55(6):3032–3047, 2007.

[DHLG14] Gaoqi Dou, Chunquan He, Congying Li, and Jun Gao. Channel estimation and symbol detection for OFDM systems using data-nulling superimposed pilots. *Electronics Letters*, 50(3):179–180, January 2014.

[DV08] D. Darsena and F. Verde. Successive NBI cancellation using soft decisions for OFDM systems. *IEEE Signal Processing Letters*, 15:873–876, 2008.

[EF10] Lloyd Emmanuel and Xavier N. Fernando. Wavelet-based spectral shaping of UWB radio signal for multisystem coexistence. *Computers & Electrical Engineering*, 36(2):261 – 268, 2010.

[FCC] *FCC. (2002, Nov.). Spectrum policy task force report. Federal Communications Commission, Tech. Rep. 02-135.*

[FRL08] Kun Fang, L. Rugini, and G. Leus. Iterative channel estimation and turbo equalization for time-varying ofdm systems. In *International Conference on Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE*, pages 2909–2912, 2008.

[GAD11] A Gomaa and N. Al-Dhahir. A sparsity-aware approach for NBI estimation in MIMO-OFDM. *IEEE Transactions on Wireless Communications*, 10(6):1854–1862, June 2011.

[GG03] M. Ghosh and V. Gadam. Bluetooth interference cancellation for 802.11g WLAN receivers. In *ICC '03.*, volume 2, pages 1169–1173, May 2003.

[GHRB12] Soukayna Ghandour-Haidar, Laurent Ros, and Jean-Marc Brossier. On the use of first-order autoregressive modeling for rayleigh flat fading channel estimation with kalman filter. *Signal Processing*, 92(2):601 – 606, 2012.

[GLL97] Alain Glavieux, Christophe Laot, and Joel Labat. Turbo equalization over a frequency selective channel. In *Proc. 1st Symp. Turbo Codes*, pages 96–102, 1997.

[GMAHS05] M. Ghogho, D. McLernon, E. Alameda-Hernandez, and A. Swami. Channel estimation and symbol detection for block transmission using data-dependent superimposed training. *IEEE Signal Processing Letters*, 12(3):226–229, March 2005.

[GS02] D. Gerakoulis and P. Salmi. An interference suppressing OFDM system for wireless communications. In *International Conference on Communications, IEEE ICC 2002.*, volume 1, pages 480–484, 2002.

[H+08] Sung-Hyun Hwang et al. Design and verification of IEEE 802.22 WRAN physical layer. In *CrownCom*, pages 1–6, May 2008.

[HR09] H. Hijazi and L. Ros. Analytical analysis of Bayesian Cramer Rao Bound for dynamical Rayleigh channel complex gains estimation in OFDM system. *IEEE Transactions on Signal Processing*, 57(5):1889–1900, May 2009.

[HR10] H. Hijazi and L. Ros. Joint data QR-detection and Kalman estimation for OFDM time-varying Rayleigh channel complex gains. *IEEE Transactions on Communications*, 58(1):170–178, January 2010.

[HSLR10] H. Hijazi, E.P. Simon, M. Lienard, and L. Ros. Channel estimation for mimo-ofdm systems in fast time-varying environments. In *4th International Symposium on Communications, Control and Signal Processing (ISCCSP), 2010*, pages 1–6, March 2010.

[HYK+08] Myeongsu Han, Takki Yu, Jihyung Kim, Kyungchul Kwak, Sungeun Lee, Seungyoup Han, and Daesik Hong. OFDM channel estimation with jammed pilot detector under narrow-band jamming. *IEEE Transactions on Vehicular Technology*, 57(3):1934–1939, 2008.

[Kay93] Steven M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory.* PRENTICE HALL, 1993.

[KML04] G. Y. Kung, M. W. Mac, and S. H. Lin. *Biometric Authentication: A Machine Learning Approach.* PRENTICE HALL, 2004.

[LGL01] Christophe Laot, A. Glavieux, and J. Labat. Turbo equalization: adaptive equalization and channel decoding jointly optimized. *IEEE Journal on Selected Areas in Communications*, 19(9):1744–1752, Sep 2001.

[LRF06] Ting-Jung Liang, W. Rave, and G. Fettweis. Iterative joint channel estimation and decoding using superimposed pilots in OFDM-WLAN. In *ICC '06.*, volume 7, pages 3140–3145, June 2006.

[Mit93] III Mitola, J. Software radios: Survey, critical evaluation and future directions. *IEEE Aerospace and Electronic Systems Magazine*, 8(4):25–36, April 1993.

[MK97] G. J. McLachlan and T. Krishnan. *The EM Algorithm and Extensions. Wiley Series in probability and Statistics*. Wiley, 1997.

[MM08] M. Morelli and M. Moretti. Robust frequency synchronization for OFDM-based cognitive radio systems. *IEEE Transactions on Wireless Communications*, 7(12):5346–5355, 2008.

[MM09] M. Morelli and M. Moretti. Channel estimation in OFDM systems with unknown interference. *IEEE Transactions on Wireless Communications*, 8(10):5338–5347, October 2009.

[NRK11] J.P. Nair and R.V. Raja Kumar. Improved superimposed training-based channel estimation method for closed-loop multiple input multiple output orthogonal frequency division multiplexing systems. *IET Communications*, 5(2):209–221, Jan 2011.

[OLLM04] A.G. Orozco-Lugo, M.M. Lara, and D.C. McLernon. Channel estimation using implicit training. *IEEE Transactions on Signal Processing*, 52(1):240–254, Jan 2004.

[PL12] The-Hanh Pham and Ying-Chang Liang. Joint channel information estimation and data detection for OFDM-based systems under unknown interference. In *IEEE Vehicular Technology Conference (VTC Spring)*, pages 1–5, May 2012.

[PR80] Abraham Peled and A. Ruiz. Frequency domain data transmission using reduced computational complexity algorithms. In *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '80.*, volume 5, pages 964–967, Apr 1980.

[Pro00] J.G. Proakis. *Digital Communications*. McGraw-Hill, 2000.

[PS99] Ian R. Petersen and Andrey V. Savkin. *Robust Kalman Filtering fo Signals and Systems with Large Uncertainties*. Birkhauser, 1999.

[Red02] Arthur J. Redfern. Receiver window design for multicarrier communication systems. *IEEE Journal on Selected Areas in Communications*, 20(5):1029–1036, 2002.

[RHS14a] Laurent Ros, Hussein Hijazi, and Eric Pierre Simon. Complex Amplitudes Tracking Loop for multipath channel estimation in OFDM systems over slow to moderate fading. *Signal Processing*, 97(April):134–145, April 2014.

[RHS14b] Laurent Ros, Hussein Hijazi, and Eric-Pierre Simon. Complex amplitudes tracking loop for multipath channel estimation in {OFDM} systems under slow to moderate fading. *Signal Processing*, 97:134 – 145, 2014.

[Say01] A.H. Sayed. A framework for state-space estimation with uncertain models. *Automatic Control, IEEE Transactions on*, 46(7):998 –1013, jul 2001.

[SC02] Ali H. Sayed and Hong Chen. A uniqueness result concerning a robust regularized least-squares solution. *Systems & Control Letters*, 46(5):361 – 369, 2002.

[SHR+10] E.P. Simon, H. Hijazi, L. Ros, M. Berbineau, and P. Degauque. Joint estimation of carrier frequency offset and channel complex gains for ofdm systems in fast time-varying vehicular environments. In *Communications Workshops (ICC), 2010 IEEE International Conference on*, pages 1–5, May 2010.

[SK13] E.P. Simon and M.A. Khalighi. Iterative soft-kalman channel estimation for fast time-varying mimo-ofdm channels. *Wireless Communications Letters, IEEE*, 2(6):599–602, December 2013.

[SKD09] S.M.-S. Sadough, M.-A. Khalighi, and P. Duhamel. Improved iterative MIMO signal detection accounting for channel-estimation errors. *IEEE Transactions on Vehicular Technology*, 58(7):3154–3167, Sept 2009.

[SN99] AliH. Sayed and VitorH. Nascimento. Design criteria for uncertain models with structured and unstructured uncertainties. In A. Garulli and A. Tesi, editors, *Robustness in identification and control*, volume 245 of *Lecture Notes in Control and Information Sciences*, pages 159–173. Springer London, 1999.

[SNB+04] F. Sjoberg, Rickard Nilsson, P.O. Borjesson, P. Odling, B. Wiese, and J. A C Bingham. Digital RFI suppression in DMT-based VDSL systems. *IEEE Transactions on Circuits and Systems I*, 51(11):2300–2312, 2004.

[SRH+11] E.P. Simon, L. Ros, H. Hijazi, Jin Fang, D.P. Gaillot, and M. Berbineau. Joint carrier frequency offset and fast time-varying channel estimation for MIMO-OFDM systems. *IEEE Transactions on Vehicular Technology*, 60(3):955–965, March 2011.

[SRHG12] E.P. Simon, L. Ros, H. Hijazi, and M. Ghogho. Joint carrier frequency offset and channel estimation for OFDM systems via the EM algorithm in the presence of very high mobility. *IEEE Transactions on Signal Processing*, 60(2):754 –765, February 2012.

[SRS14] Huaqiang Shu, L. Ros, and E.P. Simon. Simplified random-walk-model-based kalman filter for slow to moderate fading channel estimation in ofdm systems. *IEEE Transactions on Signal Processing*, 62(15):4006–4017, Aug 2014.

[SSR13] Huaqiang Shu, E.P. Simon, and L. Ros. Third-order kalman filter: Tuning and steady-state performance. *IEEE Signal Processing Letters*, 20(11):1082–1085, Nov 2013.

[SSR15] Huaqiang Shu, Eric Pierre Simon, and Laurent Ros. On the use of tracking loops for low-complexity multi-path channel estimation in {OFDM} systems. *Signal Processing*, 117:174 – 187, 2015.

[TCLB07] Zijian Tang, R.C. Cannizzaro, G. Leus, and P. Banelli. Pilot-assisted time-varying channel estimation for OFDM systems. *IEEE Transactions on Signal Processing*, 55(5):2226–2238, May 2007.

[TKS02] M. Tuchler, R. Koetter, and A.C. Singer. Turbo equalization: principles and new results. *IEEE Transactions on Communications*, 50(5):754–767, May 2002.

[TSK02] M. Tuchler, A.C. Singer, and R. Koetter. Minimum mean squared error equalization using a priori information. *IEEE Transactions on Signal Processing*, 50(3):673–683, Mar 2002.

[VPN11] Georgi Iliev Vladimir Poulkov, Miglen Ovtcharov and Zlatka Nikolova. *Narrowband Interference Suppression in MIMO Systems, MIMO Systems, Theory and Applications.* Dr. Hossein Khaleghi Bizaki (Ed.), 2011.

[WE71] S. Weinstein and P. Ebert. Data transmission by frequency-division multiplexing using the discrete Fourier transform. *IEEE Transactions on Communication Technology*, 19(5):628–634, October 1971.

[WHD+07] X. Wautelet, C. Herzet, A. Dejonghe, J. Louveaux, and L. Vandendorpe. Comparison of EM-based algorithms for MIMO

channel estimation. *IEEE Transactions on Communications*, 55(1):216–226, January 2007.

[WN05] Zhiqiang Wu and C.R. Nassar. Narrowband interference rejection in OFDM via carrier interferometry spreading codes. *IEEE Transactions on Wireless Communications*, 4(4):1491–1505, 2005.

[XM09] Huan Xu and S. Mannor. A kalman filter design based on the performance/robustness tradeoff. *Automatic Control, IEEE Transactions on*, 54(5):1171 –1175, may 2009.

[ZFC04] Dan Zhang, Pingyi Fan, and Zhigang Cao. Interference cancellation for OFDM systems in presence of overlapped narrow band transmission system. *IEEE Transactions on Consumer Electronics*, 50(1):108–114, 2004.

[ZS15] F. Zaarour and E.P. Simon. Fast time-varying channel estimation for ofdm systems with narrowband interference. *IEEE Wireless Communications Letters*, 4(4):389–392, Aug 2015.

[ZSZCD15a] F. Zaarour, E.P. Simon, M. Zwingelstein-Colin, and I. Dayoub. Comparison of superimposed pilot schemes in iterative receivers for ofdm systems. In *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, pages 1–6, June 2015.

[ZSZCD15b] Farah Zaarour, Eric Simon, Marie Zwingelstein-Colin, and Iyad Dayoub. A low complexity turbo receiver for data nulling superimposed pilots in OFDM. In *Proc. Fifth International Conference on Digital Information and Communication Technology and its Applications (DICTAP)*, pages 32–37, April 2015.

[ZZY13] Yang Zhang, Xin Zhang, and Dacheng Yang. A robust least square channel estimation algorithm for OFDM systems under interferences. In *IEEE Wireless Communications and Networking Conference (WCNC)*, pages 3122–3127, April 2013.