6-26-2020 10:00 AM

Novel Estimation and Detection Techniques for 5G Networks

Anas Saci, *The University of Western Ontario*

Supervisor: Shami, Abdallah, *The University of Western Ontario*
: Al-Dweik, Arafat, *The University of Western Ontario*
A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree
in Electrical and Computer Engineering
© Anas Saci 2020

Abstract

The thesis presents several detection and estimation techniques that can be incorporated into the fifth-generation (5G) networks. First, the thesis presents a novel system for orthogonal frequency division multiplexing (OFDM) to estimate the channel blindly. The system is based on modulating particular pairs of subcarriers using amplitude shift keying (ASK) and phase-shift keying (PSK) adjacent in the frequency domain, which enables the realization of a decision-directed (DD) one-shot blind channel estimator (OSBCE). The performance of the proposed estimator is evaluated in terms of the mean squared error (MSE), where an accurate analytical expression is derived and verified using Monte Carlo simulation under various channel conditions. The system has also extended to exploits the channel correlation over consecutive OFDM symbols to estimate the channel parameters blindly. Furthermore, a reliable and accurate approach has been introduced to evaluate the spectral efficiency of various communications systems. The metric takes into consideration the system dynamics, QoS requirements, and design constraints.

Next, a novel efficient receiver design for wireless communication systems that incorporate OFDM transmission has been proposed. The proposed receiver does not require channel estimation or equalization to perform coherent data detection. Instead, channel estimation, equalization, and data detection are combined into a single operation, and hence, the detector performs a direct data detector (D^3). The performance of the proposed D^3 is thoroughly analyzed theoretically in terms of bit error rate (BER), where closed-form accurate approximations are derived for several cases of interest, and validated by Monte Carlo simulations. The computational complexity of D^3 depends on the length of the sequence to be detected. Nevertheless, a significant complexity reduction can be achieved using the Viterbi algorithm (VA).

Finally, the thesis proposes a low-complexity algorithm for detecting anomalies in industrial steelmaking furnaces operation. The algorithm utilizes the vibration measurements collected from several built-in sensors to compute the temporal correlation using the autocorrelation function (ACF). Furthermore, the proposed model parameters are tuned by solving multi-objective optimization using a genetic algo-

rithm (GA). The proposed algorithm is tested using a practical dataset provided by an industrial steelmaking plant.

Summary for Lay Audience

In wireless communications systems, acquiring the knowledge of channel state information (CSI), commonly known as channel estimation (CE), and channel equalization are two fundamental tasks that a receiver device has to perform in order to extract the information symbols correctly. Generally speaking, the ultimate objective of designing CE methods is to maximize the data transmission reliability and data rate speed, while minimizing the complexity and processing delays. However, achieving all such conflicting objectives into one single design is generally not possible. Therefore, the system designer has to trade-off some of the objectives based on the availability of the resources and user experience requirements.

In this study, we present a new method that can perform the CE process with less data loss and power requirements as compared to conventional methods. Furthermore, this technique does not impose additional computational complexity to the existing well-known estimators and can be implemented smoothly in the current fifth-generation (5G) of mobile networks.

Besides, the thesis proposed a second receiver for wireless networks that can substantially reduce the power and the complexity of the current receivers with the advantage of the improved reliability. The proposed system was verified using probabilistic analysis and computer simulations over several environmental conditions.

Furthermore, the thesis introduces a low-complexity algorithm for detecting anomalies in industrial steelmaking furnaces operation by using vibration sensory data. Anomalous data are usually seen as alarms or an alert flag for some problems such as credit card fraud, health issues, and server crashes. Therefore, the anomaly detection can be considered as an essential diagnosing tool for manufacturing that may create many business perspectives. The proposed algorithm is tested using a practical dataset provided by an industrial steelmaking plant.

Dedication

To my dearest mother, Fatiha, and my dearest father, Djelloul.

Acknowledgements

First and foremost, I would like to express my gratitude to my supervisors, Professor Abdallah Shami and Dr. Arafat Al-Dweik for their diligent supervision throughout my Ph.D. years. Their continuous support and timely feedback were crucial to conduct my research smoothly.

I would also like to thank my examiners: Prof. Xianbin Wang, Dr. Serguei Primak, Dr. Anwar Haque, and Prof. George K. Karagiannidis for the time and efforts devoted to review and examine my thesis.

A sincere appreciation to my parents for the prayers and the unconditional support they offered during my Ph.D. journey. I extend my gratitude to my brother Abdelkader, and my sisters, Aicha and Aya, for their encouragement and support. Besides, deep gratitude and thanks to my brother Mohammad El Bashir, and I wish him a complete recovery.

My sincere thanks go to my friends Mohamad Kalil, Amr Gaballah, Moftah Ali, Osama Derbe, and Yousef Ettomi for their being true friends in times of hardship. I would also like to thank my OC2 colleagues for all the encouragement, assistance, and help that they offered over the Ph.D. years. Also, I thank my roommate, Alexander Hofkirchner, for being a considerate and kind roommate. Last but not least, I would like to thank the staff of the Electrical and Computer Engineering Department for being helpful and supportive.

Table of Contents

List of Tables

List of Figures

List of Appendices

Nomenclature

1-D   One-Dimensional

2-D   Two-Dimensional

5G   fifth-generation

ACF   Autocorrelation Function

AI   Artificial Intelligence

AIG   Autocorrelation Integrated Gaussian-based

ASK   Amplitude Shift Keying

AWGN   Additive White Gaussian Noise

BA   Basic Allocation

BER   bit error rate

BLER   Block Error Rate

BPSK   Binary Phase Shift Keying

BUx   Bad Urban

CE   Channel Estimation

CFO   Carrier Frequency Offset

CFR	Channel Frequency Response
CM	Constant Modulus
CP	Cyclic Prefix
CSI	Channel State Information
D^3	Direct Data Detector
DAB	digital audio broadcasting
DD	Decision-Directed
DS	Double-Sided
DVB-T2	Digital Video Broadcasting-Terrestrial
EC	Estimator Correlator
FB	Fully Blind
FEC	Forward Error Correction
FFT	Fast Fourier Transform
GD	Gradient Descent
GLRT	Generalized Likelihood Ratio Test
HDD	Hard Decision Decoder
i.i.d.	independent and identically distributed
IA	Incremental Allocation
ICI	Inter-Carrier Interference
IFFT	Inverse Fast Fourier Transform

IIoT	Industrial Internet of Things
IoT	Internet of Things
ISI	Inter-Symbol-Interference
k-NN	k-Nearest Neighbors
LLF	Log-Likelihood Function
LMS	Least Mean Squares
LR	Long Range
LS	Least-Squares
LSE	Least Square Estimation
LSTM	Long Short Term Memory
LTE	Long Term Evolution
LTE-A	Long Term Evolution-Advanced
M-LTE	Modified-Long Term Evolution
MASK	M-ary Amplitude Shift Keying
MIMO	Multiple-Input Multiple-Output
MINLP	Mixed-Integer Non-Linear Programming
ML	Maximum Likelihood
MLD	Maximum Likelihood Detector
MLSD	Maximum Likelihood Sequence Detector
MLSE	Maximum Likelihood Sequence Estimator

MMSE	Minimum Mean Squared Error
MPSK	M-ary Phase Shift Keying
MRC	Maximum Ratio Combiner
MSDD	Multiple Symbol Differential Detection
MSE	Mean Squared Error
MV	Moving Average
NN	Neural Network
NR	New Radio
OFDM	Orthogonal Frequency Division Multiplexing
OWC	Optical Wireless Communications
PCA	Principal Component Analysis
PDF	Probability Density Function
PEP	Pairwise Error Probability
PHY	Physical Layer
PSK	Phase Shift Keying
PSP	Per-Survivor Processing
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
QPSK	Quadrature Phase Shift Keying
RAx	Rural Area

RF	Random Forest
RNN	Recurrent Neural Network
SEP	Sequence Error Probability
SER	Symbol Error Rate
SIMO	Single-Input Multiple-Output
SISO	Single-Input Single-Output
SNR	Signal-to-Noise Ratio
SR	Short Range
SS	Single-Sided
STO	Symbol Timing Offset
SVD	Singular Value Decomposition
SVM	Support Vector Machine
SY	Synchronization
TUx	Typical Urban
VA	Viterbi Algorithm
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless Local Area Networks
ZF	Zero-Forcing

Chapter 1

Introduction

From healthcare to finance, commerce to transportation, and government to entertainment, wireless systems are shaping the world around us. Over the past decade, the deployment of Internet of Things (IoT) devices and connected sensors have experienced an upsurge due to the vast number of applications that rely on these technologies. Most IoT devices generate live data streams that can be used for detecting particular phenomena or isolated systems through indirect inference. Recent reports predict that the number of connected devices will exceed 75 billion by 2025 [1]. To accommodate such a massive number of devices, IoT and cloud computing have become key technologies that enable efficient monitoring, scalability and maintenance of the network [2]. With the deployment of the fifth-generation (5G) wireless networks, industrial 4.0 technologies, which are empowered by the industrial IoT (IIoT) have enabled a range of new application such as edge computing [3, 4], smart grids [5], and smart manufacturing [6]. IIoT allows collecting a large amount of data gathered by various sensors to monitor the progress of the manufacturing process, and predict its evolution [7]. Consequently, researchers from both academia and industry are devoting efforts to maximize the spectral efficiency of wireless networks by optimizing the spectrum utilization across all layers of the communications protocol stack.

At the physical layer, acquiring the knowledge of channel state information (CSI), commonly known as channel estimation (CE), and channel equalization are two fundamental tasks that a receiver has to perform before information symbols extraction from the received signal. In the literature, CE techniques can be classified based on their spectral efficiency, estimation accuracy, computational complexity, or the required observation window size. An efficient estimator is the one that can maximize the accuracy and spectral efficiency while minimizing the complexity and observation window size. However, achieving all such objectives simultaneously is not feasible. Hence, the system designer has to trade-off between these objectives based on the available system resources and quality of service (QoS) requirements. As a result, CE processes have to be performed accurately to prevent any system performance degradation, and with less signaling overhead to maximize the spectral efficiency.

In the context of smart manufacturing, anomaly detection is defined as the process of identifying odd or rare events, or observations flagging suspicious behavior that does not confront with the typical nature of the data in an automated manner [5]. Generally speaking, anomalous data are usually seen as alarms or an alert flag for some problems such as credit card fraud, health issues, and server crashes. Therefore, the anomaly detection can be considered as an essential diagnosing tool for manufacturing that may create many business perspectives.

1.1 Thesis Outline

This thesis presents numerous detection and estimation techniques that can be incorporated in the 5G networks and beyond from the communication and digital signal processing side and from the application side in an integrated article format. Chapter 2 proposes a novel one-shot blind channel estimation (OSBCE) technique which enables estimating the channel without the need to transmit dedicated pilot symbols

at the transmission by modulating a pair of data subcarriers to the CSI. Chapter 3 presents an extension to the OSBCE by utilizing the correlation of the subcarriers in the time domains. The results for Chapter 2 and Chapter 3 have been validated using theoretical analysis and computer simulations. Chapter 4 presents a spectral efficiency metric for orthogonal frequency division multiplexing (OFDM) that takes into account several system parameters and QoS constraints. Chapter 5 present a novel OFDM-based receiver that extracts the information directly from the FFT outputs without the need to perform channel estimation, channel interpolation, and channel equalization. The system has been thoroughly analyzed and validated over several frequency selective channel models. It is worth mentioning that Chapters 2-5 are focusing on performing the estimation at the physical layer using digital signal processing techniques to perform the channel estimation and to extract the information bits in OFDM-bases systems such as the 5G New Radio (NR). On the other hand, Chapter 6 presents a novel detection technique in the application side of 5G for anomaly detection in IoT data, which is considered as one of the key applications of 5G networks. The proposed detection algorithm uses vibration IoT sensor measurements based on statistical and signal processing techniques. Chapter 7 concludes the thesis and provides some suggestions for future research directions.

1.2 Thesis Contributions

The main contributions of each chapter are listed below:

1.2.1 Contribution of Chapter 2

Chapter 2 proposes a novel OFDM-based system is based on modulating particular pairs of subcarriers using amplitude shift keying (ASK) and phase shift keying (PSK), which enables the realization of a decision-directed (DD) OSBCE, with complexity

and accuracy that are comparable to pilot-based channel estimation techniques. The performance of the proposed estimator is evaluated in terms of the mean squared error (MSE), where an accurate analytical expression is derived and verified using Monte Carlo simulation under various channel conditions. The obtained results show that the MSE of the proposed OSBCE is comparable to pilot-based estimators, which confirms the efficiency of the proposed OSBCE. The chapter main contributions are:

1. Novel blind OSBCE is introduced for OFDM that enables DD estimation without the need for pilot symbols.

2. Analytical analysis for the SER at MPSK symbols, which are used as pilots for the subsequent steps. Also, a closed-form expression for the probability density function (PDF) of the phase error ψ_k^ℓ in Rayleigh fading channels with imperfect channel knowledge at the receiver

3. An accurate analytical expression for the MSE in DD manner has been derived.

4. The system is verified under the Long Term Evolution (LTE) downlink physical (PHY) layer specifications and over practical frequency selective channels.

1.2.2 Contribution of Chapter 3

The chapter presents an estimator that exploits the channel correlation over consecutive OFDM symbols to estimate the channel parameters blindly. In the new estimator, particular subcarriers re modulated using M-ary phase shift keying (MPSK), and subcarriers with the same indices in some consecutive OFDM symbols are modulated using M-ary amplitude shift keying (MASK), replacing the pilots in pilot-aided systems. Consequently, all subcarriers are data-bearing, which leads to spectral efficiency improvement. The proposed estimator uses the feature that MPSK and MASK modulated symbols have sufficient CSI that enables them to cooperate in order to detect the MPSK symbols coherently and blindly. Then, the CSI at the corresponding

MPSK symbols can be acquired in a DD fashion. The performance of the proposed estimator is evaluated in term of symbol error rate (SER) and mean-squared error (MSE), where an exact analytical formula is obtained for the SER of binary phase shift keying (BPSK) symbols in mobile radio channels with various time-varying rates. The obtained results show that the proposed estimator produces accurate channel estimates as compared to pilot-aided and state-of-the-art systems without additional complexity. The chapter's main contributions are:

1. The chapter presents a blind channel estimation technique that utilizes the correlation of consecutive OFDM symbols in the time domain.

2. Closed form SER analysis is present and validated for several channel mobility conditions.

3. The system was evaluated over several channel doppler frequencies according to the LTE downlink specifications.

4. A hybrid scheme can be adopted according to the channel conditions on to maximize the channel estimation accuracy.

1.2.3 Contribution of Chapter 4

In the literature, the spectral efficiency is considered as a key performance indicator that is used to classify various communications systems, algorithms, and techniques. However, the classification is typically performed while assuming that all systems have a static structure, and without considering the quality of service (QoS) requirements or the constraints imposed by the system design. Therefore, this work presents a new reliable and accurate approach to evaluate the spectral efficiency of communications systems while considering the system dynamics, QoS requirements and design constraints. To demonstrate its effectiveness, the proposed metric is used to eval-

uate the spectral efficiency of various blind and pilot-aided channel estimation and synchronization algorithms. The chapter's main contributions are:

1. The proposed system can be incorporated effectively and efficiently in practical systems such as LTE-A standard.

2. A new fair and reliable approach has been proposed to compute and compare the spectral efficiency of various blind and non-blind communications systems.

3. A sub-optimal solution for the bit loading problem has been considered, where the modulation order for each subcarrier is assigned independently of all other subcarriers based on its signal-to-noise ratio (SNR), this approach will be denoted as the basic allocation (BA) algorithm, and incremental allocation (IA) algorithm. In such scenarios, the bit error rate (BER) constraint is satisfied when the BER for each subcarrier is satisfied.

4. The new approach considers the fact that different subcarriers in OFDM systems may be modulated using different modulation types and orders to satisfy QoS requirements.

1.2.4   Contribution of Chapter 5

The chapter presents a novel efficient receiver design for wireless communication systems that incorporate OFDM transmission scheme. The proposed receiver does not require channel estimation or equalization to perform coherent data detection. Instead, channel estimation, equalization, and data detection are combined into a single operation, and hence, the detector performs a direct data detector (D^3). The performance of the proposed D^3 is thoroughly analyzed theoretically in terms of BER, where closed-form accurate approximations are derived for several cases of interest, and validated by Monte Carlo simulations. The D^3 is applied for key practical wireless systems such as the LTE and the NR of the 5G system. The obtained theoretical

and simulation results demonstrate that the BER of the proposed D^3 is only 3 dB away from coherent detectors with perfect knowledge of the CSI in flat and frequency-selective fading channels for a wide range of SNRs. If CSI is not known perfectly, then D^3 outperforms the coherent detector substantially, particularly at high SNRs with linear interpolation. The computational complexity of D^3 depends on the length of the sequence to be detected. Nevertheless, a significant complexity reduction can be achieved using the Viterbi algorithm. The chapter's main contributions are:

1. Unlike conventional OFDM detectors, this chapter presents a new detector to regenerate the information symbols directly from the received samples at the FFT output, which is denoted as the direct data detector (D^3). By implementing the D^3, there is no need to perform channel estimation, interpolation, equalization, or symbol decision operations.

2. An efficient implementation using 1-D decomposition of the 2-D search space and the implementation of Viterbi algorithm (VA) results in a substantial complexity reduced as compared to conventional detectors.

3. Comprehensive complexity analysis comparison of the D^3 system as compared to the pilot-based system with linear interpolation for various modulation orders.

4. The D^3 performance is evaluated in terms of complexity, computational power, BER, where analytic expressions are derived for several channel models and system configurations. The D^3 BER is compared to other widely used detectors such as the maximum likelihood (ML) coherent detector with perfect and imperfect CSI, multiple symbol differential detector (MSDD), the ML sequence detector (MLSD) with no CSI, and the per-survivor processing detector.

1.2.5 Contribution of Chapter 6

In this chapter, a novel low-complexity anomaly detection algorithm is designed by exploiting the temporal signals collected from multiple sensors. The algorithm is based on modeling the healthy processes as a Gaussian distribution, and hence, the anomalous processes can be distinguished based on a threshold set for the Gaussian PDF. Also, it captures the temporal data obtained by creating new autocorrelation features. The main contributions of this chapter can be summarized as:

1. The design of a low-complexity and accurate automated anomaly detection algorithm, based on the joint Gaussian modeling of the sensor vibration signals from the historical data.

2. The proposed algorithm can detect temporal fluctuations of the vibration signals by constructing new features that carry autocorrelation information. Therefore, the algorithm is denoted as autocorrelation integrated Gaussian (AIG) algorithm.

3. An objective and informative metrics for anomaly detection is proposed in a time-series data that reflects a better insight into the overall performance of the anomaly detection of temporal data.

4. Formulate and solve an optimization framework for the hyper-parameter tuning of the proposed algorithm. The obtained results show a reliable and timely estimation of the anomalous process against the true anomaly labels while keeping the processing time complexity substantially low both in training and execution phases.

Bibliography

[1] I. L. BRIEF, "Internet of things," 2019.

[2] Y. Xu and A. Helal, "Scalable cloud–sensor architecture for the internet of things," *IEEE Internet Things J*, vol. 3, no. 3, pp. 285–298, 2015.

[3] T. Wang, H. Ke, X. Zheng, K. Wang, A. K. Sangaiah, and A. Liu, "Big data cleaning based on mobile edge computing in industrial sensor-cloud," *IEEE Trans. Ind. Informat.*, vol. 16, no. 2, pp. 1321–1329, Feb 2020.

[4] Y. Kang, I. Park, J. Rhee, and Y. Lee, "Mongodb-based repository design for iot-generated rfid/sensor big data," *IEEE Sensors J.*, vol. 16, no. 2, pp. 485–497, 2016.

[5] S. Chen, H. Wen, J. Wu, W. Lei, W. Hou, W. Liu, A. Xu, and Y. Jiang, "Internet of things based smart grids supported by intelligent edge computing," *IEEE Access*, vol. 7, pp. 74 089–74 102, 2019.

[6] J. Leng, D. Yan, Q. Liu, K. Xu, J. L. Zhao, R. Shi, L. Wei, D. Zhang, and X. Chen, "Manuchain: Combining permissioned blockchain with a holistic optimization model as bi-level intelligence for smart manufacturing," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 50, no. 1, pp. 182–192, Jan 2020.

[7] T. Hussain, K. Muhammad, J. Del Ser, S. W. Baik, and V. H. C. de Albuquerque, "Intelligent embedded vision for summarization of multi-view videos in iiot," *IEEE Trans. Ind. Informat.*, 2019.

[8] M. S. Mahdavinejad, M. Rezvan, M. Barekatain, P. Adibi, P. Barnaghi, and A. P. Sheth, "Machine learning for internet of things data analysis: A survey," *Digital Commun. and Networks*, vol. 4, no. 3, pp. 161–175, 2018.

[9] M. Kwon, J. Lee, and H. Park, "Intelligent iot connectivity: Deep reinforcement learning approach," *IEEE Sensors J.*, vol. 20, no. 5, pp. 2782–2791, 2020.

[10] T. Yu, X. Wang, and A. Shami, "A novel fog computing enabled temporal data reduction scheme in iot systems," in *2017 IEEE Global Commun. Conf. (GLOBECOM)*. IEEE, 2017, pp. 1–5.

[11] M. Injadat, F. Salo, A. B. Nassif, A. Essex, and A. Shami, "Bayesian optimization with machine learning algorithms towards anomaly detection," in *2018 IEEE Global Commun. Conf. (GLOBECOM)*. IEEE, 2018, pp. 1–6.

[12] J. Park, "Anomaly detection from the signal of low-cost laser device without the false alarm and the missing," *IEEE Sensors J.*, vol. 18, no. 10, pp. 4275–4285, 2018.

[13] P. Ni, T. Haglund, and M. Ersson, "Study on slopping prevention in the bof steelmaking process," *steel research international*, vol. 88, no. 8, p. 1600399, 2017.

[14] M. Shakirov, A. Boutchenkov, G. Galperine, and B. Schrader, "Prediction and prevention of slopping," in *Proceedings from ISSTech 2003*, 2003.

[15] J. Kafie, B. Babaei, and V. Scipolo, "Bof slop detection - savings potential with real-time slop detection and mitigation technology," in *Proceedings from AISTech 2015*, 2015.

[16] D. Huang, D. Mu, L. Yang, and X. Cai, "Codetect: Financial fraud detection with anomaly feature detection," *IEEE Access*, vol. 6, pp. 19 161–19 174, 2018.

[17] I. A. Khan, D. Pi, Z. U. Khan, Y. Hussain, and A. Nawaz, "Hml-ids: A hybrid-multilevel anomaly prediction approach for intrusion detection in scada systems," *IEEE Access*, vol. 7, pp. 89 507–89 521, 2019.

[18] S. Jin, Z. Zhang, K. Chakrabarty, and X. Gu, "Toward predictive fault tolerance in a core-router system: Anomaly detection using correlation-based time-series analysis," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 37, no. 10, pp. 2111–2124, 2017.

[19] J. Jiang, J. Chen, T. Gu, K.-K. R. Choo, C. Liu, M. Yu, W. Huang, and P. Mohapatra, "Anomaly detection with graph convolutional networks for insider threat and fraud detection," in *MILCOM 2019-2019 IEEE Military Commun. Conference (MILCOM)*. IEEE, 2019, pp. 109–114.

[20] W. Rasheed and T. B. Tang, "Anomaly detection of moderate traumatic brain injury using auto-regularized multi-instance one-class svm," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 1, pp. 83–93, 2020.

[21] S. Rajendran, W. Meert, V. Lenders, and S. Pollin, "Unsupervised wireless spectrum anomaly detection with interpretable features," *IEEE Trans. Cogn. Commun. Netw*, vol. 5, no. 3, pp. 637–647, 2019.

[22] S. Naseer, Y. Saleem, S. Khalid, M. K. Bashir, J. Han, M. M. Iqbal, and K. Han, "Enhanced network anomaly detection based on deep neural networks," *IEEE Access*, vol. 6, pp. 48 231–48 246, 2018.

[23] Y. Lin and J. Wang, "Probabilistic deep autoencoder for power system measurement outlier detection and reconstruction," *IEEE Trans. Smart Grid*, vol. 11, no. 2, pp. 1796–1798, 2020.

[24] A. E. Mekki, A. Bouhoute, and I. Berrada, "Improving driver identification for the next-generation of in-vehicle software systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 7406–7415, 2019.

[25] K. Zhu, Z. Chen, Y. Peng, and L. Zhang, "Mobile edge assisted literal multi-dimensional anomaly detection of in-vehicle network using lstm," *IEEE Trans. Veh. Technol.*, vol. 68, no. 5, pp. 4275–4284, 2019.

[26] W. Yu, T. S. Dillon, F. Mostafa, W. Rahayu, and Y. Liu, "A global manufacturing big data ecosystem for fault detection in predictive maintenance," *IEEE Trans. Ind. Informat.*, 2019.

[27] Q. Jiang, X. Yan, and B. Huang, "Performance-driven distributed pca process monitoring based on fault-relevant variable selection and bayesian inference," *IEEE Trans. Ind. Electron.*, vol. 63, no. 1, pp. 377–386, 2015.

[28] W. Hu, J. Gao, B. Li, O. Wu, J. Du, and S. J. Maybank, "Anomaly detection using local kernel density estimation and context-based regression," *IEEE Trans. Knowl. Data Eng.*, 2018.

[29] M. Desforges, P. Jacob, and J. Cooper, "Applications of probability density estimation to the detection of abnormal conditions in engineering," *Proceedings of the Institution of Mechanical Engineers, Part C: J. of Mechanical Engineering Science*, vol. 212, no. 8, pp. 687–703, 1998.

[30] R. Moghaddass and J. Wang, "A hierarchical framework for smart grid anomaly detection using large-scale smart meter data," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 5820–5830, 2017.

[31] W. Haider, J. Hu, Y. Xie, X. Yu, and Q. Wu, "Detecting anomalous behavior in cloud servers by nested arc hidden semi-markov model with state summarization," *IEEE Trans. Big Data*, 2017.

[32] X. Yang and W. Wen, "Ridge and lasso regression models for cross-version defect prediction," *IEEE Trans. Rel.*, vol. 67, no. 3, pp. 885–896, 2018.

[33] R. Sakia, "The box-cox transformation technique: a review," *J. of the Royal Statistical Society: Series D (The Statistician)*, vol. 41, no. 2, pp. 169–178, 1992.

[34] Proakis, *Digital Commun. 5th Edition.* McGraw Hill, 2007.

[35] H.-L. Liu, L. Chen, K. Deb, and E. D. Goodman, "Investigating the effect of imbalance between convergence and diversity in evolutionary multiobjective algorithms," *IEEE Trans. Evol. Comput.*, vol. 21, no. 3, pp. 408–425, 2016.

[36] S. H. Ebenuwa, M. S. Sharif, M. Alazab, and A. Al-Nemrat, "Variance ranking attributes selection techniques for binary classification problem in imbalance data," *IEEE Access*, vol. 7, pp. 24 649–24 666, 2019.

Chapter 2

One-Shot Blind Channel Estimation for OFDM Systems Over Frequency Selective Fading Channels

2.1 Preamble

[1]Orthogonal frequency division multiplexing (OFDM) is a multicarrier modulation technique that received tremendous interest from the industry and academia over the past decade. The catalyst for such interest, is the special features that OFDM has such as spectral efficiency, immunity to multipath propagation, efficient implementation using the fast Fourier transform (FFT) pairs, and low-complexity equalization [2]. Therefore, OFDM is currently adopted by several commercial standards, such as the second generation digital video broadcasting-terrestrial (DVB-T2) [4], wireless local area networks (WLAN) IEEE 802.11 [4], Worldwide Interoperability for Microwave Access (WiMAX) [5] and Long Term Evolution Advanced (LTE-A) [3]. Moreover, OFDM is adopted in the fifth-generation (5G) new radio (NR) wireless

---

[1]A version of this chapter has been published in [1].

14

communications standard [7].

Acquiring the knowledge of channel state information (CSI), commonly known as channel estimation (CE), and channel equalization are two fundamental tasks that a receiver has to perform prior to information symbols extraction from the received signal. Generally speaking, most CE techniques reported in the literature can be classified based on their spectral efficiency, estimation accuracy, computational complexity, or the required observation window size. An efficient estimator is the one that can maximize the accuracy and spectral efficiency, while minimizing the complexity and observation window size. However, achieving all such conflicting objectives into one single design is generally infeasible, and hence, the system designer has to compromise some of these objectives based on the available system resources and quality of service (QoS) requirements. For most practical applications, a reasonable compromise can be achieved by using training symbols, where CSI is estimated at the receiver side by modulating particular subcarriers at the transmitter side using known symbols, denoted as reference or pilot symbols, and then, use such pilots for CE purposes [8]-[6]. In LTE-A [3], comb-type pilots are deployed in a time-frequency subcarrier grid as shown in Fig. 1a, where the pilots occupy about 4.7% of the total number of subcarriers. The spectral efficiency can be even lower for some other systems, such as the IEEE 802.11n, where pilot symbols constitute 7.1% of the total subcarriers. Moreover, in communications systems that involve burst transmission, such as frequency hopping and cognitive radio, the channel coefficients over consecutive OFDM symbols can be uncorrelated, and hence, pilot symbols are needed in every OFDM symbol, which degrades the spectral efficiency even further. Therefore, many algorithms have been proposed to estimate the CSI blindly, by utilizing only the received data symbols [24]-[26], and consequently, improve the spectral efficiency.

In the literature, blind CE is one of the widely addressed issues in wireless communications [9]-[27]. However, the problem remains open because, to the best of our
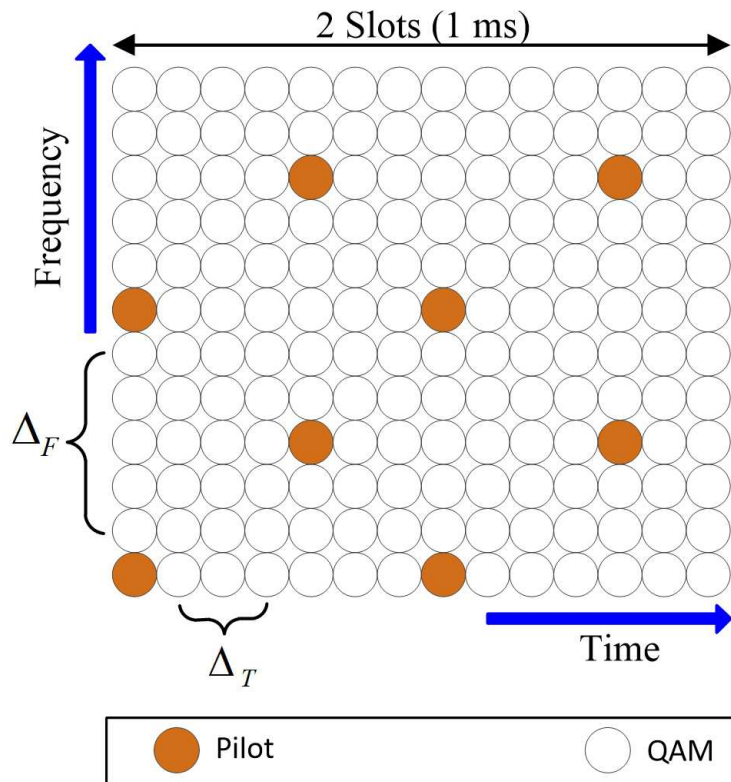
Figure 2.1: Time-frequency grid of LTE-A transmission grid.

knowledge, none of the techniques reported in the literature managed to resolve the conflicting objectives problem perfectly. For example, several blind estimators are designed based on the assumption that all subcarriers should be modulated using a constant modulus (CM) constellation, such as M-ary phase shift keying (MPSK) [10]-[12]. Although such techniques do not require pilot symbols, they indirectly degrade the spectral efficiency of the system because they prohibit using spectrally efficient modulation schemes, such as quadrature amplitude modulation (QAM). Therefore, it would be more factual to denote such techniques as conditionally-blind, to distinguish them from fully blind (FB) systems, which do not require pilot symbols and do not have any constraints on the modulation type or order. It is worth noting that conditionally-blind systems are different from semi-blind systems, in which the CE is performed using both, the pilot and data symbols [25]-[27].

Computational complexity is another critical performance indicator used to com-

pare various channel estimation techniques. Generally speaking, blind estimation techniques have higher computational complexity than pilot-based techniques [13], [14]. The excessive computational complexity is mainly caused by the iterative structure of such algorithms [14]-[24], or due to the requirements to perform an extensive search over the solution space [28]. Although the complexity of the system reported in [28] becomes comparable to pilot-based estimation at high signal-to-noise ratios (SNRs), such condition can be frequently violated in practical scenarios.

The observation window size specifies the number of OFDM symbols required to compute the CSI estimates. In CE techniques that require large observation window, the channel is assumed to be fixed over the observation period [10], [11], [12], [14]. Although such assumption is acceptable for static and slow fading channels, it is not actually the case for fast fading channels. Moreover, if the observation window size is very large, then such assumption becomes realizable only for static channels. Channel estimators that can perform CE within one OFDM block, denoted as one-shot estimators, usually outperform other estimators with multiple-symbols observation window [28].

The accuracy of CE techniques is typically evaluated using mean squared error (MSE), which is supposed to be sufficiently low to minimize the bit error rate (BER) degradation caused by CE errors [28]. In the literature, the performance of pilot-based CE is commonly used as a benchmark for comparison [8], [19], [28], because pilot-based estimation techniques offer reliable estimates and their impact on the BER is tolerable.

As it can be noted from the aforementioned discussion, pilot-based estimators have several attractive features in terms of complexity, estimation accuracy, and observation window size. However, the spectral efficiency remains the major concern. Practically speaking, the dominant standards such as DVB-T2 [4], WiMAX [5] and LTE-A [3], are using pilot-based CE, which implies that systems' designers prefer to
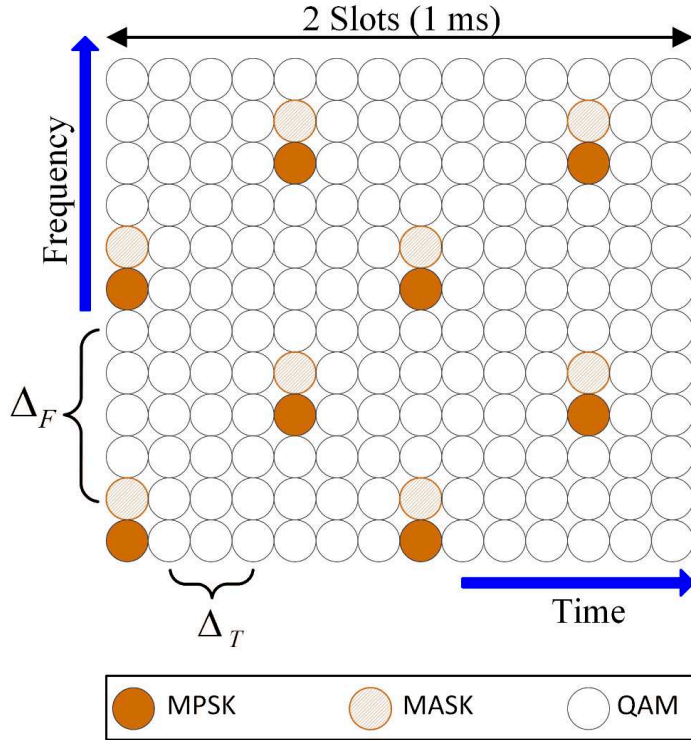
Figure 2.2: Time-frequency grid of OSBCE transmission grid.

compromise the spectral efficiency, but gain the other advantages. Unlike most of the work reported in the literature, this work presents a novel one-shot blind channel estimation (OSBCE) technique which is conditionally-blind, but only a small fraction of the subcarriers have the modulation-type constraint, and hence, it is spectrally efficient. The OSBCE is based on replacing the pilot symbols in conventional pilot-based systems with MPSK symbols, and the modulation type of the adjacent subcarriers is limited to M-ary amplitude shift keying (MASK) modulation. Such configuration allows detecting the MPSK symbols coherently using the partial channel information embedded in the MASK symbols. Then, the detected MPSK symbols are used to obtain the full channel information in a decision-directed (DD) manner. The complexity, observation window and accuracy of the OSBCE are comparable to pilot-based estimators while the spectral efficiency of the OSBCE is higher than pilot systems at moderate and high SNRs. The system performance is evaluated in terms of the symbol error rate (SER) and MSE, where a closed-form formula is derived

for the SER while the MSE can be efficiently evaluated numerically. The obtained analytical results are confirmed using Monte Carlo simulation.

Notation: The notation used in this chapter is as follows. Uppercase boldface letters such as $\mathbf{H}$ will denote $N \times N$ matrices, whereas lowercase boldface letters such as $\mathbf{x}$ will denote row or column vectors with N elements. Calligraphic letters with a hat such as $\hat{\mathcal{A}}$ will denote initial estimates of the variable A, while a regular symbol with a hat such as $\hat{A}$ will denote the final estimate of A. Blackboard symbols with the subscript such as $\mathbb{M}_Q$ will denote the set $\{0, 1, \ldots, M_Q - 1\}$. The Euclidean norm is denoted as $\|.\|$, and the operators '$|$' and '$\mid$' will denote the conditioning operation, interchangeably. Furthermore, $\mathrm{E}[\cdot]$ denotes the expectation process, the complex conjugate, transpose, and Hermitian transpose will be denoted as $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$, respectively.

2.2   OFDM System and Channel Models

Consider an OFDM system with N subcarriers modulated by a sequence of N complex data symbols $\mathbf{a} = [A_0, \ A_1, \ ...., \ A_{N-1}]^T$. The data symbols are selected uniformly from a general constellation such as MPSK, QAM or MASK, with modulation orders M_P, M_Q and M_A, respectively. In pilot-based practical OFDM systems [3], N_P of the subcarriers are reserved for pilot symbols, which can be used for channel estimation and synchronization purposes. For the purpose of this work, we define three sets of indices for the subcarriers, namely, the set of pilots' indices $\mathbb{V} = \{v_1, v_2, \ldots, v_{N_P}\}$, the set of subcarriers' indices adjacent to the pilots $\mathbb{T} = \{t_1, t_2, \ldots, t_{N_P}\}$, and the set of indices of the remaining data symbols is denoted as $\mathbb{U}$. It is worth noting that the three sets are disjoint, $\mathbb{V} \cap \mathbb{T} \cap \mathbb{U}$, and $t_i = v_i + 1$.

The modulation process can be implemented efficiently using N-point inverse FFT (IFFT). The output of the IFFT process during the ℓth OFDM block is given by

$\mathbf{x}(\ell) = \mathbf{F}^H \mathbf{a}(\ell)$, where $\mathbf{F}$ is the normalized $N \times N$ FFT matrix, and hence, $\mathbf{F}^H$ is the IFFT matrix. The elements of $\mathbf{F}^H$ are defined as $F_{i,k} = (1/\sqrt{N})e^{j2\pi ik/N}$ where i and k denote the row and column indices $i, k \in \{0, 1, ..., N-1\}$, respectively. For convenience, we drop the block index notation ℓ in the remaining parts unless it is necessary to include it. To eliminate inter-symbol-interference (ISI) between consecutive OFDM symbols and maintain the subcarriers' orthogonality in frequency-selective multipath fading channels, a cyclic prefix (CP) of length N_{CP} samples no less than the channel maximum delay spread (L_{h}) is formed by copying the last N_{CP} samples of $\mathbf{x}$ and appending them in front of the IFFT output to compose the OFDM symbol with a total length $N_{\mathrm{t}} = N + N_{\mathrm{CP}}$ samples and a duration of T_{t} seconds. Then, the complex baseband OFDM symbol during the ℓth signaling period $\tilde{\mathbf{x}}$ is upsampled, filtered and up-converted to a radio frequency centered at f_{c} before transmission through the antenna.

At the receiver front-end, the received signal is down-converted to baseband and sampled at a rate $T_{\mathrm{s}} = T_{\mathrm{t}}/N_{\mathrm{t}}$. In this work, we assume that the channel is composed of $L_{\mathrm{h}}+1$ independent multipath components each of which has a gain $h_m \sim \mathcal{CN}\left(0, 2\sigma_{h_m}^2\right)$ and delay $m \times T_{\mathrm{s}}$, where $m \in \{0, 1, ..., L_{\mathrm{h}}\}$. The channel taps are assumed to be constant over one OFDM symbol, but they may change over two consecutive symbols, which corresponds to a quasi-static multipath channel [22]. Then, the received signal after discarding the first N_{CP} CP samples, and computing the FFT can be expressed as

$$\mathbf{r} = \mathbf{Ha} + \mathbf{w} \tag{2.1}$$

where $\mathbf{r} \in \mathbb{C}^{N \times 1}, \mathbf{w}$ denotes the additive white gaussian noise (AWGN) vector, whose samples are independent and identically distributed (i.i.d.) $w_k \sim \mathcal{CN}(0, 2\sigma_w^2)$, and $\mathbf{H}$ denotes the channel frequency response (CFR), which is defined as

$$\mathbf{H} = \mathrm{diag}\left\{[H_0, H_1, \ldots, H_{N-1}]\right\} \tag{2.2}$$

where $H_k = \sum_{m=0}^{L_h} h_m e^{-j2\pi mk/N}$.

It is worth noting that the diagonal elements of $\mathbf{H}$ are highly correlated, particularly the adjacent elements where the correlation coefficient $\varrho \triangleq \mathrm{E}\left[H_k H_{k+1}^*\right]$ is defined as

$$\varrho = \mathrm{E}\left[\sum_{n=0}^{L_h}\sum_{m=0}^{L_h} h_n h_m^* e^{j2\pi \frac{-(n-m)k+n}{N}}\right]. \tag{2.3}$$

Given that h_m and h_n are mutually independent $\forall m \neq n$, then $E\left[|h_n|^2\right] = 2\sigma_{h_n}^2$ and $\mathrm{E}\left[h_n h_m^*\right]\big|_{n\neq m} = 0$. Thus

$$\varrho = \sum_{n=0}^{L_h} \sigma_{h_n}^2 \; e^{j2\pi \frac{n}{N}}. \tag{2.4}$$

The elements of the FFT output are then fed to a single-tap zero-forcing (ZF) or minimum mean squared error (MMSE) equalizer, followed by a maximum likelihood detector (MLD). In this work we consider the ZF equalizer, and hence, the estimated kth symbol can be expressed as

$$\hat{A}_k = \arg\min_{A_k^{(i)}, i \in \mathbb{M}_Q} \left| \frac{\hat{H}_k^*}{|\hat{H}_k|^2} r_k - A_k^{(i)} \right|^2, \; k \notin \mathbb{V} \tag{2.5}$$

where $\hat{H}_k$ is the estimated CFR at the kth subcarrier, $A_k^{(i)}$ are the trial values of the data symbols. It is worth noting that single-tap frequency-domain ZF and MMSE equalizers offer approximately the same performance in quasi-static single-input single-output (SISO) systems [30], [31]. However, the mathematical analysis of the MMSE equalizer is more intractable.

In OFDM-based systems such as LTE-A, the data symbols are arranged in a time-frequency grid as shown in Fig. 1a, and channel estimation based on such structure is typically performed over two steps. First, initial channel estimates are obtained at the positions of pilot symbols using least square estimation (LSE),

$$\hat{\mathcal{H}}_k = \frac{r_k}{A_k}, \, k \in \mathbb{V} \tag{2.6}$$

where the pilot symbols values are assumed to be known at the receiver side. By noting that $r_k = H_k A_k + w_k$, the CFR estimates can be written as

$$\hat{\mathcal{H}}_k = H_k + q_k \tag{2.7}$$

where $H_k \sim \mathcal{CN}\left(0, 2\sigma_H^2\right)$ and $q_k \sim \mathcal{CN}\left(0, \frac{2\sigma_w^2}{|A_k|^2}\right)$.

Once the initial CFR $\hat{\mathcal{H}}_k \, \forall k \in \mathbb{V}$ is obtained, the CFR $\hat{H}_k(\ell) \, \forall k, \ell$ can be obtained as well given that the pilot grid density satisfies the two-dimensional (2-D) sampling theorem. Then, optimal interpolation using a 2-D Wiener filter that exploits the time and frequency correlation of the channel can be invoked at the expense of substantial implementation complexity [32]. The complexity can be reduced by decomposing the 2-D interpolation process into two cascaded one dimensional (1-D) processes, and then use less computationally involved interpolation schemes [33], [34]. The channel estimation in 1-D can be obtained using various techniques such as linear interpolation [13], parametric estimation [35], or least-square-fitting [36]. It is worth noting that when nonlinear interpolation is invoked, the initial channel estimates at the pilots positions $\hat{\mathcal{H}}_k$ will be replaced by new estimates obtained from the fitting polynomial, and thus $\hat{H}_k \neq \hat{\mathcal{H}}_k$ [36]. Furthermore, in the special case where A_k belongs to CM constellation, it is sufficient to know the phase of the CFR to perform coherent MLD, which can be expressed as

$$\hat{A}_k = \arg \min_{A_k^{(i)}, i \in \mathbb{M}_P} \left| e^{-j\hat{\theta}_k} r_k - A_k^{(i)} \right|^2 \tag{2.8}$$

where $\hat{\theta}_k \triangleq \arg\left\{\hat{H}_k\right\}$ is the estimated version of $\theta_k \triangleq \arg\{H_k\}$.
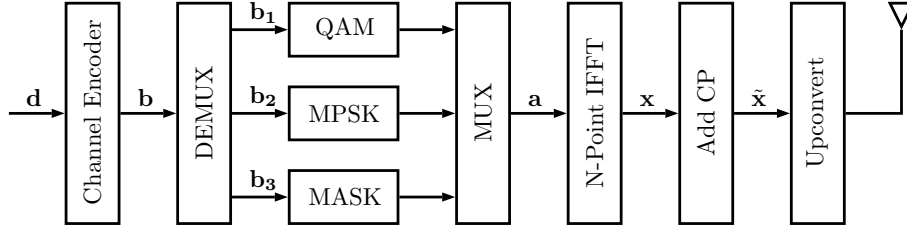
Figure 2.3: Block diagram of the proposed transmitter.
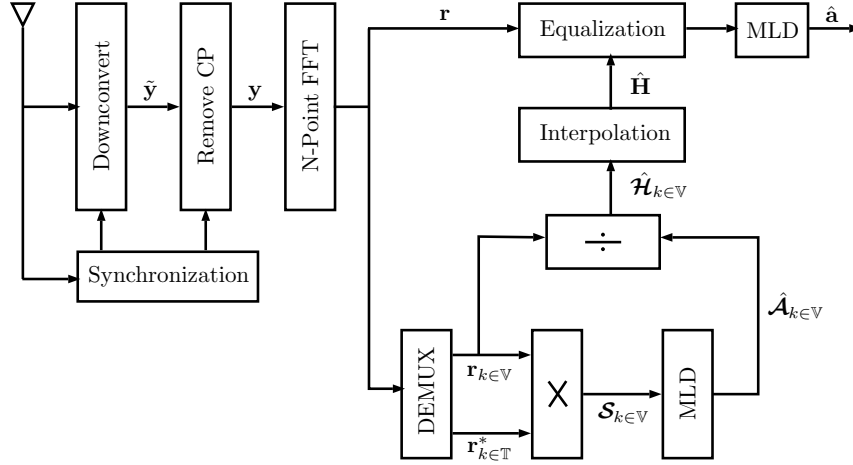
Figure 2.4: Block diagram of the proposed receiver.

2.3   The Proposed OSBCE

For most practical channel models, as indicated in Table 2.1, it can be noted that $|\varrho| \approx 1$ and $\arg\{\varrho\} \approx 0$, which implies that $H_k \approx H_{k+1}$. Therefore, the FFT output at subcarriers k and $k+1$ can be written as

$$r_k = H_k A_k + w_k \tag{2.9}$$

and

$$r_{k+1} \approx H_k A_{k+1} + w_{k+1}. \tag{2.10}$$

Moreover, the AWGN at high SNRs can be ignored, and hence (2.10) can be further simplified to $r_{k+1} \approx H_k A_{k+1}$. Consequently, by carefully choosing the modulation types for the data symbols A_k and A_{k+1}, the information symbol A_k can be

Table 2.1: Correlation coefficient ϱ for common channel models ($N = 256$).

Channel model	Channel Profile	$|\varrho|$	$\arg\{\varrho\}$
cost207RAx4	Rural Area (RAx), 4 taps	0.99998	0.00169
cost207TUx6	Typical Urban (TUx), 6 taps	0.99475	0.05616
cost207TUx12	Typical Urban (TUx), 12 taps	0.99556	0.07981
cost207BUx6	Bad Urban (BUx), 6 taps	0.97447	0.19145

recovered blindly without explicit knowledge of H_k. Towards this goal, consider the case where A_{k+1} is modulated using unipolar MASK, $A_{k+1} \in \mathbb{R}^+$, the set of positive real numbers excluding zero, and A_k is modulated using MPSK. Consequently, MLD of A_k requires only the knowledge of $\hat{\theta}_k$, which can be obtained by noting that $\arg\{r_{k+1}\} \triangleq \hat{\vartheta}_{k+1} \approx \theta_k$. Therefore, the preliminary estimate of the CFR phase $\hat{\vartheta}_{k+1}$ can be used to obtain a preliminary estimate of the symbol A_k, where

$$\hat{\mathcal{A}}_k = \arg \min_{A_k^{(i)}, i \in \mathbb{M}_P} \left| e^{-j\hat{\vartheta}_{k+1}} r_k - A_k^{(i)} \right|^2. \tag{2.11}$$

Furthermore, because $A_{k+1} \in \mathbb{R}^+$ and A_k has CM, the MLD in (2.11) can be also implemented as

$$\hat{\mathcal{A}}_k = \arg \min_{A_k^{(i)}, i \in \mathbb{M}_P} \left| r_k r_{k+1}^* - A_k^{(i)} \right|^2. \tag{2.12}$$

Once $\hat{\mathcal{A}}_k$ is obtained, we can compute $\hat{\mathcal{H}}_k$ in a DD fashion as described in (3.6),

$$\hat{\mathcal{H}}_k = \frac{r_k}{\hat{\mathcal{A}}_k}. \tag{2.13}$$

Therefore, the proposed technique can be implemented by replacing the pilot symbols with data symbols that have CM, and using MASK to modulate the adjacent subcarriers, $A_{k+1} \in \mathbb{R}^+$. Finally, $\hat{H}_k$ can be obtained from $\hat{\mathcal{H}}_k$ using any technique that is originally used in conjunction with pilot-based systems [13], [35], [36]. The proposed channel estimator transmitter and receiver block diagrams are depicted in Figs. 2 and 3, respectively. At the transmitter, the information bits are applied to a channel encoder, which is an optional function, then the encoded bits are split into

three parallel streams each of which is modulated using the corresponding modulation scheme. The three types of symbols are combined to form one block, and the rest of the process is similar to conventional OFDM transmission.

As compared to other practical OFDM-based systems such as LTE-A, it can be noted that the proposed transmitter is generally similar to such systems except that the pilot subcarriers are replaced with MPSK information symbols and the adjacent subcarriers are modulated using MASK as shown in Fig. 1b. It is worth noting that pilot symbols in LTE-A are originally modulated using QPSK, however, they do not bear information. Moreover, the symbols adjacent to the pilots can be modulated using QPSK, 16 or 64 QAM, which does not allow extracting the channel information directly from such modulation schemes. Therefore, MASK is used to estimate the channel phase directly from the received symbols, and hence allow coherent detection of the QPSK symbols in the pilots locations.

The MASK modulated symbols A_{k+1} can be designed such that the average power is normalized to unity, $\mathcal{P}_s = \frac{1}{M_A} \sum_{i=0}^{M_A-1} \left(A_{k+1}^{(i)} \right)^2 = 1$. Assuming equally spaced amplitudes, the amplitude spacing $\delta \triangleq A_{k+1}^{(i+1)} - A_{k+1}^{(i)}$ [24] can be written as

$$A_{k+1}^{(i)} = (i+1) \times \delta, \ i \in \{0, 1, \ldots, M_A - 1\} \tag{2.14}$$

where,

$$\delta = \sqrt{\frac{6}{(2M_A + 1)(M_A + 1)}}. \tag{2.15}$$

For all other subcarriers other than the pilots' and their adjacent subcarriers, the modulation type and order can be chosen arbitrarily. Moreover, although the OSBCE is applied to LTE-A, it can be applied to other OFDM-based systems where the frequency spacing Δ_F and time spacing Δ_T can be changed based on the corresponding system specifications.

As it can be noted from the system description, the computational complexity

of the OSBCE is equivalent to LSE using pilots. Particularly, when considering that $\hat{\mathcal{A}}_k = \hat{A}_k$, which is mainly the case because for most interpolation techniques $\hat{H}_k \approx \hat{\mathcal{H}}_k$. In the case that a new detection is applied after the estimation of $\hat{\mathbf{H}}$, then the only additional complexity, as compared to pilot systems, is N_P complex multiplications and MLD detections applied to the N_P MPSK symbols. Therefore, the proposed OSBCE complexity is generally low, and it is equivalent to pilot-based channel estimators.

2.4 Symbol Error Rate (SER) Analysis

Because the OSBCE operates as a DD estimator, its MSE depends on the initial SER, $\Pr\left(\hat{\mathcal{A}}_k \neq A_k\right)$. Therefore, this section is dedicated to the derivation of the initial SER, then the MSE derivation is presented in the next section.

As it can be noted from (3.11), eliminating the fading channel effect using the product $r_k r_{k+1}^*$ is equivalent to conventional equalization of MPSK symbols, which is independent of the CFR magnitude $|H_k|$. However, the equalized samples can be also written as $r_k / r_{k+1} \triangleq \mathcal{S}_k$, where

$$
\begin{aligned}
\mathcal{S}_k &= \frac{H_k A_k + w_k}{H_{k+1} A_{k+1} + w_{k+1}} \\
&= \frac{A_k \left|H_k\right| e^{j(\theta_k - \hat{\vartheta}_{k+1})} + \left|w_k\right| e^{j\left(\arg(w_k) - \hat{\vartheta}_{k+1}\right)}}{\left|H_{k+1} A_{k+1} + w_{k+1}\right|} \\
&= \frac{1}{\left|r_{k+1}\right|} \left(\left|H_k\right| e^{j\psi_k} A_k + \acute{w}_k\right)
\end{aligned}
\tag{2.16}
$$

where $\psi_k = \theta_k - \hat{\vartheta}_{k+1}$, $\acute{w}_k = \left|w_k\right| e^{j\left(\arg(w_k) - \hat{\vartheta}_{k+1}\right)}$ is a zero-mean complex Gaussian random variable whose variance is the same as that of w_k. Therefore, the equalized sample $\mathcal{S}_k$ in (2.16) has the form of MPSK symbol equalized using imperfect channel estimate. However, the conditional BER of BPSK and QPSK modulations in fading channels with imperfect knowledge of the CFR parameters can be expressed as [26,

10.14a]

$$P_{B,k}|\left[\psi_k, \alpha_k\right] = Q\left(\frac{\alpha_k}{\sigma_w}\cos(\psi_k)\right) \tag{2.17}$$

and

$$P_{B,k}|\left[\psi_k, \alpha_k\right] = \frac{1}{2}Q\left(\frac{\alpha_k}{\sqrt{2}\sigma_w}\left[\cos(\psi_k) - \sin(\psi_k)\right]\right)$$

$$+ \frac{1}{2}Q\left(\frac{\alpha_k}{\sqrt{2}\sigma_w}\left[\cos(\psi_k) + \sin(\psi_k)\right]\right) \tag{2.18}$$

where $\alpha_k = |H_k|$ is the fading envelope of the kth subcarrier, and $Q(x) \triangleq 1/\sqrt{2\pi}\int_x^\infty e^{-t^2/2}dt$. Consequently, the conditioning on α_k and ψ_k can be eliminated by averaging (3.20) and (2.18) over the joint probability density function (PDF) $P(\psi_k, \alpha_k)$. To simplify the solution, we initially assume that the MASK symbol A_{k+1} is fixed, and hence, the conditional BER can be computed as

$$P_{B,k} = \int_0^\infty \int_{-\pi}^\pi P_{B,k}|\left[\psi_k, \alpha_k\right] \; P(\alpha_k, \psi_k) \; d\psi_k \; d\alpha_k. \tag{2.19}$$

Although A_{k+1} is not written explicitly in (3.20), (2.18) and the joint PDF $P(\alpha_k, \psi_k)$, it is actually included in the conditioning on ψ_k, because $\psi_k = \theta_k - \arg\left\{H_{k+1}A_{k+1} + w_{k+1}\right\}$.

The joint PDF $P(\alpha_k, \psi_k)$ can be obtained by following the approach reported in [28], where r_k and r_{k+1} are expressed as

$$r_k = H_k A_k + w_k, \;\; H_k \triangleq x_1 + jx_2 = \alpha_k e^{j\theta_k} \tag{2.20}$$

$$r_{k+1} = H_{k+1}A_{k+1} + w_{k+1} \triangleq x_3 + jx_4 = \beta_k e^{j\vartheta_k}. \tag{2.21}$$

The random variables $\{x_1,\, x_2,\, x_3,\, x_4\}$ are all zero-mean Gaussian random variables, where x_1 and x_2 are independent, x_3 and x_4 are independent as well, thus $\mathrm{E}\left[x_1 x_2\right] = \mathrm{E}\left[x_3 x_4\right] = 0$. Therefore, the joint PDF $P(x_1,\, x_2,\, x_3,\, x_4)$ can be described

by

$$P(x_1, x_2, x_3, x_4) = \frac{1}{4\pi^2(\sigma_1^2\sigma_2^2 - \mu_1^2 - \mu_2^2)}$$
$$\times \exp\left[-\frac{\sigma_2^2(x_1^2 + x_2^2) + \sigma_1^2(x_3^2 + x_4^2) - 2\mu_1(x_1x_3 + x_2x_4) - 2\mu_2(x_1x_4 - x_2x_3)}{2(\sigma_1^2\sigma_2^2 - \mu_1^2 - \mu_2^2)}\right].$$
(2.22)

where σ_1, σ_2, μ_1, and μ_2 are given by [35]

$$\sigma_1^2 \triangleq \mathrm{E}\left[x_1^2\right] = \mathrm{E}\left[x_2^2\right] = \frac{1}{2}\mathrm{E}\left[|H_k|^2\right] = \sigma_H^2 \tag{2.23}$$

$$\sigma_2^2 \triangleq \mathrm{E}\left[x_3^2\right] = \mathrm{E}\left[x_4^2\right] = \frac{1}{2}\mathrm{E}\left[|r_{k+1}|^2\right]$$

$$= \frac{1}{2}\mathrm{E}\left[|H_{k+1}A_{k+1} + w_{k+1}|^2\right]$$

$$= \sigma_H^2|A_{k+1}|^2 + \sigma_w^2 \tag{2.24}$$

$$\mu_1 + j\,\mu_2 \triangleq \frac{1}{2}\left\{\mathrm{E}\left[(r_{k+1})\,H_k^*\right]\right\}$$

$$= A_{k+1}\sum_{n=0}^{L_h}\sigma_{h_n}^2\cos\left(2\pi\frac{n}{N}\right)$$

$$+jA_{k+1}\sum_{n=0}^{L_h}\sigma_{h_n}^2\sin\left(2\pi\frac{n}{N}\right). \tag{2.25}$$

By making the transformation from rectangular (x_1, x_2, x_3, x_4) to polar coordinates $(\alpha_k, \beta_k, \theta_k, \hat{\vartheta}_k)$, and applying the following change of variables

$$x_1 = \alpha_k\cos(\theta_k),\ x_2 = \alpha_k\sin(\theta_k)$$

$$x_3 = \beta_k\cos(\hat{\vartheta}_k),\ x_4 = \beta_k\sin(\hat{\vartheta}_k),\ \psi_k = \theta_k - \hat{\vartheta}_k \tag{2.26}$$

the following joint PDF can be obtained

$$
P(\alpha_k, \psi_k) = \int_0^\infty \frac{\alpha_k \beta_k}{2\pi \sigma_1^2 \sigma_2^2 (1 - \rho^2)} \exp\left\{ - \frac{1}{2(1 - \rho^2)} \right.
$$
$$
\left. \times \left[\frac{\alpha_k^2}{\sigma_1^2} + \frac{\beta_k^2}{\sigma_2^2} - 2\frac{\alpha_k \beta_k}{\sigma_1 \sigma_2} \left(\rho_1 \cos(\psi_k) - \rho_2 \sin(\psi_k) \right) \right] \right\} d\beta_k \quad (2.27)
$$

where the correlation coefficients are defined by

$$
\rho_1 \triangleq \frac{\mu_1}{\sigma_1 \sigma_2}, \ \rho_2 \triangleq \frac{\mu_2}{\sigma_1 \sigma_2}, \ \rho \triangleq \sqrt{\rho_1^2 + \rho_2^2}. \quad (2.28)
$$

Substituting (3.20), (2.18) and (3.15) into (2.19), and using the integral identity established in [28, Appendix B], we get the closed-form BER conditioned on A_{k+1} for BPSK as

$$
P_{B,k} = \frac{1}{2} \left[1 - \frac{\rho_1}{\sqrt{1 + \frac{1}{\text{SNR}} - \rho_2^2}} \right] \quad (2.29)
$$

and for QPSK

$$
P_{B,k} = \frac{1}{2} \left[1 - \frac{1}{2} \frac{\rho_1 + \rho_2}{\sqrt{2 + \frac{1}{\text{SNR}} - (\rho_1 - \rho_2)^2}} + \frac{\rho_1 - \rho_2}{\sqrt{2 + \frac{1}{\text{SNR}} - (\rho_1 + \rho_2)^2}} \right] \quad (2.30)
$$

It is worth noting that the SNR per bit γ_b is defined as

$$
\gamma_b \triangleq \frac{\text{E}\left[|H_k A_k|^2\right]}{\log_2 (M_P) \ \text{E}\left[|w_k|^2\right]}
$$
$$
= \frac{\mathcal{P}_s}{\sigma_w^2} \frac{\sigma_1^2}{\log_2 (M_P)}
$$
$$
= \frac{\text{SNR}}{\log_2 (M_P)}. \quad (2.31)
$$

Given that the average symbol power $\mathcal{P}_s = \text{E}\left[|A_k|^2\right] = 1$, the SNR per bit γ_b can be written as

$$\gamma_b \triangleq \frac{\sigma_1^2}{\log_2{(M_P)}\ \sigma_w^2}. \tag{2.32}$$

The parameters σ_2, μ_1 and μ_2 are functions of the MASK symbol A_{k+1}, and hence the average unconditioned BER can be obtained by averaging (3.22) and (2.30) over all possible values of A_{k+1}. Since $A_{k+1}^{(i)}$ are equally probable, the average $\bar{P}_{B,k}$ becomes,

$$\bar{P}_{B,k} = \frac{1}{M_A} \sum_{i=0}^{M_A-1} \left(P_{B,k}|A_{k+1}^{(i)} \right). \tag{2.33}$$

Moreover, assuming Gray coding, the SER for QPSK is given by [39, 8.7]

$$\bar{P}_{S,k} = 2\bar{P}_{B,k}. \tag{2.34}$$

2.5 Mean Squared Error (MSE) Analysis

The MSE of the initial CFR estimate $\hat{\mathcal{H}}_k$ is given by

$$\mathrm{MSE}(\hat{\mathcal{H}}_k) = \mathrm{E}\left[\left| \hat{\mathcal{H}}_k - H_k \right|^2 \right] \tag{2.35}$$

where

$$\hat{\mathcal{H}}_k = \frac{r_k}{\hat{\mathcal{A}}_k} = H_k \frac{A_k}{\hat{\mathcal{A}}_k} + \frac{w_k}{\hat{\mathcal{A}}_k}. \tag{2.36}$$

To simplify the notations, $\mathrm{MSE}(\hat{\mathcal{H}}_k)$ is written as MSE, unless it is necessary to write the full expression. Substituting (2.36) into (2.35) gives

$$\mathrm{MSE} = \mathrm{E}\left[\left| H_k \frac{A_k}{\hat{\mathcal{A}}_k} + \frac{w_k}{\hat{\mathcal{A}}_k} - H_k \right|^2 \right]. \tag{2.37}$$

Because $\hat{\mathcal{A}}_k$ may take one of two states, the law of total probability can be used to decompose the MSE as the sum of two conditional scenarios

$$\mathrm{MSE} = (\mathrm{MSE}|D_C)\Pr(D_C) + (\mathrm{MSE}|D_I)\Pr(D_I) \tag{2.38}$$

where the events of correct and incorrect decisions are denoted by D_C and D_I, respectively, and $\Pr(D_I) = \Pr\left(\hat{\mathcal{A}}_k \neq A_k\right) = P_{S,k}$ and $\Pr(D_C) = \Pr\left(\hat{\mathcal{A}}_k = A_k\right) = P_{C,k} = 1 - P_{S,k}$. To simplify the presentation, the two conditional cases are presented into the following two subsections.

2.5.1   MSE $\left(\hat{\mathcal{H}}_k | D_C\right)$

For the case of $\hat{\mathcal{A}}_k = A_k$, the conditional initial channel estimate can be expressed as

$$\hat{\mathcal{H}}_k | D_C = H_k | D_C + \frac{w_{k|D_C}}{A_k}. \tag{2.39}$$

Therefore, the conditional MSE can be computed as

$$\mathrm{MSE} | D_C = \mathrm{E}\left[\left|\left(\hat{\mathcal{H}}_k | D_C\right) - (H_k | D_C)\right|^2\right]. \tag{2.40}$$

It is worth noting that the conditioning on the right hand side of (2.40) is necessary due to the correlation between the estimated MPSK symbol $\hat{\mathcal{A}}_k$, H_k and the AWGN samples w_k, which is due to the fact that the event $\hat{\mathcal{A}}_k = A_k$ is generally realized at high H_k values and low w_k values. Therefore, substituting (2.39) in (2.40) gives

$$\mathrm{MSE} | D_C = \mathrm{E}\left[\left|\frac{w_k | D_C}{A_k}\right|^2\right]. \tag{2.41}$$

The result in (2.41) is expected because the process is generally similar to the conventional LSE. However, $w_k | D_C$ is a sampled version of w_k because the conditioning process on D_C eliminates most of the high power noise samples, and hence, $\mathrm{E}\left[|w_k | D_C|^2\right] \neq \mathrm{E}\left[|w_k|^2\right]$. Fig. 4 shows the AWGN variance conditioned on D_C versus SNR as compared to the unconditional AWGN. The channel is modeled as a typical urban (TUx) multipath channel [45], which is described in Section 5.7. It can be noticed from the figure that $\mathrm{E}\left[|w_k | D_C|^2\right] < \mathrm{E}\left[|w_k|^2\right]$ at low SNRs, and they converge for SNR $\gtrsim$ 10 dB. Such behavior is due to the fact that the effect of noise sampling disappears at high SNRs because the channel fading is the dominant pa-

rameter that determines the outcome of the detection process. Therefore, it can be assumed without sacrificing the accuracy that

$$
\begin{aligned}
\text{MSE}|D_C &= \mathrm{E}\left[\left|\frac{w_k \mid D_C}{A_k}\right|^2\right] \\
&\approx \mathrm{E}\left[\left|\frac{w_k}{A_k}\right|^2\right] \\
&= \frac{1}{\text{SNR}}.
\end{aligned}
\tag{2.42}
$$

2.5.2 MSE $\left(\hat{\mathcal{H}}_k|D_I\right)$

For the case where $\hat{\mathcal{A}}_k \neq A_k$, the MSE can be obtained by substituting

$$
\hat{\mathcal{H}}_k|D_I = (H_k|D_I)\frac{A_k}{\hat{\mathcal{A}}_k} + \frac{w_k|D_I}{\hat{\mathcal{A}}_k}
\tag{2.43}
$$

in (2.35), which gives

$$
\text{MSE}|D_I = \mathrm{E}\left[\left|(H_k \mid D_I)\frac{A_k}{\hat{\mathcal{A}}_k} + \frac{w_k|D_I}{\hat{\mathcal{A}}_k} - (H_k \mid D_I)\right|^2\right].
\tag{2.44}
$$

By assuming that all transmitted symbols are equiprobable [42], we assume, without loss of generality, that the MPSK symbol $A_k^{(0)} = e^{j\pi/M_P}$ is transmitted, and hence, $\hat{\mathcal{A}}_k = e^{j\pi(2i+1)/M_P}$, $i \in \{1, 2, \ldots, M_P - 1\}$. By defining $\varphi_k \triangleq \arg\left\{A_k^{(0)}/\hat{\mathcal{A}}_k\right\}$, then (2.44), after some straightforward manipulations becomes

$$
\text{MSE}|D_I = \mathrm{E}\left[\left|\frac{w_k \mid D_I}{\hat{\mathcal{A}}_k}\right|^2\right] + 2\mathrm{E}\left[|H_k|^2 \mid D_I\right]\left(1 - E\left[\cos(\varphi_k)|D_I\right]\right).
\tag{2.45}
$$

Contrary to the D_C case, the conditional noise variance $E\left\{|w_k|D_I|^2\right\}$ is slightly higher than the unconditional noise variance $\mathrm{E}\left[|w_k|^2\right]$ at low SNRs, then the difference disappears for SNR $\gtrsim$ 40 dB as shown in Fig. 4. Therefore, $\mathrm{E}\left[|w_k|D_I|^2\right] \approx \mathrm{E}\left[|w_k|^2\right] = \frac{1}{\text{SNR}}$. Similar to the AWGN case, the incorrect decision event D_I and

the CFR H_k are correlated as well. Since most error events occur in deep fading conditions, then $\mathrm{E}\left[\alpha_k^2|D_I\right] \ll \mathrm{E}\left[\alpha_k^2\right]$. Therefore, to evaluate $\mathrm{MSE}|D_I$ in (2.45), the terms $\mathrm{E}\left[\cos(\varphi_k)|D_I\right]$ and $\mathrm{E}\left[\alpha_k^2|D_I\right]$ should be evaluated first.

- $\mathrm{E}\left[\cos(\varphi_k)|D_I\right]$:

Because φ_k is a discrete random variable, the conditional expectation can be expressed as

$$\mathrm{E}\left[\cos(\varphi_k)|D_I\right] = \sum_{\varphi_k} \mathrm{Pr}\left(\varphi_k|D_I\right)\cos(\varphi_k) \tag{2.46}$$

where $\varphi_k = \frac{-2\pi i}{M_P}$, $i \in \{1,2,\ldots,M_P-1\}$. For example, $\mathrm{Pr}\left(\varphi_k = \frac{-2\pi}{M_P}\right) = \mathrm{Pr}\left(\hat{\mathcal{A}}_k = A_k^{(1)}\right)$. Recalling Bayes' theorem with mixed distributions, the conditional probability $\mathrm{Pr}\left(\varphi_k|D_I\right)$ can be evaluated as

$$\mathrm{Pr}\left(\varphi_k|D_I\right) = \frac{1}{\mathrm{Pr}\left(D_I\right)} \mathrm{Pr}\left(D_I|\varphi_k\right)\mathrm{Pr}\left(\varphi_k\right). \tag{2.47}$$

By noting the definition of φ_k and the fact that the transmitted symbol is $A_k^{(0)}$, then $\mathrm{Pr}\left(D_I|\varphi_k\right) = 1$, and hence,

$$\mathrm{Pr}\left(\varphi_k|D_I\right) = \frac{1}{P_{S,k}} \mathrm{Pr}\left(\varphi_k\right) \tag{2.48}$$

where

$$\mathrm{Pr}\left(\varphi_k\right) = \int_{\frac{2\pi i}{M_P}}^{\frac{2\pi(i+1)}{M_P}} P(\Theta_k)d\Theta_k,\ i \in \{1,2,...,M_P-1\} \tag{2.49}$$

and

$$\Theta_k = \arg\left\{r_{k+1}^* r_k\right\}. \tag{2.50}$$

However, at high SNRs $\Theta_k \approx \arg\{A_k\} + \psi_k$, and since it is assumed that $A_k = A_k^{(0)} = e^{j\pi/M_P}$, then $\arg\{A_k\} = \pi/M_P$. Therefore (2.49) can be expressed as

$$\Pr(\varphi_k) = \int_{\frac{\pi(2i-1)}{M_P}}^{\frac{\pi(2i+1)}{M_P}} P(\psi_k) d\psi_k, \ i \in \{1, 2, ..., M_P - 1\} \tag{2.51}$$

where $P(\psi_k)$, which can be obtained from (3.15),

$$\begin{aligned}
P(\psi_k) &= \int_0^\infty P(\alpha_k, \psi_k) \ d\beta_k \ d\alpha_k \\
&= \frac{1 - \rho^2}{2\pi \left[1 - \zeta_k^2\right]^{\frac{3}{2}}} \zeta_k \ \arctan\left(\frac{\zeta_k}{\sqrt{1 - \zeta_k^2}}\right) \\
&\quad + 2\sqrt{1 - \zeta_k^2} + \pi \zeta_k
\end{aligned} \tag{2.52}$$

and $\zeta_k = \rho_1 \cos(\psi_k) + \rho_2 \sin(\psi_k)$. Finally, $\Pr(\varphi_k | D_I)$ is given by

$$\Pr(\varphi_k | D_I) = \frac{1}{P_{S,k}} \int_{\frac{\pi(2i-1)}{M_P}}^{\frac{\pi(2i+1)}{M_P}} P(\psi_k) d\psi_k. \tag{2.53}$$

To the best of our knowledge, the integral in (2.53) has no closed form expression, and hence, it will be evaluated numerically.

- $\mathrm{E}\left[\alpha_k^2 | D_I\right]$:

The conditional variance of H_k can be computed as

$$\mathrm{E}\left[\alpha_k^2 | D_I\right] = \int_0^\infty \alpha_k^2 P(\alpha_k | D_I) d\alpha_k \tag{2.54}$$

Using Bayes' rule with mixed distributions, the conditional PDF $P(\alpha_k | D_I)$ can be evaluated as

$$\begin{aligned}
P(\alpha_k | D_I) &= \frac{\Pr(D_I | \alpha_k) P(\alpha_k)}{\Pr(D_I)} \\
&= \frac{(P_{S,k} | \alpha_k) P(\alpha_k)}{P_{S,k}}.
\end{aligned} \tag{2.55}$$

By noting that the $P_{S,k} = P_{B,k}$ for the BPSK and using (2.34) for QPSK, then $P(\alpha_k | D_I)$ which is given by (2.55) can be simplified to

$$P(\alpha_k|D_I) = \frac{\log_2(M_P)}{P_{S,k}} (P_{B,k}|\alpha_k) P(\alpha_k). \tag{2.56}$$

Using the joint PDF $P(\alpha_k, \psi_k)$ in (3.15), and the conditional BER for BPSK (3.20) or QPSK (2.18), then $\mathrm{E}\left[(\alpha_k^2|D_I)\right]$ can be computed as depicted in (2.57),

$$\mathrm{E}\left[\alpha_k^2|D_I\right] = \frac{\log_2(M_P)}{P_{S,k}} \int_0^\infty \int_0^\infty \int_{-\pi}^\pi \alpha_k^2 (P_{B,k}|\psi_k, \alpha_k) \ P(\alpha_k, \beta_k, \psi_k)d\psi_k \ d\beta_k \ d\alpha_k \tag{2.57}$$

It is worth mentioning that evaluating the expression in (2.57) analytically is intractable. However, the expression in (2.57) can be computed numerically. Finally, the $\mathrm{MSE}(\hat{\mathcal{H}}_k)$ can be written as

$$\begin{aligned} \mathrm{MSE}(\hat{\mathcal{H}}_k) &= (\mathrm{MSE}|D_C)P_{C,k} + (\mathrm{MSE}|D_I)P_{S,k} \\ &= \frac{1}{\mathrm{SNR}} + 2P_{S,k}\mathrm{E}\left[\alpha_k^2|D_I\right](1 - \mathrm{E}\left[\cos(\varphi_k)|D_I\right]). \end{aligned} \tag{2.58}$$

By noting that most of the terms in (2.58) are dependent on A_{k+1}, then the unconditional MSE can be expressed as

$$\overline{\mathrm{MSE}}(\hat{\mathcal{H}}_k) = \frac{1}{M_A} \sum_{i=0}^{M_A-1} \left(\mathrm{MSE}(\hat{\mathcal{H}}_k) \mid A_{k+1}^{(i)}\right). \tag{2.59}$$

Moreover, by noting that $\mathrm{E}\left[\alpha_k^2|D_I\right] \le 2\sigma_H^2$, and noting that $-1 \le \cos(\varphi_k) \le 1$, and hence an upper and lower bounds can be expressed as

$$\frac{1}{\mathrm{SNR}} \le \mathrm{MSE} \le \frac{1}{\mathrm{SNR}} + 8\sigma_H^2 \bar{P}_{S,k}. \tag{2.60}$$

2.6   Spectral Efficiency

In this section, the spectral efficiency of the OSBCE is evaluated in terms of the throughput, which is defined as the average number of information bits per subcarrier. In the most general case, each subcarrier in an OFDM grid can be modulated using

a certain modulation scheme with a particular modulation order. Therefore, the average number of information bits over one OFDM grid can be expressed as

$$\eta = \frac{1}{N_T \times N_F} \sum_{\ell=0}^{N_T-1} \sum_{k=0}^{N_F-1} \log_2 \left[M\left(\ell, k\right) \right]. \tag{2.61}$$

where $M\left(\ell, k\right)$ is the modulation order at a given location $\left(\ell, k\right)$, N_T and N_F are the time and frequency dimensions of the OFDM grid, respectively. Fig. 1a shows an example where $N_T = 14$ and $N_F = 12$.

In practical systems, the value of $M\left(\ell, k\right)$ is dynamically selected based on the system QoS requirements, the system resources, modulation order constrains, and the channel state information (CSI) [27]. Without loss of generality, consider the case where the values of $M\left(\ell, k\right)$ can be selected dynamically with the aim of maximizing the throughput of a particular system under BER, and modulation type/order constraints. Therefore, the problem can be formulated as

$$\max_{M(\ell,k)\in\mathbb{M}} \frac{1}{N_T \times N_F} \sum_{\ell=0}^{N_T-1} \sum_{k=0}^{N_F-1} \log_2 \left[M\left(\ell, k\right) \right] \tag{2.62a}$$

Subject to:

$$\bar{P}_B \leq P_{th} \tag{2.62b}$$

where $\mathbb{M}$ is the set of all possible modulation orders, and (4.6c) is used to guarantee that the average BER $\bar{P}_B$ is less than a prescribed threshold P_{th},

$$\bar{P}_B = \frac{\sum_{\ell=0}^{N_T-1} \sum_{k=0}^{N_F-1} \log_2 \left(M\left(\ell, k\right) \right) P_B \left(\gamma_{\ell,k} \,|_{M,T} \right)}{\sum_{\ell=0}^{N_T-1} \sum_{k=0}^{N_F-1} \log_2 \left(M\left(\ell, k\right) \right)} \leq P_{th} \tag{2.63}$$

where $P_B \left(\gamma_{\ell,k} \,|_{M,T} \right)$ is the instantaneous BER for a given value of M and modulation type T, and the instantaneous SNR is denoted by $\gamma_{\ell,k} = |A_{k,\ell}|^2 \cdot |H_{k,\ell}|^2 / \left(2\sigma_w^2 \right)$. Assuming that the possible modulation types are QAM, MPSK and MASK, then $P_B \left(\gamma_{\ell,k} \,|_{M_Q, \text{QAM}} \right)$ and $P_B \left(\gamma_{\ell,k} \,|_{M_P, \text{MPSK}} \right)$ are given in [35], while $P_B \left(\gamma_{\ell,k} \,|_{M_A, \text{MASK}} \right)$ under the Gray coding assumption is given by

$$P_B\left(\gamma_{\ell,k}\,|_{M_A,\,\text{MASK}}\right) = \frac{2\left(M_A - 1\right)}{M_A\,\log_2\left(M_A\right)}Q\left(\sqrt{\frac{3\,\gamma_{k,\ell}}{4M_A^2 + 6M_A + 2}}\right). \tag{2.64}$$

It is worth noting that P_B for the considered modulation schemes is computed while assuming perfect knowledge of the instantaneous SNRs $\gamma_{\ell,k}$. Therefore, the obtained throughput in (2.62) can be considered an upper bound because channel estimation errors may degrade the SNRs, and hence, reduce the throughput for all the considered techniques.

For a fair comparison between various systems, the power efficiency should be also taken into consideration. Therefore, the power per information subcarrier in the blind system $P_{SC,blind}$ should be equal to that in the pilot-based system $P_{SC,pilot}$. Consequently,

$$P_{SC,blind} = \frac{N_F \times N_T}{N_F \times N_T - N_{P,\text{RB}}}P_{SC,pilot} \tag{2.65}$$

where $N_{P,\text{RB}}$ is the number of pilot subcarriers per OFDM grid. However, since the MASK symbols in the OSBCE are the most sensitive symbols to noise, the additional power is allocated to them. Consequently, the power of each subcarrier with MASK modulation will be assigned an additional 3 dB gain.

2.7   Numerical Results

The performance of the proposed OSBCE is evaluated over quasi-static flat and frequency-selective multipath fading channels, where the channel remains fixed over one transmission frame that consists of 7 OFDM symbols, but may vary randomly over consecutive frames. The proposed estimator is compared to pilot-based OFDM systems with the LTE transmission grid shown in Fig. 2.1. The performance of the proposed estimator is evaluated in terms of SER and MSE.
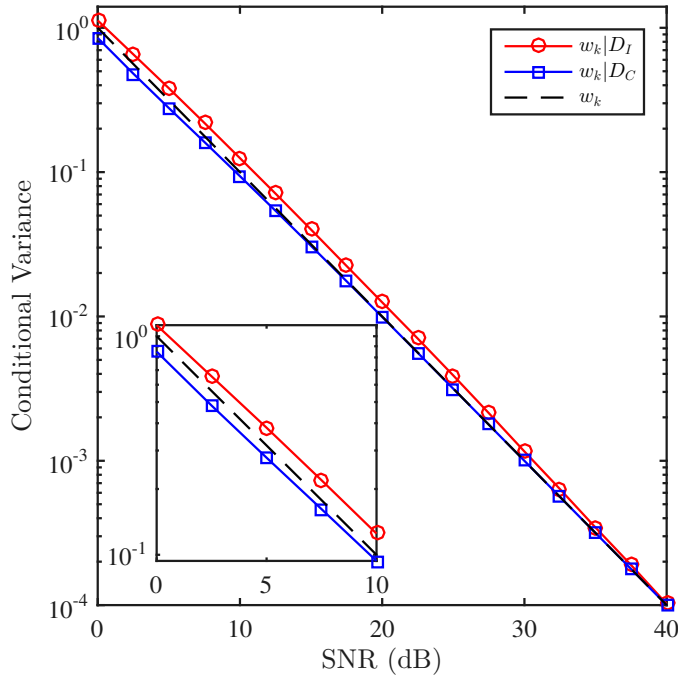
Figure 2.5: Conditional AWGN variance in the case of correct and incorrect decisions compared to the unconditional variance of AWGN for the TUx channel, $M_P = 4$, $M_A = 4$.

The OFDM system considered follows the LTE downlink physical layer specifications [3] where the sampling frequency is 3.836 MHz, $N = 256$ subcarriers, $N_{\mathrm{CP}} = 18$ samples, subcarrier spacing is 15 kHz, total OFDM symbol period is 71.3 μ sec, and CP period is 4.69 μ sec. Two channel models are used, the flat fading and the TUx multipath fading model [45], which consists of 9 taps with normalized delays of [0, 1, ..., 8] samples, and the average taps' gains $2\sigma_{h_n}^2$, $n = [0, 1, ..., 8]$ are [0.269, 0.174, 0.289, 0.117, 0.023, 0.058, 0.036, 0.026, 0.008]. In each simulation run, 30×10^4 OFDM symbols are processed.

The PDF of the phase error ψ_k over the TUx and flat channels is given in Figs. 2.6 and 2.7, respectively, for SNR = 10, 20 and 30 dB given that $A_{k+1} = 1$. As expected, the figures show that the phase-error variance (σ_ψ^2) decreases as SNR increases, which implies that the channel estimates can be improved by increasing SNR. However for the TUx channel, Fig. 2.6 shows that the phase error is biased where $\mathrm{E}\,[\psi_k] = \arg\{\varrho\} \neq 0$. Such bias is due to the fact that $\theta_k \neq \theta_{k+1}$ in frequency-selective
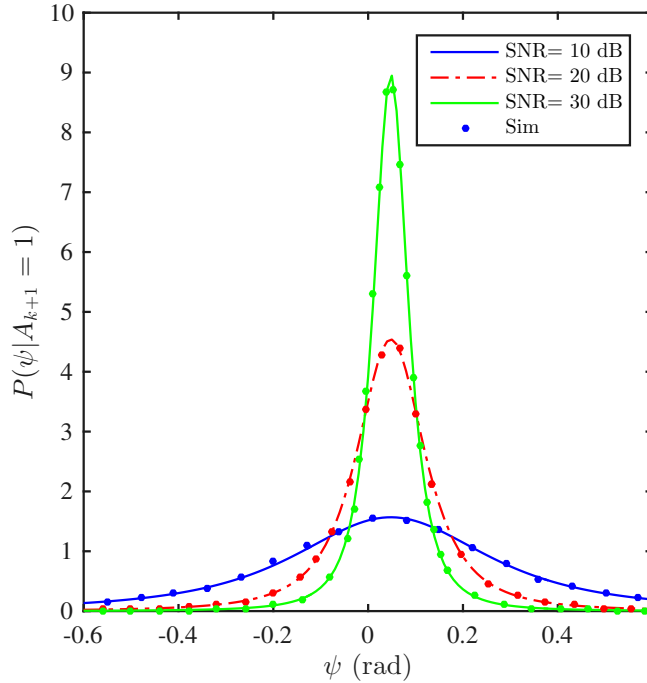
Figure 2.6: $P(\psi)$ for different values of SNR over the TUx channel where $A_{k+1} = 1$, $M_P = 4$, and $M_A = 4$.

channels, and hence, the multiplication process of $r_k r_{k+1}^*$ is equivalent to equalization with biased channel estimates, which causes SER error floors at high SNRs. More-over, the figures show that σ_ψ^2 for the TUx channel is larger than that of the flat fading channel because the instantaneous difference between θ_k and θ_{k+1} is random, and hence contributes to σ_ψ^2. The figures also show the perfect match between the analytical and simulation results.

Fig. 2.8 shows $P(\psi_k)$ for different values of A_{k+1} at SNR $= 20$ dB. It can be noted that ψ_k is dependent on A_{k+1} where the variance decreases for high values of A_{k+1}. Such results imply that minimizing ψ_k can be obtained by maximizing A_{k+1}. However, since the possible values of A_{k+1} will affect $P_{S,k+1}$, then the optimal MASK amplitudes should be selected such that

$$\mathbf{a}_{k+1} = \arg \min_{\mathbf{a}_{k+1} \in \mathbb{R}^+} \sum_{i=0}^{N-1} P_{S,i} \tag{2.66}$$
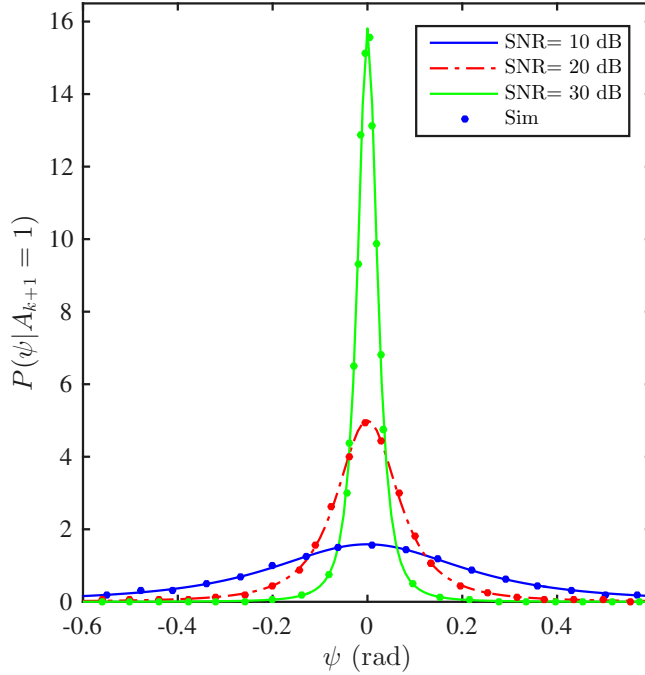
Figure 2.7: $P(\psi)$ for different values of SNR over a flat fading channel where $A_{k+1} = 1$, $M_P = 4$, and $M_A = 4$.

where $\mathbf{a}_{k+1} = \left[A_{k+1}^{(0)},\, A_{k+1}^{(1)},\, ...,\, A_{k+1}^{(M_A-1)} \right]$ and $\frac{1}{M_A} \sum_{i=0}^{M_A-1} \left(A_{k+1}^{(i)} \right)^2 = 1$. However, solving such multi-objective optimization problem is computationally expensive, and hence we consider the equally spaced amplitudes for the purposes of this work.

The initial theoretical and simulated SER, $\Pr(\hat{\mathcal{A}}_k \neq A_k) \ \forall \ k \in \mathbb{V}$, of the MPSK symbols detected using (3.11) is depicted in Fig. 2.9. The results are presented for BPSK and QPSK using $M_A = 4, 8$ and 16, over the TUx and flat fading channels. The figure shows that the theoretical and simulation results match very well. However, a little difference can be observed for the QPSK case due to the approximation made in (2.18) [26], and the Gray coding assumption. Moreover, the impact of the channel frequency selectivity is observed only at high SNRs where an error floor starts to appear at SNR $\gtrsim 25$ dB.

Fig. 2.10 shows the impact of increasing the modulation order M_A, as well as the effect of the frequency selectivity of the channel on the initial SER, $\Pr(\hat{\mathcal{A}}_k \neq A_k)$
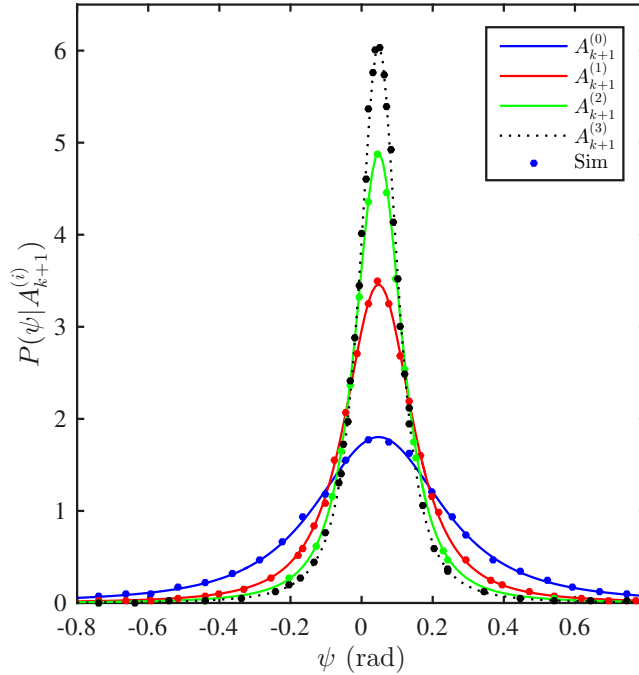
Figure 2.8: $P(\psi)$ for different values of A_{k+1} over the TUx channel, SNR $= 20$ dB, $M_P = 4$, $M_A = 4$.

$\forall\, k \in \mathbb{V}$. The figure shows that SER increases by increasing M_A which can be justified by noting that for $M_{A_1} < M_{A_2}$, $A_{k+1}^{(i)}|M_{A_1} > A_{k+1}^{(i)}|M_{A_2} \,\forall\, i \le \min\{M_{A_1}, M_{A_2}\}$. In other words, increasing M_A will introduce more amplitudes with small values, which is equivalent to equalization using less reliable channel estimates. The figure also shows that at high SNRs, the impact of changing M_A vanishes for the TUx channel, because SER is determined by the frequency selectivity of the channel, which caused the bias of ψ_k. The flat channel kept at the same trends since SER will always be the dominant factor. Nevertheless, the initial SER remains sufficiently low to produce reliable channel estimates.

Fig. 2.11 shows the conditional variance of the CFR for $M_A = 4$, and for $M_P = 2$ and 4. As it can be noted from the figure, the difference between the conditional and unconditional variances is substantial, and it is inversely proportional to $A_{k+1}^{(i)}$, which is due to the fact that higher values of $A_{k+1}^{(i)}$ are more immune to fading, i.e.,

41
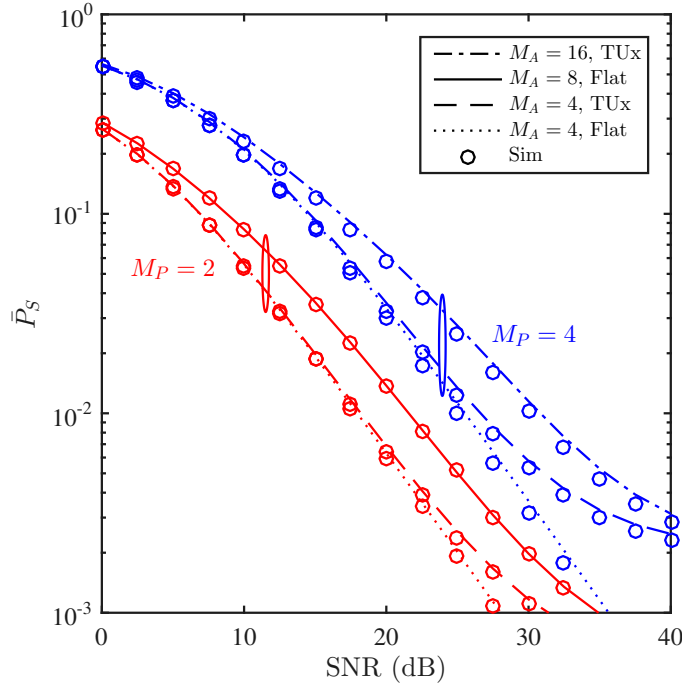
Figure 2.9: Initial theoretical and simulated average SER $\Pr(\hat{\mathcal{A}}_k \neq A_k) \ \forall \ k \in \mathbb{V}$ over the TUx and flat channels.

making an incorrect decision requires lower values of α_k. The results in the figure confirm the accuracy of the derived analytical expressions, however, the simulation results for the BPSK show better match because they are exact, while the QPSK results demonstrate some deviation due to the Gray coding approximation.

Fig. 2.12 presents the theoretical and simulated MSE versus SNR for BSK and QPSK respectively. As it can be noted from the figure, the analytical and simulation results match almost perfectly for SNR $\gtrsim$ 10 dB. However, the theoretical MSE of the QPSK slightly deviates from the simulation results at low SNRs due to the approximations made. Nevertheless, the results confirm the efficiency of the OSBCE because its MSE is comparable to pilot-based LSE for a wide range of SNRs, which is the lower bound for the MSE of the OSBCE. At very high SNRs, the MSE starts to diverge again from the lower bound because $P_{S,k}$ suffers from an error floor at such SNRs. However, the MSE at such high SNRs is much less than what is required to
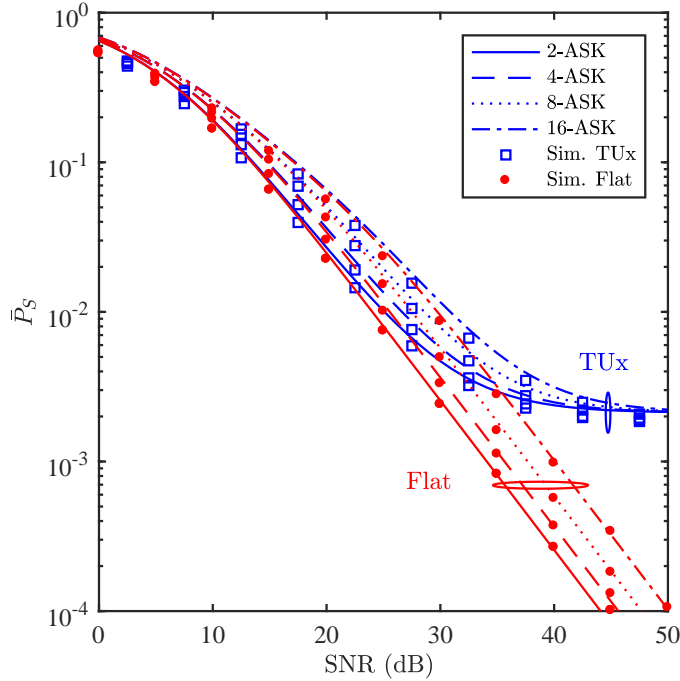
Figure 2.10: Initial theoretical and simulated average SER $\Pr(\hat{\mathcal{A}}_k \neq A_k) \ \forall \ k \in \mathbb{V}$ over the TUx and flat channels for different values of M_A, and $M_P = 4$.

provide BERs close to those with perfect knowledge of CSI. Although the derived upper bound is loose, it is useful to get a general idea about the system MSE with minimum effort.

Fig. 2.13 presents the final SER, $\Pr(\hat{A}_k \neq A_k)$, of the OSBCE for two different groups of subcarriers. In one case $\bar{P}_{S,k}$ is computed for all subcarriers, and in the other case, $\bar{P}_{S,k}$ is computed for subcarriers with index $k \in \mathbb{U}$, i.e., all subcarriers except the MPSK data symbols at pilot locations and the adjacent MASK symbols. In all cases, the final channel estimates $[\hat{H}_0, \hat{H}_1, \ldots, \hat{H}_{N-1}]$ are obtained using spline interpolation of the initial channel estimates $\hat{\mathcal{H}}_k \ \forall k \in \mathbb{V}$. The SER results are compared to pilot-based systems where the initial channel estimates are obtained using LSE, and the final estimates are obtained using spline interpolation. For all OSBCE cases, the MASK modulation order $M_A = 4$, and the SNRs per information bit γ_b for all systems are identical. As it can be noticed from the figure, the SER of the OSBCE where
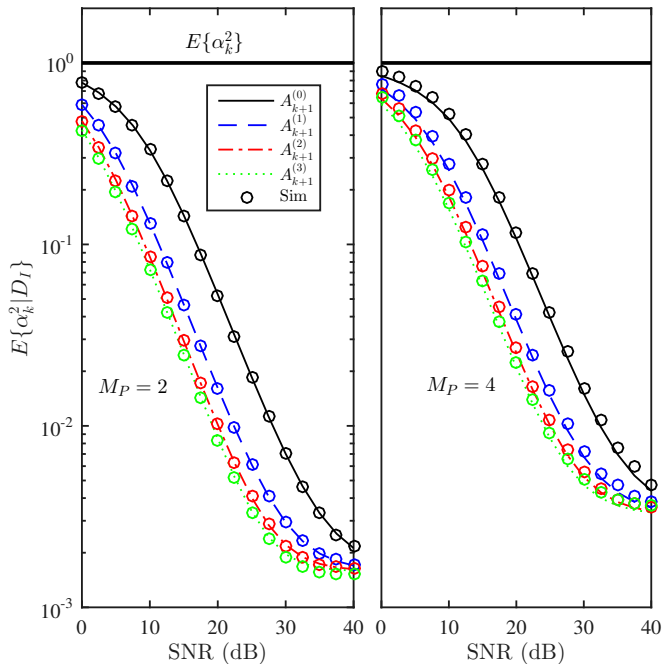
Figure 2.11: Theoretical and simulated conditional CFR variance in the case of incorrect decisions as compared to the unconditional variance of CFR for the TUx channel, $M_P = 2$ and 4, $M_A = 4$.
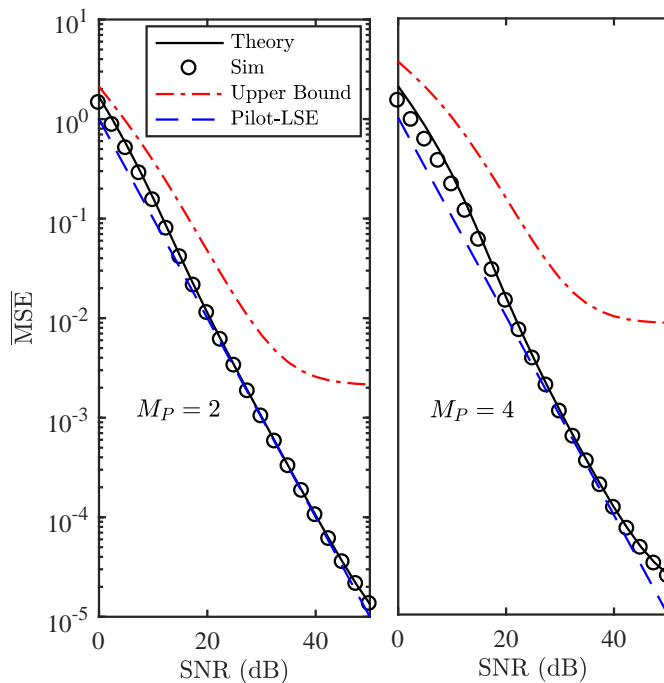
Figure 2.12: MSE of the initial channel estimate $\hat{\mathcal{H}}_k$, $M_P = 2$ and 4, $M_A = 4$.
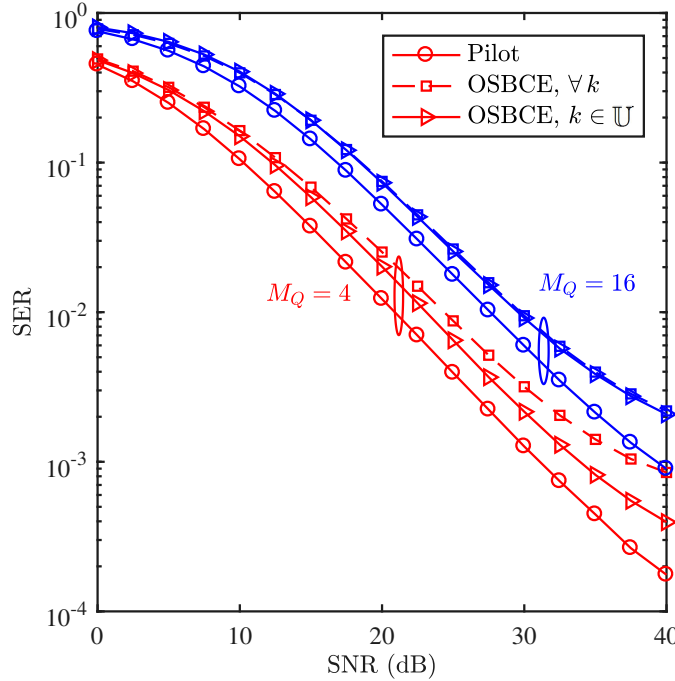
Figure 2.13: Overall system SER, $\Pr(\hat{A}_k \neq A_k)$, using spline interpolation for different M_Q values, $M_P = 4$ and $M_A = 4$.

$k \in \mathbb{U}$ is about 3 dB worse than the pilot-based system at moderate SNR values for $M_Q = 4$ and 16. The difference increases at high SNRs due to the SER floors of the MASK modulation. For the SER of all subcarriers case, it can be noted that the SER when $M_Q = 4$ experiences an additional increase, which is due to the difference between the SER for MASK and QAM when $M_A = M_Q$. For the case of $M_A = 4$ and $M_Q = 16$, it can be noticed that the SER for $\forall k$ and $k \in \mathbb{U}$ is almost identical, which implies that the average joint SER of the 4-PSK and 4-ASK pairs is similar to the 16-QAM [43]. Therefore, under SER constraints, the OSBCE spectral efficiency could be less than the FB system.

Fig. 2.14 presents the block error rate (BLER) of the OSBCE and pilot-based systems using turbo coding [3]. The block length is 160 bits, and the channel interleaver is modeled as 512×512 random interleaver. As it can be noted from the figure, the BLER of the OSBCE is comparable to the pilot based system where the coding
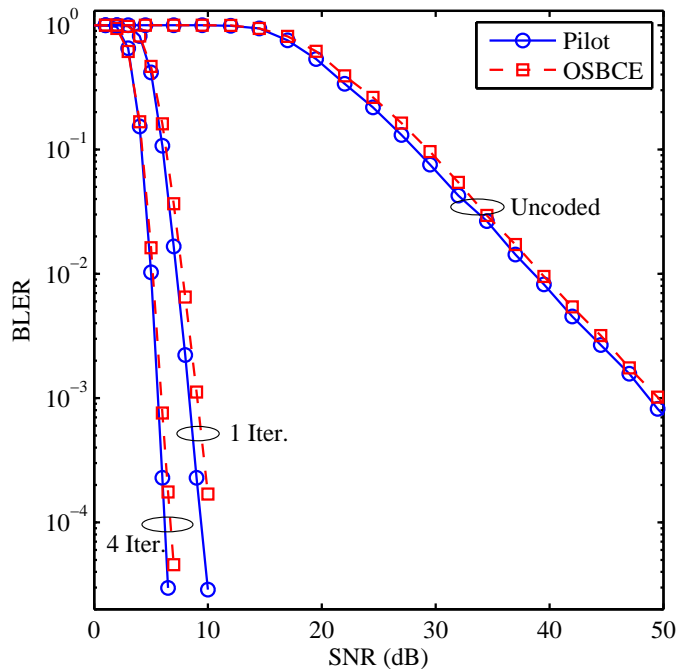
Figure 2.14: Coded BLER for the OSBCE and pilot-based systems using one and four decoding iterations, $M_A = M_P = M_Q = 4$.
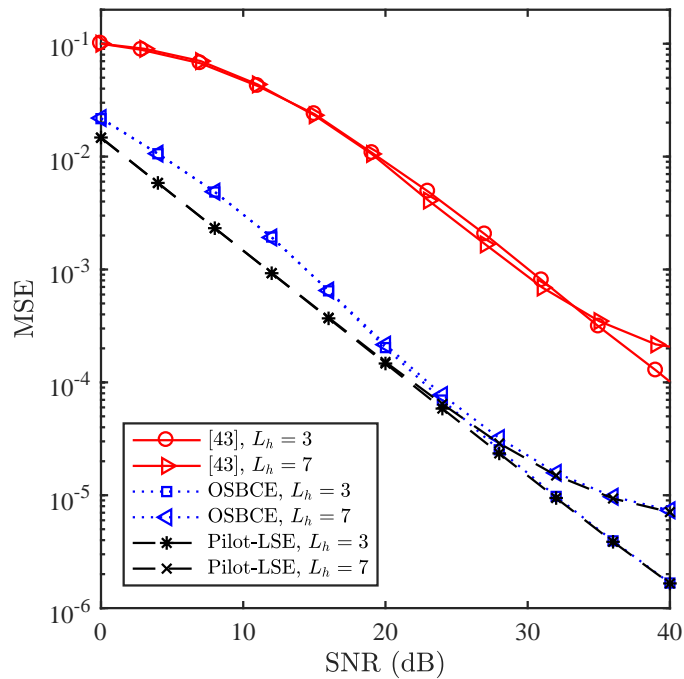
Figure 2.15: MSE of the OSBCE and the subspace blind estimator [44], where $M_A = M_P = 4$.
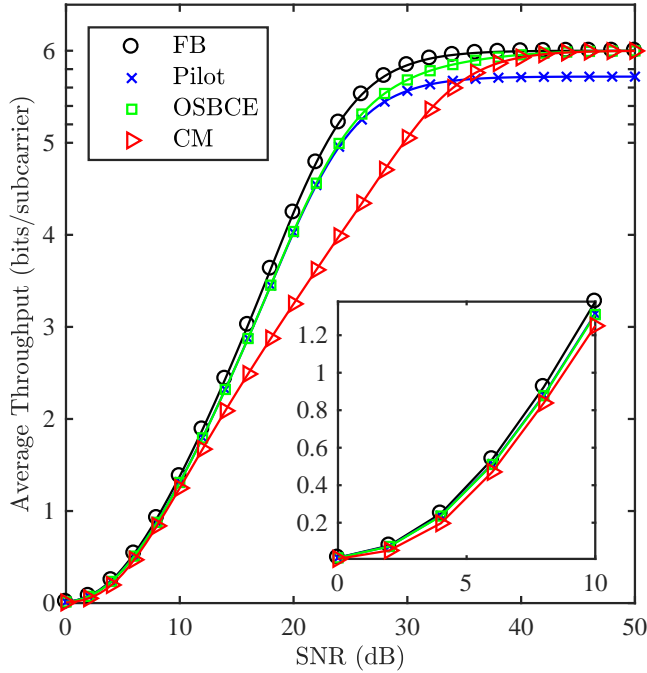
Figure 2.16: Average throughput per subcarrier for $P_{th} = 10^{-3}$.

gain difference is less than 0.5 dB.

Fig. 2.15 presents the MSE of the OSBCE and the subspace estimator reported in [44]. Because the estimation process in [44] yields the channel impulse response, the channel impulse response using the OSBCE are obtained by interpolating the initial channel estimates, and then computing the IFFT. The results are obtained based on the system and channel models used in [44]. As it can be noted from the figure, the proposed OSBCE significantly outperforms the subspace estimator for the entire range of SNRs. Moreover, it is worth noting that the observation window in [44] is N/N_{CP} OFDM symbols, while it is only one OFDM symbol for the OSBCE.

The average throughput of the OSBCE, CM, pilot and FB system is shown in Fig. 2.16, where the FB system is used as an upper limit to compare the spectral efficiency of the considered systems. The modulation order for all the considered systems varies from $1 \leq M(\ell, k) \leq 64$, where the Incremental Allocation Algorithm [36] is used to solve the optimization problem in (2.62). For the pilot-based system,

the distribution of the data and pilot subcarriers follows the LTE-A transmission grid as in Fig. 2.1. As it can be noted from Fig. 2.16, the throughput for all the considered systems is approximately equal for low SNR values, namely for SNR $\lesssim 10$ dB. In the moderate SNR range, $10 \lesssim \text{SNR} \lesssim 25$ dB, it can be noted that the OSBCE and the pilot-based system provide equivalent throughput while the CM throughput is well below the OSBCE and pilot-based systems. At high SNRs, the OSBCE and CM throughput approaches the FB system while the pilot-based saturates at 5.71 bit/subcarrier. Therefore, the OSBCE outperforms the pilot-based and CM for a wide range of SNRs.

It is worth noting that various approaches have been considered in the literature to provide the transmitter with the instantaneous SNRs that are required for applications such as bit loading, beamforming and precoding. Examples for such approaches are channel feedback [46], channel sounding [47] and channel reciprocity [48]. In channel feedback techniques, the channel estimates are obtained at the receiver side, and then fed back to the transmitter. On the contrary, channel sounding and channel reciprocity techniques do not rely on the receiver because the transmitter directly estimates the channel coefficients. Therefore, the optimization processes performed at the transmitter become independent of the channel estimation accuracy at the receiver side if one of the last two techniques is incorporated.

2.8   Chapter Summary

In this work, a novel blind channel estimator was introduced for OFDM systems with single transmit antenna based on a hybrid OFDM symbol structure, where pilot subcarriers in conventional OFDM systems are replaced by MPSK symbols, and the adjacent subcarriers are modulated using MASK. Therefore, the MASK symbol can be considered equivalent to the channel frequency response with respect to the

MPSK symbol, and hence, the MPSK symbol can be immediately detected, and then used to estimate the channel in a DD manner. The chapter also showed that the proposed OSBCE can be incorporated effectively and efficiently in practical systems such as LTE-A standard. The proposed estimator requires one OFDM symbol to estimate the CFR, which makes it suitable for mobile channels, where the channel frequently varies in the time domain. Monte Carlo simulation was used to verify the analytical results, which also confirmed that the OSBCE can produce reliable channel estimates as compared to pilot-based systems, with similar complexity, but with improved spectral efficiency. In future work, the proposed technique will be extended to the multiple transmit antenna case.

Bibliography

[1] A. Saci, A. Al-Dweik, A. Shami and Y. Iraqi, "One-Shot Blind Channel Estimation for OFDM Systems Over Frequency-Selective Fading Channels," *IEEE Trans. Commun.*, vol. 65, no. 12, pp. 5445-5458, Dec. 2017.

[2] E. Dahlman, S. Parkvall and J. Skold, 4G: LTE/LTE-Advanced for Mobile Broadband, Academic Press, 2nd Edition, London, UK, 2013.

[3] Digital Video Broadcasting (DVB); Frame Structure, Channel Coding and Modulation for a Second Generation Digital Terrestrial Television Broadcasting System (DVB-T2), ETSI Standard (EN) 302 755, 2011.

[4] IEEE Standard for Information technology, Telecommunications and information exchange between systems Local and metropolitan area networks, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, IEEE Std. 802.11-2012.

[5] IEEE Standard for Wireless MAN-Advanced Air Interface for Broadband Wireless Access Systems, Std. IEEE 802.16.1-2012.

[6] LTE; Evolved Universal Terrestrial Radio Access (E-UTRA), LTE physical layer, 3GPP TS 36.300, 2011.

[7] H. Lin, "Flexible configured OFDM for 5G air interface," *IEEE Access*, vol. 3, pp. 1861-1870, Sep. 2015.

[8] M. Morelli and U. Mengali, "A comparison of pilot-aided channel estimation methods for OFDM systems," *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 3065-3073, Dec. 2001.

[9] Ye Li, "Pilot-symbol-aided channel estimation for OFDM in wireless systems," *IEEE Trans. Veh. Technol.*, vol. 49, no. 4, pp. 1207-1215, July 2000.

[10] W. Chen and C. Chung, "Spectrally efficient OFDM pilot waveform for channel estimation," *IEEE Trans. Commun.*, vol. 65, no. 1, pp. 387-402, Jan. 2017.

[11] X. He, R. Song and W. Zhu, "Pilot allocation for distributed-compressed-sensing-based sparse channel estimation in MIMO-OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 65, no. 5, pp. 2990-3004, May 2016.

[12] R. Mohammadian, A. Amini and B. H. Khalaj, "Compressive sensing-based pilot design for sparse channel estimation in OFDM systems," *IEEE Commun. Lett.*, vol. 21, no. 1, pp. 4-7, Jan. 2017.

[13] X. Dong, W. Lu and A. Soong, "Linear interpolation in pilot symbol assisted channel estimation for OFDM," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, pp. 1910-1920, May 2007.

[14] M. X. Chang and T. D. Hsieh, "Detection of OFDM signals in fast-varying channels with low-density pilot symbols," *IEEE Trans. Veh. Technol.,* vol. 57, no. 2, pp. 859-872, Mar. 2008.

[15] C. Qi, G. Yue, L. Wu, Y. Huang and A. Nallanathan, "Pilot design schemes for sparse channel estimation in OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 64, no. 4, pp. 1493-1505, Apr. 2015.

[16] A. Al-Dweik and Y. Iraqi, "Error probability analysis and applications of amplitude-coherent detection in flat Rayleigh fading channels," *IEEE Trans. Commun.,* vol. 64, no. 5, pp. 2235-2244, May 2016.

[17] C. Tu and B. Champagne, "Subspace-based blind channel estimation for MIMO-OFDM systems with reduced time averaging," *IEEE Trans. Veh. Technol.,* vol. 59, no. 3, pp. 1539–1544, Mar. 2010.

[18] F. Gao, Y. Zeng, A. Nallanathan, and T. S. Ng, "Robust subspace blind channel estimation for cyclic prefixed MIMO OFDM systems," *IEEE J. Sel. Areas Commun.,* vol. 26, no. 2, pp. 378–388, Feb. 2008.

[19] M. C. Necker and G. L. Stüber, "Totally blind channel estimation for OFDM on fast varying mobile radio channels," *IEEE Trans. Wireless Commun.,* vol. 3, no. 5, pp. 1514–1525, Sep. 2004.

[20] F. Gao and A. Nallanathan, "Blind channel estimation for MIMO OFDM systems via non redundant linear precoding," *IEEE Trans. Signal Process.,* vol. 55, no. 2, pp. 784–789, Feb. 2007.

[21] T. Al-Naffouri, A. Dahman, M. Sohail, W. Xu and B. Hassibi, "Low-complexity blind equalization for OFDM systems with general constellations," *IEEE Trans. Signal Process.,* vol. 60, no. 12, pp. 6395-6407, Dec. 2012.

[22] S. Banani and R. Vaughan, "OFDM with iterative blind channel estimation," *IEEE Trans. Veh. Technol.,* vol. 59, no. 9, Nov. 2010.

[23] S. Park, B. Shim and J. Choi, "Iterative channel estimation using virtual pilot signals for MIMO-OFDM systems," *IEEE Trans. Signal Process.,* vol. 63, no. 12, pp. 3032-3045, June 2015.

[24] M. Song, D. Kim and G. Im, "Recursive channel estimation method for OFDM-based cooperative systems," *IEEE Commun. Lett.*, vol. 14, no. 11, pp. 1029-1031, Nov. 2010.

[25] Tao Cui and C. Tellambura, "Joint data detection and channel estimation for OFDM systems," *IEEE Trans. Commun.*, vol. 54, no. 4, pp. 670-679, Apr. 2006.

[26] M. Muck, M. de Courville and P. Duhamel, "A pseudorandom postfix OFDM modulator - semi-blind channel estimation and equalization," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 1005-1017, Mar. 2006.

[27] S. Abdallah and I. Psaromiligkos, "Semi-blind channel estimation with superimposed training for OFDM-based AF two-way relaying," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2468-2467, May 2014.

[28] Ming-Xian Chang and Y. T. Su, "Performance analysis of equalized OFDM systems in Rayleigh fading," *IEEE Trans. Wireless Commun.*, vol. 1, no. 4, pp. 721-732, Oct. 2002.

[29] A. Al-Dweik, A. Hazmi, S. Younis, B. Sharif, and C. Tsimenidis, "Carrier frequency offset estimation for OFDM systems over mobile Radio channels," *IEEE Trans. Veh. Technol.*, vol. 59, pp. 974-979, Feb. 2010.

[30] M. Mirahmadi, A. Al-Dweik and A. Shami, "BER reduction of OFDM based broadband communication systems over multipath channels with impulsive noise," *IEEE Trans. Commun.*,vol. 61, no. 11, pp. 4602-4615, Nov. 2013.

[31] M. S. Ahmed, S. Boussakta, B. S. Sharif and C. C. Tsimenidis, "OFDM based on low complexity transform to increase multipath resilience and reduce PAPR," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5994-6007, Dec. 2011.

[32] S. Tomasin and M. Butussi, "Analysis of interpolated channel estimation for mobile OFDM systems," *IEEE Trans. Commun.* vol. 58, no. 5, pp. 1578-1588, May 2010.

[33] F. D'Agostini, S. Carboni, M. De Castro, F. De Castro, and D. Trindade, "Adaptive concurrent equalization applied to multicarrier OFDM systems," *IEEE Trans. Broadcast.,* vol. 54, no. 3, pp. 441-447, Sep. 2008.

[34] M. Henkel, C. Schilling and W. Schroer, "Comparison of channel estimation methods for pilot aided OFDM systems," in *Proc. IEEE VTC. Spring,* Dublin, 2007, pp. 1435-1439.

[35] S. Liu, F. Wang, R. Zhang and Y. Liu, "A simplified parametric channel estimation scheme for OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 5082-5090, Dec. 2008.

[36] M. Chang, "A new derivation of least-squares-fitting principle for OFDM channel estimation," *IEEE Trans. Wireless Commun.*, vol. 5, no. 4, pp. 726-731, Apr. 2006.

[37] M. K. Simon, S. M. Hinedi, and W. C. Lindsey, *Digital communication techniques: signal design and detection*, Prentice Hall PTR, 1995.

[38] J. G. Proakis, and M. Salehi, *Digital communications*, 4th ed. New York: McGraw-Hill, 2001.

[39] M. K. Simon and M. S. Alouini, *Digital Communication over Fading Channels–A Unified Approach to Performance Analysis*, 2nd Ed., Wiley, 2005.

[40] M. Kalil, A. Shami, A. Al-Dweik, and S. Muhaidat, "Low-complexity power-efficient schedulers for LTE uplink with delay-sensitive traffic," *IEEE Trans. Veh. Technol*, vol. 64, no. 10, pp. 4551-4564, Oct. 2015.

[41] ETSI TR 125 943 V9.0.0 (2010-02), Universal Mobile Telecommunications System (UMTS) Deployment Aspects, 3GPP TR 25.943, Release 9.

[42] R. Mallik and N. Sagias, "Distribution of inner product of complex Gaussian random vectors and its applications," *IEEE Trans. Commun.*, vol. 59, no. 12, pp. 3353-3362, Dec. 2011.

[43] F. Xiong, "M-ary amplitude shift keying OFDM system," *IEEE Trans. Commun.,* vol. 51, no. 10, pp. 1638-1642, Oct. 2003.

[44] C. H. Tseng, Y. C. Cheng and C. D. Chung, "Subspace-based blind channel estimation for OFDM by exploiting cyclic prefix," *IEEE Wireless Commun. Lett.*, vol. 2, no. 6, pp. 691-694, Dec. 2013.

[45] A. Wyglinski, F. Labeau, and P. Kabal, "Bit loading with BER-constraint for multicarrier systems," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1383-1387, July 2005.

[46] A. Alkhateeb and R. W. Heath, "Frequency selective hybrid precoding for limited feedback millimeter wave systems," *IEEE Trans. Commun.,* vol. 64, no. 5, pp. 1801-1818, May 2016.

[47] N. B. Mehta, S. Kashyap and A. F. Molisch, "Antenna selection in LTE: from motivation to specification," *IEEE Commun. Mag.,* vol. 50, no. 10, pp. 144-150, Oct. 2012.

[48] F. E. Dorcheh and S. Shahbazpanahi, "Jointly optimal pre- and post-channel equalization and distributed beamforming in asynchronous bidirectional relay networks," *IEEE Trans. Signal Process.*, vol. 65, no. 17, pp. 4593-4608, Sept. 2017.

Chapter 3

Blind Channel Estimation using Cooperative Subcarriers for OFDM Systems

3.1 Introduction

[1] Estimation of channel state information (CSI) and equalization are among the primary operations at the receiver side to perform coherent information detection. Consequently, CSI estimation has to be performed accurately to avoid bit error rate (BER) degradation [2]. In the literature, numerous techniques have been proposed to obtain the CSI information accurately and effectively. However, pilot-aided CSI estimation remains the dominant approach in most industrial standards [2]-[7]. The density and distribution of pilot symbols depend on the channel conditions such as frequency selectivity and time variation of the channel [3]. For example, the pilot symbols in LTE-A [3] are distributed in comb-type arrangement in time-frequency subcarrier grid, where the pilots utilize 4.7% of the system bandwidth. The bandwidth

---

[1]A version of this chapter has been published in [1].

efficiency becomes even lower in systems where pilots are inserted in every OFDM block [9]. Therefore, blind channel estimation techniques have been considered widely in the literature [10]-[17] to avoid the spectral losses caused by the pilots. Despite the claimed spectral efficiency of blind algorithms, a significant number of blind estimators impose constraints on the modulation type [15]-[17], and hence, degrades the spectral efficiency indirectly. Moreover, computational complexity is another major concern for blind estimators [7], [14]. The additional computational complexity is generally due to the iterative nature of most blind algorithms [14], or because some estimators have to perform an exhaustive search over the solution space [28], which makes a wide range of blind estimators computationally prohibitive. Moreover, it is worth noting that obtaining accurate CSI is crucial for upper layers system adaptation and dynamic resource allocation operations, which can play a significant role in enhancing the spectral efficiency [18]-[20].

More recently, Saci *et al.* [21] proposed a low complexity blind channel estimator for OFDM-based communication systems, denoted as one-shot blind channel estimator (OSBCE). The proposed estimator exploits the channel correlation between adjacent subcarriers to estimate the CSI blindly. However, the performance of the OSBCE is mostly determined by the frequency selectivity of the channel. In this work, we propose a novel blind channel estimation technique by modifying the frame structure of the OSBCE. In the proposed scheme, the pilot symbols are replaced with M-ary phase shift keying (MPSK) symbols, and the subcarriers with the same subcarriers' indices in consecutive OFDM symbols are modulated using M-ary amplitude shift keying (MASK). The remaining subcarriers can be modulated using any efficient modulation scheme such as quadrature amplitude modulation (QAM). The analytical and simulation results show that the proposed estimator provides accurate channel estimates in channels with sever frequency selectivity.The performance of the proposed estimator is evaluated using the symbol error rate (SER), where an exact

closed-form formula is obtained. Moreover, the mean-squared error (MSE) of the channel frequency response estimates at the MPSK symbols is evaluated using Monte Carlo Simulations.

3.2 OFDM System and Channel Models

In OFDM systems, a serial data stream is converted into a parallel stream $\mathbf{a}^\ell = [A_0^\ell,$ $A_1^\ell, ..., A_{N-1}^\ell]^T$ to modulate N data symbols during the ℓth OFDM block. The data symbols A_k^ℓ are chosen from constellation such as MPSK, MASK, and QAM, with modulation orders M_P, M_A, and M_Q, respectively. In pilot-aided OFDM systems [3], the transmitted data $\mathbf{a}^\ell$ vector is composed of N_P subcarriers, which are allocated for pilot symbols for channel estimation purposes. Consequently, $N - N_P$ subcarriers are bearing information, and therefore, used for data transmission. The set of pilot and data subcarrier indices are denoted by $\mathbb{V}$ and $\mathbb{U}$, respectively.

The modulation process in OFDM is implemented efficiently by feeding the data $\mathbf{a}^\ell$ to an N-point inverse fast Fourier transform (IFFT) process. The output of the IFFT process during the ℓth OFDM block is given by,

$$\mathbf{x}^\ell = \mathbf{F}^H \mathbf{a}^\ell \tag{3.1}$$

where $\mathbf{F}$ and $\mathbf{F}^H$ are the normalized $N \times N$ FFT and IFFT matrices, respectively. The elements of $\mathbf{F}^H$ are defined as $F_{i,k}^H = \left(1/\sqrt{N}\right) e^{j2\pi ik/N}$ where i and k are the row and column indices $\{i, k\} \in \{0, 1, ..., N-1\}$, respectively. To combat inter-symbol-interference (ISI) between successive OFDM symbols and preserve the orthogonality among subcarriers, a cyclic prefix (CP) of length N_{CP} samples and duration of T_{CP} that has to exceed the channel maximum delay spread $\mathcal{D}_h$, is added to compose the transmitted OFDM symbol $\tilde{\mathbf{x}}^\ell$. The symbol $\tilde{\mathbf{x}}^\ell$ has a total length $N_{\mathrm{t}} = N + N_{\mathrm{CP}}$ samples and a total duration of $T_t = T_u + T_{\mathrm{CP}}$, where T_u is the useful OFDM frame duration.

At the receiver end, the received signal will propagate through a multipath fading channel and additive white Gaussian noise (AWGN) will be added to the signal. In the scope of this work, the channel is assumed to be composed of $\mathcal{D}_h + 1$ independent and identically distributed (i.i.d.) multipath components, each of which has a gain $h_m^\ell \sim \mathcal{CN}\left(0, 2\sigma_{h_m}^2\right)$, and delay $m \times T_{\mathrm{s}}$, where $2\sigma_{h_m}^2$ is the normalized power of the mth multipath component $\sum_{n=0}^{\mathcal{D}_h} 2\sigma_{h_n}^2 = 1$, $m \in \{0, 1, ..., \mathcal{D}_h\}$, and $T_s \triangleq T_t/N_t$. Also, we assume that a quasi-static multipath channel is adopted [22], in which the channel taps are constant over one OFDM symbol, but they may vary over two consecutive symbols. The AWGN samples are modeled as $z_n^\ell \sim \mathcal{CN}\left(0, 2\sigma_z^2\right)$.

After discarding the CP symbols at the receiver side, the received samples at the ℓth OFDM symbol are given by,

$$y_n^\ell = \sum_{m=0}^{\mathcal{D}_h} h_m^\ell \, x_n^\ell + z_n^\ell. \tag{3.2}$$

Then, the receiver applies the FFT to the received non-CP sequence $\mathbf{y}^\ell$, where $\mathbf{y}^\ell = \mathbb{H}^\ell \mathbf{x}^\ell + \mathbf{z}^\ell$, $\mathbb{H}$ is an $N \times N$ circulant matrix, and $\mathbf{z} \sim \mathcal{CN}\left(\mathbf{0}, 2\sigma_z^2\mathbf{i}\right)$ is the AWGN vector, where $\mathbf{0}$ and $\mathbf{i}$ are $N \times 1$ zero and unit vectors, respectively. Therefore, the FFT output can be expressed as,

$$\mathbf{r}^\ell = \mathbf{F}\,\mathbf{y}^\ell = \mathbf{F}\mathbb{H}^\ell\mathbf{F}^H\mathbf{a}^\ell + \mathbf{F}\mathbf{z}^\ell. \tag{3.3}$$

Since the matrix $\mathbb{H}^\ell$ is circulant, then, it will be diagonalized by FFT and IFFT matrices. Hence, we can write (3.3) as,

$$\mathbf{r}^\ell = \mathbf{H}^\ell\mathbf{a}^\ell + \mathbf{w}^\ell \tag{3.4}$$

where $\mathbf{r}^\ell \in \mathbb{C}^{N \times 1}$, $\mathbf{w}^\ell \sim \mathcal{CN}\left(\mathbf{0}, 2\sigma_w^2\mathbf{i}\right)$ is the FFT of the noise vector $\mathbf{z}^\ell$, and $\mathbf{H}^\ell$ denotes the channel frequency response (CFR), $\mathbf{H}^\ell = \mathrm{diag}\left\{\left[H_0^\ell, H_1^\ell, \ldots, H_{N-1}^\ell\right]\right\}$, where $H_k^\ell \sim \mathcal{CN}\left(0, 2\sigma_H^2\right)$. In order to extract the information symbols from the received channel, the received symbols r_k^ℓ will be fed to a channel equalizer and a maximum likelihood detector (MLD). In this work, zero forcing (ZF) equalizer is

used, and then, the estimated output can be expressed as,

$$\hat{\mathcal{A}}_k^\ell = \arg\min_{A_k^\ell(i)} \left| \frac{\left(\hat{\mathcal{H}}_k^\ell\right)^*}{|\hat{\mathcal{H}}_k^\ell|^2} r_k^\ell - \hat{\mathcal{A}}_k^\ell(i) \right|^2 , \ i = [0, ..., M-1], \ \{k, \ell\} \in \mathbb{V} \tag{3.5}$$

where $\hat{\mathcal{H}}_k^\ell$ is the estimated CFR at the subcarrier location $\{k, \ell\} \in \mathbb{V}$. In OFDM-based standards such as LTE-A [3], the channel estimation is performed over two steps:

1. Initial channel estimates are computed at the pilot subcarriers using least squares estimation (LSE) as,

$$\hat{\mathcal{H}}_k^\ell = \frac{r_k^\ell}{A_k^\ell}, \ \{k, \ell\} \in \mathbb{V} \tag{3.6}$$

   where the pilot symbols values A_k^ℓ are perfectly known at the receiver side.

2. Estimating the channel at data subcarriers $\{k, \ell\} \in \mathbb{U}$ using interpolation and fitting techniques such as linear, spline, and two-dimensional (2D) interpolation [7].

3.3 The Proposed System

Assuming that the channel variations over time are modeled using Jake's model [23], and given that h_m and h_n are mutually independent $\forall m \neq n$, then $E\left\{|h_m|^2\right\} = 2\sigma_{h_m}^2$ and $E\left\{h_n h_m^*\right\}|_{n \neq m} = 0$. Then, we can define the time correlation coefficient $\varrho_t \triangleq E\left\{h_n^\ell \left(h_m^{\ell+1}\right)^*\right\}$ as,

$$\varrho_t = \begin{cases} 2\sigma_{h_n}^2 J_0\left(2\pi f_d T_t\right), & n = m \\ 0, & otherwise \end{cases} \tag{3.7}$$

where $J_0\left(\cdot\right)$ is the Bessel function of the first kind and zero order, and f_d is the maximum Doppler shift. Therefore, for most practical values of T_t and f_d in (3.7), it can be noted that $|\varrho_t| \approx 1$ and $\arg\{\varrho_t\} \approx 0$, which implies that $H_k^\ell \approx H_k^{\ell+1}$.
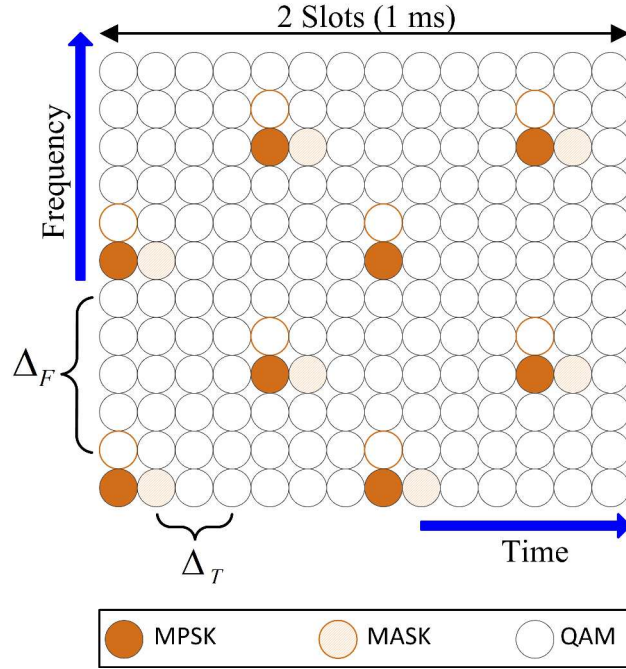
Figure 3.1: Proposed frame structure.

Consequently, the FFT output at kth subcarrier and OFDM symbols ℓ and $\ell+1$ can be written as,

$$r_k^\ell = H_k^\ell A_k^\ell + w_k^\ell \tag{3.8}$$

$$r_k^{\ell+1} \approx H_k^\ell A_k^{\ell+1} + w_k^{\ell+1}. \tag{3.9}$$

Moreover, at high signal-to-noise ratios (SNRs), the noise component $w_k^{\ell+1}$ in (3.9) can be neglected such that it can be simplified to $r_k^{\ell+1} \approx H_k^\ell A_k^{\ell+1}$. If the symbols $A_k^{\ell+1}$ are modulated using a unipolar MASK, $A_k^{\ell+1} \in \mathbb{R}^+$ as in [24], and A_k^ℓ using MPSK as illustrated in Fig. 3.1. Then, the equalization process of the MPSK symbols can be expressed by,

$$\hat{\mathcal{S}}_k^\ell \triangleq r_k^\ell \left(r_k^{\ell+1} \right)^*. \tag{3.10}$$

In this work, we denote the set of MASK subcarrier indices as $\mathbb{T}$. The equalization in (3.10) is feasible using the multiplication process because A_k^ℓ requires only the knowledge of $\theta_k^\ell \triangleq \arg\left\{ H_k^\ell \right\}$ to perform the coherent detection at the MPSK

symbols r_k^ℓ, which can be achieved by noting that $A_k^{\ell+1} \in \mathbb{R}^+$ and $\hat{\vartheta}_k^\ell \triangleq \arg\left\{r_k^{\ell+1}\right\} \approx$ $\arg\left\{H_k^\ell A_k^{\ell+1}\right\} = \arg\left\{H_k^\ell\right\}$, where $\hat{\vartheta}_k^\ell$ is the estimate of θ_k^ℓ. Therefore, the MLD output can be expressed as,

$$\hat{\mathcal{A}}_k^\ell = \arg\min_{A_k^\ell(i)} \ \left|r_k^\ell\left(r_k^{\ell+1}\right)^* - A_k^\ell(i)\right|^2, \ i = [0, ..., M_P - 1]. \tag{3.11}$$

Once $\hat{\mathcal{A}}_k^\ell$ is obtained, we can estimate the CFR $\hat{\mathcal{H}}_k^\ell$ at locations $\{k, \ell\} \in \mathbb{V}$ using LSE similar to (3.6) as,

$$\hat{\mathcal{H}}_k^\ell = \frac{r_k^\ell}{\hat{\mathcal{A}}_k^\ell}. \tag{3.12}$$

3.4   System Performance

In this section, we evaluate the SER of MPSK analytically at the locations $\{k, \ell\} \in \mathbb{V}$. For the proposed system, we can express the true and the estimated CFR as,

$$H_k^\ell \triangleq x_1 + jx_2 = \alpha_k^\ell e^{j\theta_k^\ell} \tag{3.13}$$
$$\mathcal{H}_k^\ell = r_k^{\ell+1} = H_k^{\ell+1} A_k^{\ell+1} + w_k^{\ell+1} \triangleq x_3 + jx_4 = \beta_k^\ell e^{j\hat{\vartheta}_k^\ell} \tag{3.14}$$

where α_k^ℓ and β_k^ℓ are the fading envelopes of the true and the estimated CFR, respectively. The random variables $\{x_1, x_2, x_3, x_4\}$ are all zero-mean Gaussian random variables, where x_1 and x_2 are independent, likewise x_3 and x_4. By making the transformation of the joint Gaussian PDF in [35] from rectangular (x_1, x_2, x_3, x_4) to polar coordinates $(\alpha_k^\ell, \beta_k^\ell, \theta_k^\ell, \hat{\vartheta}_k^\ell)$ and defining the phase error as $\psi_k^\ell \triangleq \theta_k^\ell - \hat{\vartheta}_k^\ell$, then, the joint PDF can be expressed by,

$$P(\alpha_k^\ell, \beta_k^\ell, \psi_k^\ell) = \frac{\alpha_k^\ell \beta_k^\ell}{2\pi \sigma_1^2 \sigma_2^2 (1 - \rho^2)} \exp\left\{-\frac{1}{2(1 - \rho^2)}\right.$$
$$\left. \times \left[\frac{\left(\alpha_k^\ell\right)^2}{\sigma_1^2} + \frac{\left(\beta_k^\ell\right)^2}{\sigma_2^2} - 2\frac{\alpha_k^\ell \beta_k^\ell}{\sigma_1 \sigma_2}(\rho_1 \cos(\psi_k^\ell) - \rho_2 \sin(\psi_k^\ell))\right]\right\} \tag{3.15}$$

where (ρ_1, ρ_2, ρ) are defined as,

$$\rho_1 \triangleq \frac{\mu_1}{\sigma_1 \sigma_2}, \; \rho_2 \triangleq \frac{\mu_2}{\sigma_1 \sigma_2}, \; \rho \triangleq \sqrt{\rho_1^2 + \rho_2^2} \tag{3.16}$$

and $(\sigma_1^2, \sigma_2^2, \mu_1, \mu_2)$ are given in [27] as,

$$\sigma_1^2 \triangleq \frac{1}{2} \mathrm{E}\left[\left|H_k^\ell\right|^2\right] = \sigma_H^2 \tag{3.17}$$

$$\sigma_2^2 \triangleq \frac{1}{2} \mathrm{E}\left[\left|\mathcal{H}_k^\ell\right|^2\right] = \left(A_{k+1}^\ell\right)^2 \sigma_H^2 + \sigma_w^2 \tag{3.18}$$

$$\mu_1 + j\, \mu_2 \triangleq \frac{1}{2} \mathrm{E}\left[\mathcal{H}_k^\ell H_k^{\ell*}\right] = A_k^{\ell+1} \sigma_H^2 \left[J_0\left(2\pi f_d T_s\right)\right]. \tag{3.19}$$

However, BER of BPSK modulation, conditioned on a fixed fading envelope α_k^ℓ and phase error ψ_k^ℓ, is given by [26, 10.14a],

$$P_{B,k}^\ell|\left[\psi_k^\ell, \alpha_k^\ell\right] = Q\left(\frac{\alpha_k^\ell}{\sigma_w} \cos(\psi_k^\ell)\right) \tag{3.20}$$

where $Q(x) \triangleq 1/\sqrt{2\pi} \int_x^\infty e^{-t^2/2} dt$. Therefore, the conditioning on $\left[\psi_k^\ell, \alpha_k^\ell\right]$ can be removed by evaluating the following integral,

$$P_{B,k}^\ell = \int_0^\infty \int_0^\infty \int_{-\pi}^\pi P_{B,k}^\ell|\left[\psi_k^\ell, \alpha_k^\ell\right] \cdot P(\alpha_k^\ell, \beta_k^\ell, \psi_k^\ell)\, d\alpha_k^\ell\, d\beta_k^\ell\, d\psi_k^\ell. \tag{3.21}$$

Using the correlation coefficients obtained in (3.16)-(3.19), and the integral identity established in [27, Appendix B], we get the closed-form BER conditioned on $A_k^{\ell+1}$ for BPSK as,

$$P_{B,k}^\ell|A_k^{\ell+1} = \frac{1}{2}\left[1 - \frac{A_k^{\ell+1}\left[J_0\left(2\pi f_d T_s\right)\right]}{1 + \frac{1}{\mathrm{SNR}}}\right]. \tag{3.22}$$

Since the transmitted MASK symbols $A_k^{\ell+1}$ are equally probable, and given that SER is equal to BER for the BPSK modulation, the unconditional SER can be obtained by averaging over all values of $A_k^{\ell+1}$ as,

$$\bar{P}_{S,k}^\ell = \bar{P}_{B,k}^\ell = \frac{1}{M_A} \sum_{i=0}^{M_A-1} \left(P_{B,k}^\ell|A_k^{\ell+1}(i)\right). \tag{3.23}$$
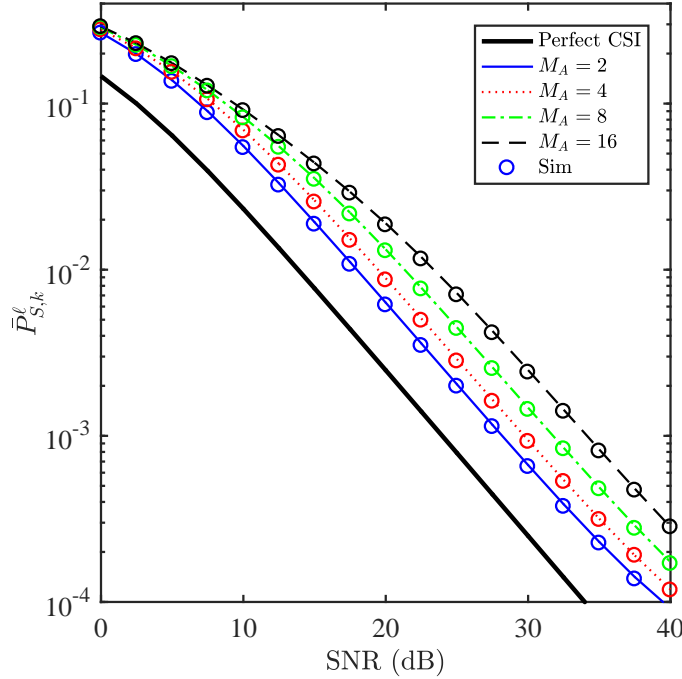
Figure 3.2: Theoretical and simulated $\bar{P}_{S,k}^{\ell}$ at MPSK locations $\{k,\ell\} \in \mathbb{V}$ for vehicle speed $V = 20$ km/h, $M_P = 2$, and different values of M_A.

3.5 Complexity Analysis

In this section, we can analyze the complexity of the proposed estimator by calculating the number of complex multiplication/division and complex addition/subtraction required to estimate the CFR in one OFDM symbol. Then, we compare its complexity to the pilot-aided LSE OFDM. To simplify the discussion, the channel estimates in one OFDM symbol of the proposed system can be obtained using the following stages:

1. Channel equalization at $\{k,\ell\} \in \mathbb{V}$: which is represented by the multiplication $r_k^{\ell} \left(r_k^{\ell+1} \right)^*$ process: This step requires one complex multiplication per MPSK symbol. This step is equivalent to the channel equalization in the pilot-aided system. Therefore, it is not considered an additional complexity.

2. Information detection at MPSK locations $\{k,\ell\} \in \mathbb{V}$: which is performed by
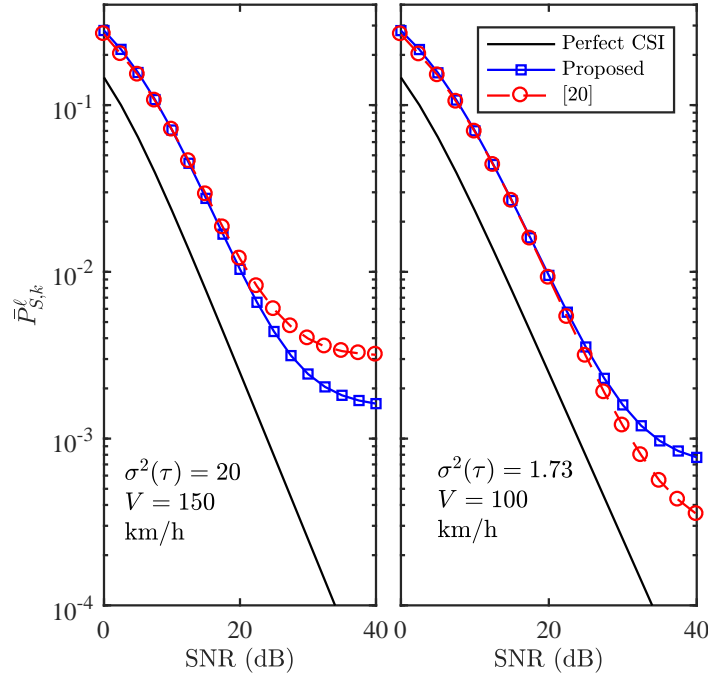
Figure 3.3: $\bar{P}_{S,k}^{\ell}$ at MPSK locations $\{k, \ell\} \in \mathbb{V}$ for $M_P = 2$, $M_A = 4$ and different vehicle speeds V as compared to [21] for different delay spreads $\sigma^2(\tau)$.

computing $\hat{\mathcal{A}}_k^{\ell}$ using maximum likelihood criterion in (3.11). This step is also not considered an excess complexity because it is required anyway to recover the information symbol in the pilot-aided systems.

3. Channel estimation at the MPSK symbols $\hat{\mathcal{H}}_k^{\ell}$: this step requires one complex division as in (3.12). It is worth noting that this step equivalent to the LSE in the pilot-aided as in (3.11).

4. Channel estimation at the remaining subcarriers, $\forall k \in \{\mathbb{T} \cap \mathbb{U}\}$: the complexity of this step depends on the interpolation technique used. However, the same step is needed for the pilot-aided with the same complexity.

Therefore, the proposed system complexity is equivalent to pilot-aided system with LSE, which confirms the claimed low complexity. In addition, the complexity the proposed estimator is identical to the OSBCE.

3.6   Numerical Results

To evaluate the system performance under various operating conditions, we consider the downlink physical layer specifications of the LTE-A standard [3], where the carrier frequency $f_c = 1.9$ GHz, the sampling frequency $f_s \triangleq 1/T_s$ is 3.836 MHz, $N = 256$ subcarriers, $N_{\mathrm{CP}} = 18$ samples, subcarrier spacing is 15 KHz, total OFDM symbol period is $T_s = 71.3 \,\mu$ sec, and CP duration is $T_{\mathrm{CP}} = 4.69 \,\mu$ sec. Unless it is mentioned otherwise, the channel model considered is the typical urban (TUx) multipath fading model [45], which consists of 9 taps with normalized delays of [0, 1, ..., 8] samples, and the normalized path gains of $2\sigma_{h_n}^2$, $n = [0, 1, ..., 8]$ are [0.269, 0.174, 0.289, 0.117, 0.023, 0.058, 0.036, 0.026, 0.008], which corresponds to a delay spread $\sigma^2(\tau) = 3.49$, where the delay spread is defined as,

$$\sigma^2(\tau) = \frac{\sum_i g_i \tau_i^2}{\sum_i g_i} - \left(\frac{\sum_i g_i \tau_i}{\sum_i g_i}\right)^2 \tag{3.24}$$

where g_i is the gain of path index i, and τ_i is the delay of path index j. For the sake of comparison, other two channel models are used with the delay spreads $\sigma^2(\tau) = 1.73$, $\sigma^2(\tau) = 20$, respectively as defined in [22]. The maximum Doppler frequency is modeled as a user equipment (UE) in a moving vehicle, which is given by $f_d = \frac{V}{c} f_c$, where V is the speed of the vehicle, and c is the speed of light, $c = 3 \times 10^8$ m/s. The simulation results are obtained using Monte Carlo simulations, where in each run, 2×10^6 OFDM symbols are generated.

Fig. 3.2 shows the effect of changing the modulation order of the MASK symbols M_A on the SER at MPSK symbols $P_{S,k}^\ell$ at a vehicle speed of $V = 20$ km/h and $M_P = 2$. As it can be seen from the figure, increasing M_A will increase the SER as well. This is due to the fact that MASK symbols are used as channel estimates for the MPSK symbols, and hence, increasing M_A will introduce more small amplitudes, which leads to less reliable estimates that cause SER degradation.
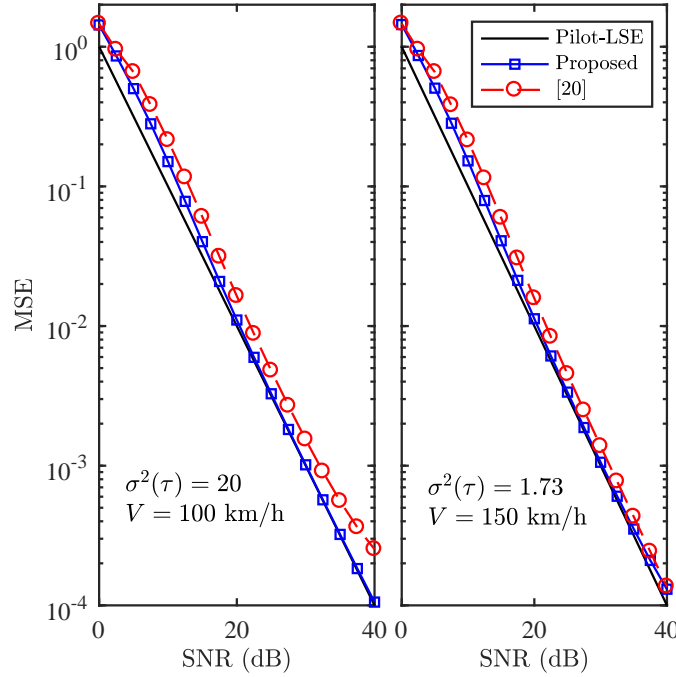
Figure 3.4: MSE of the channel estimate $\mathcal{H}_k^\ell$ at locations $\{k, \ell\} \in \mathbb{V}$ for $M_P = 2$, $M_A = 4$ and different vehicle speeds V as compared to [21] for different delay spreads $\sigma^2(\tau)$.

The effect of the vehicle speed V and the channel frequency selectivity on $P_{S,k}^\ell$ is given in Fig. 3.3. On the left side subfigure, the channel is severely frequency selective and the vehicle speed is high, which causes error floors for both, the proposed and OSBCE. However, the proposed estimator error floor is lower than the OSBCE. On the right side subfigure, the channel frequency selectivity is very low while V moderate. As it can be noted from the figure, the OSBCE starts to outperform the proposed system at SNR$\approx$ 26 dB. Such behavior can be justified by the fact that $\mathrm{E}\left[\arg\left(H_k^{\ell+1}\right)\right] \neq \mathrm{E}\left[\arg\left(H_k^\ell\right)\right]$ at high f_d values while $\mathrm{E}\left[\arg\left(H_k^\ell\right)\right] \neq \mathrm{E}\left[\arg\left(H_{k+1}^\ell\right)\right]$ at high $\sigma^2(\tau)$ values. Therefore, the proposed estimator is more robust at low and moderate vehicle speeds, while the OSBCE is more robust in channels with mild and low frequency selectivity. Consequently, the MASK-MPSK grid should be adaptive based on the channel conditions to achieve the best SER. Furthermore, Figs. 3.2 and 3.3 show perfect matching between the theoretical and simulations results.
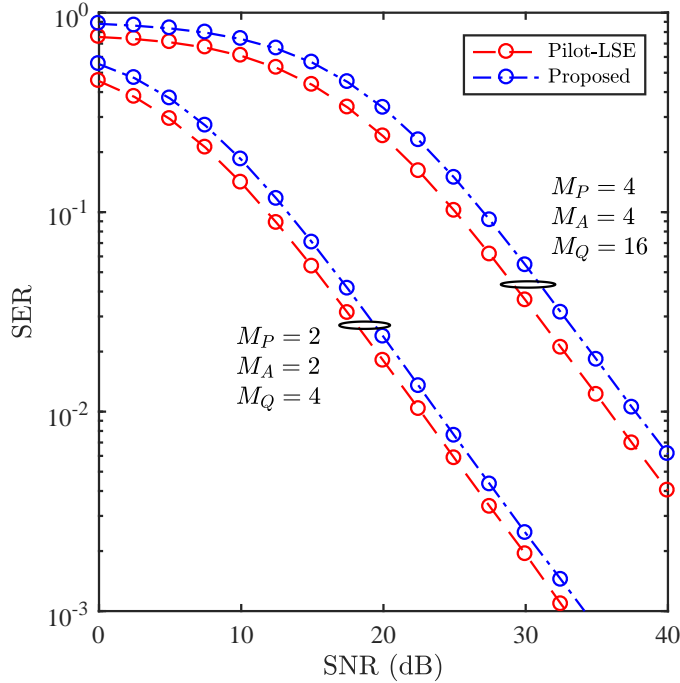
Figure 3.5: SER after spline interpolation at QAM locations $\{k, \ell\} \in \mathbb{U}$ for the proposed system as compared to LSE, at $V = 50$ km/h and for different values of M_P, M_A and M_Q.

Fig. 3.4 presents the MSE as a performance measure for the accuracy of the proposed channel estimator. As it can be noted, the results are benchmarked against MSE of the conventional pilot-aided LSE, which is considered a lower bound for the proposed estimator, that is given by MSE $= 1/\text{SNR}$. As the figure indicates, the proposed estimator produces channel estimates that are comparable to the pilot-aided LSE estimates, particularly, at moderate and high values of SNR. Also, the figure shows that the proposed estimator provides more accurate estimates than the OSBCE in [21], even when the Doppler frequency is high and channel selectively is low.

Fig. 3.5 depicts the SER at QAM subcarriers $\{k, \ell\} \in \mathbb{U}$ for $V = 50$ km/h after spline interpolation for two different configurations of (M_A, M_P, M_Q). SER results are compared to the pilot-aided LSE, where spline interpolation is used to obtain CFR at QAM symbols from the channel estimates $\mathcal{H}_k^\ell$. The results show that

proposed system has a reliable performance by which the difference is only about 1-2 dB as compared to the pilot-aided LSE. However, the proposed system gains more power and spectral efficiencies that are consumed by the pilot symbols in the conventional systems without any additional complexity. It is worth noting that a thorough study of the spectral efficiency analysis of blind against pilot-aided schemes has been discussed in the work reported in [21].

In the context of comparing blind algorithms to the pilot-aided estimators, it is worth mentioning that a fair comparison between the compared systems, the power efficiency should also be taken into account. Therefore, the power per information subcarrier in a blind system $P_{SC,blind}$ should be expressed as,

$$P_{SC,blind} = \frac{N_F \times N_T}{N_F \times N_T - N_{P,\text{RB}}} P_{SC,pilot} \tag{3.25}$$

where $N_{P,\text{RB}}$ is the number of pilot subcarriers per OFDM grid, N_F, N_T are the frequency and time dimensions per OFDM grid, receptively. However, the additional power gained in (3.25) are allocated to the MASK symbols since they are the less reliable modulation scheme, and hence, the power of each MASK subcarrier will gain an additional 3 dB gain as derived in Appendix A. However, it is worth noting that OSBCE system requries one OFDM symbol to perform the channel estimation process, and hence, denoted as one-shot estimation. On the other hand, the proposed system in this work needs two OFDM symbols to estation the channel, which is twice the observation window size of the OSBCE.

3.7   Chapter Summary

In this chapter, a novel blind channel estimator was introduced for OFDM based systems over mobile radio channels. The proposed transmission frame uses a hybrid OFDM frame structure to estimate the CSI blindly. In this arrangement, MPSK data

symbols replace pilot subcarriers in the conventional OFDM systems, and the subcarrier with the same pilot subcarrier index in the following OFDM symbol is modulated using MASK. The MASK symbols are used as CFR estimates to detect the MPSK symbols. The detected MPSK symbols are then used as pilot symbols to estimate the CFR at the remaining subcarriers. The system performance was evaluated using LTE downlink specifications and for various speeds of the UE. Monte Carlo simulations, verified by the analytical results, showed that the proposed system supports high mobility speeds up to 150 km/h and produces accurate channel estimates comparable to pilot-aided systems, with improved power and spectral efficiencies. However, significant performance improvement can be achieved by using an adaptive time/frequency configuration.

Bibliography

[1] A. Saci, A. Al-Dweik and A. Shami, "Blind channel estimation using cooperative subcarriers for OFDM systems," *IEEE Int. Conf. Commun. (ICC)*, Kansas City, USA, May 2018.

[2] Z. Sheng, H. D. Tuan, H. H. Nguyen and Y. Fang, "Pilot optimization for estimation of high-mobility OFDM channels," *IEEE Trans. Veh. Technol.*, vol. 66, no. 10, pp. 8795-8806, Oct. 2017.

[3] M. X. Chang and T. D. Hsieh, "Detection of OFDM signals in fast-varying channels with low-density pilot symbols," *IEEE Trans. Veh. Technol.*, vol. 57, no. 2, pp. 859-872, March 2008.

[4] X. He, R. Song and W. Zhu, "Pilot allocation for distributed-compressed-sensing-based sparse channel estimation in MIMO-OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 65, no. 5, pp. 2990-3004, May 2016.

[5] B. Gong, L. Gui, Q. Qin, X. Ren and W. Chen, "Block distributed compressive sensing-based doubly selective channel estimation and pilot design for large-scale MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 66, no. 10, pp. 9149-9161, Oct. 2017.

[6] R. Mohammadian, A. Amini and B. H. Khalaj, "Deterministic pilot design for sparse channel estimation in MISO/Multi-User OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 129-140, Jan. 2017.

[7] M. Jiang, S. Huang and W. Wen, "Adaptive polar-linear interpolation aided channel estimation for wireless communication systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 3, pp. 920-926, March 2012.

[8] LTE; Evolved Universal Terrestrial Radio Access (E-UTRA), LTE physical layer, 3GPP TS 36.300, 2011.

[9] ETSI EN 302 755, Digital Video Broadcasting (DVB); Frame structure channel coding and modulation, for a second generation digital terrestrial Television broadcasting system (DVB-T2), V1.4.1, 2015.

[10] Y. C. Pan and S. M. Phoong, "An improved subspace-based algorithm for blind channel identification using few received blocks," *IEEE Trans. Commun.,* vol. 61, no. 9, pp. 3710-3720, September 2013.

[11] F. Gao, Y. Zeng, A. Nallanathan, and T. S. Ng, "Robust subspace blind channel estimation for cyclic prefixed MIMO OFDM systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 2, pp. 378–388, Feb. 2008.

[12] A. Ladaycia, A. Mokraoui, K. Abed-Meraim and A. Belouchrani, "Performance bounds analysis for semi-blind channel estimation in MIMO-OFDM communications systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 5925-5938, Sept. 2017

[13] W. Zhang, F. Gao and Q. Yin, "Blind channel estimation for MIMO-OFDM systems with low order signal constellation," *IEEE Commun. Lett.*, vol. 19, no. 3, pp. 499-502, March 2015.

[14] P. Zhang, S. Chen and L. Hanzo, "Embedded iterative semi-blind channel estimation for three-stage-concatenated MIMO-aided QAM turbo transceivers," *IEEE Trans. Veh. Technol.*, vol. 63, no. 1, pp. 439-446, Jan. 2014.

[15] N. Sarmadi, S. Shahbazpanahi and A. B. Gershman, "Blind channel estimation in orthogonally coded MIMO-OFDM systems: a semidefinite relaxation approach," *IEEE Trans. Signal Process.*, vol. 57, no. 6, pp. 2354-2364, June 2009.

[16] T. Y. Al-Naffouri and A. A. Quadeer, "Cyclic prefix based enhanced data recovery in OFDM," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3406-3410, June 2010.

[17] T. Cui and C. Tellambura, "Efficient blind receiver design for orthogonal space-time block codes," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, pp. 1890-1899, May 2007.

[18] K. Hammad, S. L. Primak, M. Kalil and A. Shami, "QoS-aware energy-efficient downlink predictive scheduler for OFDMA-based cellular devices," *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 1468-1483, Feb. 2017.

[19] M. Kalil, A. Shami, A. Al-Dweik, and S. Muhaidat, "Low-complexity power-efficient schedulers for LTE uplink with delay-sensitive traffic," *IEEE Trans. Veh. Technol*, vol. 64, no. 10, pp. 4551-4564, Oct. 2015.

[20] E. Aqeeli, A. Moubayed and A. Shami, "Power-aware optimized RRH to BBU allocation in C-RAN," *IEEE Trans. Wireless Commun.*, vol. PP, no. 99, pp. 1-1. doi: 10.1109/TWC.2017.2777825.

[21] A. Saci, A. Al-Dweik, A. Shami and Y. Iraqi, "One-shot blind channel estimation for OFDM systems over frequency-selective fading channels," *IEEE Trans. Commun.,* vol. 65, no. 12, pp. 5445-5458, Dec. 2017.

[22] A. Al-Dweik, A. Hazmi, S. Younis, B. Sharif, and C. Tsimenidis, "Carrier frequency offset estimation for OFDM systems over mobile radio channels," *IEEE Trans. Veh. Technol.*, vol. 59, pp. 974-979, Feb. 2010.

[23] W. C. Jakes, and D. C. Cox, *Microwave Mobile Communications,* Wiley-IEEE Press, 1994.

[24] A. Al-Dweik and Y. Iraqi, "Error probability analysis and applications of amplitude-coherent detection in flat Rayleigh fading channels," *IEEE Trans. Commun.,* vol. 64, no. 5, pp. 2235-2244, May 2016.

[25] J. G. Proakis, and M. Salehi, *Digital communications*, 4th ed. New York: McGraw-Hill, 2001.

[26] M. K. Simon, S. M. Hinedi, and W. C. Lindsey, *Digital communication techniques: signal design and detection*, Prentice Hall PTR, 1995.

[27] Ming-Xian Chang and Y. T. Su, "Performance analysis of equalized OFDM systems in Rayleigh fading," *IEEE Trans. Wireless Commun.*, vol. 1, no. 4, pp. 721-732, Oct. 2002.

[28] ETSI TR 125 943 V9.0.0 (2010-02), Universal Mobile Telecommunications System (UMTS) Deployment Aspects, 3GPP TR 25.943, Release 9.

Chapter 4

Cross-Layer Spectral Efficiency of Adaptive Communications Systems with QoS Constraints

4.1   Introduction

[1]Bandwidth scarcity is one of the chronic limitations that wireless networks have been suffering from since the early days of mobile data emergence. The massive and persistent increase of the traffic volume over the past few years made the network scalability process highly challenging, more frequent and costly. For example, the mobile data traffic in the second quarter of 2015 has increased by a factor of 18 as compared to the second quarter of 2010, and is expected to grow by a factor of 10 between 2015 and 2020 [2]. In other words, the data growth in a single decade can be as much as 180 fold. Consequently, researchers from the academic and industrial sectors are devoting extraordinary efforts to maximize the spectral efficiency of wireless networks by optimizing the spectrum utilization across all layers of the communications protocol

---

[1]A version of this chapter has been published in [1].

stack.

At the physical layer (PHY), tremendous efforts were focused on developing blind channel estimation and synchronization (CE/SY) techniques. For example, various blind carrier frequency offset (CFO) estimation, blind symbol timing offset (STO) estimation and blind channel estimation algorithms are reported in [3]-[9], and [10]-[24], respectively. At the upper layers, system adaptation and dynamic resource allocation have recently received considerable attention due to their potential to improve the spectral efficiency [25]-[30]. Typically, the spectral gain achieved at the PHY is considered independent of the operations performed at the upper layers. Therefore, CE/SY algorithms that do not involve pilot symbols are denoted as blind, and hence are spectrally efficient. However, a plethora of blind CE/SY algorithms have strict constraints regarding the modulation type [3]-[24], which might affect the effectiveness of using various link adaptation techniques at the upper layers, and hence, cross-layer spectral efficiency evaluation is necessary.

Generally speaking, pilot-aided CE/SY are computationally efficient, produce reliable estimates, robust under various operating conditions, and have short observation windows. Therefore, pilot-aided CE/SY is adopted in most current broadband wired/wireless applications such as digital audio broadcasting (DAB) [31], fourth generation (4G) LTE-Advanced [3], and second generation digital video transmission over cable (DVB-C2) [33]. It is worth noting that all these standards have orthogonal frequency division multiplexing (OFDM) as a modulation scheme. However, the low spectral efficiency limitation of pilot-aided techniques has triggered extensive research to develop blind CE/SY techniques [3]-[24].

To the best of the authors knowledge, no work in the literature has considered the overall spectral efficiency across multiple layers. Therefore, this work presents a cross-layer spectral evaluation when particular spectral enhancement techniques are applied at different layers. Without loss of generality, we apply the proposed approach

to compute the relative spectral efficiency in the context of OFDM. However, it can be applied to other systems with different modulation schemes. Contrary to the case where different layers are treated independently, the obtained results show that several blind techniques might degrade the overall spectral efficiency as compared to pilot-aided techniques. Moreover, the results show that the spectral efficiency is a function of the signal-to-noise ratio (SNR), QoS, channel characteristics, and the link adaptation algorithm used.

4.2   System and Channel Models

Because OFDM is widely adopted for several practical systems [31]-[33], it will be used, without loss of generality, in this work to demonstrate the spectral efficiency of several systems with different frame structures, QoS requirements, and system design constraints.

4.2.1   OFDM System and Channel models

In OFDM systems, a sequence of N complex symbols is used to modulate N orthogonal subcarriers during the ℓth OFDM block $\mathbf{d}(\ell) = [d_0(\ell), d_1(\ell), ..., d_{N-1}(\ell)]^T$. However, N_P symbols, denoted as pilots, do not carry information because they are known at the receiver side. The symbols d_k, including the pilots, are usually drawn uniformly from a quadrature amplitude modulation (QAM), phase shift keying (PSK) or amplitude shift keying (ASK) constellation. The sequence of data and pilot symbols is modulated using an N-point inverse discrete Fourier transform (IDFT) process that produces the sequence $\mathbf{x}(\ell) = [x_0(\ell), x_1(\ell), \cdots, x_{N-1}(\ell)]^T$. Thus, $\mathbf{x}(\ell) = \mathbf{F}^H \mathbf{d}(\ell)$, where $\mathbf{F}$ is the normalized $N \times N$ DFT matrix. Then, the cyclic prefix (CP) is created by copying the last N_g samples of the IDFT output and appending them at the beginning of the symbol to be transmitted. Therefore, the transmitted OFDM block

consists of $N_t = N + N_g$ samples. The useful part of the OFDM symbol does not include the N_g prefix samples and has a duration of T_u seconds.

At the receiver front-end, the received signal is applied to a matched filter and is then sampled at a period $T_s = T_u/N$. Assuming that the channel is fixed within one OFDM symbol, dropping the CP samples, and applying the DFT to the received sequence gives,

$$\mathbf{y}(\ell) = \mathbf{H}(\ell)\mathbf{d}(\ell) + \mathbf{z}(\ell) \tag{4.1}$$

where $\mathbf{H}(\ell)$ denotes the channel frequency response during the ℓth OFDM block

$$\mathbf{H}(\ell) = \mathrm{diag}\left([H_0(\ell),\, H_1(\ell),\ldots,\, H_{N-1}(\ell)]^T\right)$$

and $\mathbf{z}(\ell) = [z_0(\ell),\, z_1(\ell),\cdots,\, z_{N-1}(\ell)]^T$ denotes the additive system noise, which is modeled as a white Gaussian process with zero mean and variance σ_z^2.

To maximize the efficiency of OFDM-based communication systems, the modulation types/orders of the information symbols in $\mathbf{d}(\ell)$ are chosen based on the channel matrix $\mathbf{H}(\ell)$ [30], which is assumed to be known at the transmitter side via a feedback channel, and the instantaneous SNR of each subcarrier,

$$\gamma_{\ell,k} = \frac{|H_k(\ell)|^2}{\sigma_z^2}\,|d_k(\ell)|^2. \tag{4.2}$$

However, to minimize the signaling over the feedback channel, and to exploit the time/frequency correlation of the channel, the channel information is grouped into blocks, each of which has N_F subcarriers in frequency domain and N_T subcarriers in time domain, which forms one resource block of size $N_B = N_F \times N_T$. Therefore, all subcarriers within a particular block are assigned to the same modulation type/order. In LTE, the resource block for FDD 1.4 MHz with normal CP has $N_F = 12$ and $N_T = 7$, and hence $N_B = 84$. More generally, each subcarrier can be modulated

using different modulation type and order.

4.2.2 Spectral Efficiency of OFDM Systems

Generally speaking, the spectral efficiency ξ of a communications system is defined as $\xi = R_B/W$, where R_B is the bit rate and W is the bandwidth, and hence, the unit of ξ is bit/s/Hz. Consequently, the relative spectral efficiency of two different systems can be defined as $\eta_R = \xi^{(1)}/\xi^{(2)}$. In the case that the two systems utilize the same bandwidth then, $\eta_R = R_B^{(1)}/R_B^{(2)}$. Moreover, if R_B is normalized by the total transmission time, then η_R becomes the ratio of the total number of bits transmitted by the first system to the total number of bits transmitted by the second system.

In OFDM-based systems, the spectral efficiency is usually computed as the ratio of the number of data-bearing subcarriers to the total number of subcarriers, and thus

$$\eta_R = 1 - \frac{N_P}{N}. \tag{4.3}$$

However, such definition is valid only when all subcarriers in the system are modulated using the same modulation type and order. In practice, different subcarriers can be modulated using different modulation schemes and orders. Therefore, the relative spectral efficiency between two OFDM-based systems should be computed as the ratio between the total number of information bits of the first system to those in the second system over one information (resource) block [3]. Therefore, we can define the relative spectral efficiency as,

$$\eta_R = \frac{\eta^{(1)}|\mathbf{B}}{\eta^{(2)}|\mathbf{B}} = \frac{\sum_{\ell=0}^{N_T-1}\sum_{k=0}^{N_F-1}\log_2[M_{n,m}(\ell,k)],\ [n,m]\in\mathcal{M}^{(1)}}{\sum_{\ell=0}^{N_T-1}\sum_{k=0}^{N_F-1}\log_2[M_{n,m}(\ell,k)],\ [n,m]\in\mathcal{M}^{(2)}} \tag{4.4}$$

where

Table 4.1: $M_{n,m}$ example for $n = 1, 2, 3$ and different values of m.

$n \setminus m$	1	2	3	4	5	6	7
1 (ASK)	1	2	4	8	—	—	—
2 (PSK)	1	2	4	8	16	—	—
3 (QAM)	1	—	4	—	16	—	64

Table 4.2: Time-frequency modulation map example.

$\ell \setminus k$	0	1	2	3
0	$2, 1$	$2, m$	$2, m$	$2, 1$
1	n, m	n, m	n, m	n, m
2	$1, m$	$1, m$	$1, m$	$1, m$
3	$2, 1$	$2, m$	$2, m$	$2, 1$

$$
\mathbf{B} = \begin{bmatrix}
b_{0,0} & b_{0,1} & \cdots & b_{0,N_F-1} \\
b_{1,0} & b_{1,1} & \cdots & b_{1,N_F-1} \\
\vdots & \vdots & \ddots & \vdots \\
b_{N_T-1,0} & b_{N_T-1,1} & \cdots & b_{N_T-1,N_F-1}
\end{bmatrix}. \tag{4.5}
$$

Each element in $\mathbf{B}$ represents the number of bits allocated for the corresponding subcarrier, $M_{n,m}(\ell, k)$ is the modulation order at a given location (ℓ, k), and for a configuration $[n, m]$, where n and m denote the modulation type and order sequence number, respectively. The set $\mathcal{M}^{(\cdot)}$ is the set of all possible values of n and m for a particular system. For example, assume that a particular system supports three different modulation schemes with different modulation orders as depicted in Table 4.1, and the time-frequency grid has $N_F = 4$ and $N_T = 4$. Consequently, the modulation map $\mathcal{M}$ will have the structure given in Table 4.2. As it can be noted from the table, the four subcarriers in the corners of the table carry no information because $m = 1$, the remaining subcarriers in row-0 and row-3 are limited to PSK, but with any order. The subcarriers in the first row can have any combination of $[n, m]$, and the subcarriers in the second row are limited to ASK modulation with any m value.

In practical systems, the map $\mathcal{M}$ is specified at the initial stages of the system design, and then, the values of n and m are dynamically selected based on the system

QoS requirements, the system resources, and the channel state information (CSI) [27]. Without loss of generality, consider the case where the values of m and n can be selected dynamically with the aim of maximizing the spectral efficiency of a particular system under bit error rate (BER), and modulation type/order constraints. Therefore, the problem can be formulated as

$$\max_{m,n} \sum_{\ell=0}^{N_T-1} \sum_{k=0}^{N_F-1} \log_2\left[M_{n,m}\left(\ell,k\right)\right] \tag{4.6a}$$

Subject to:

$$[n,m] \in \mathcal{M} \tag{4.6b}$$

$$\bar{P} \leq P_T \tag{4.6c}$$

where (4.6b) is used to guarantee that the system uses only the allowed modulation types and orders, and (4.6c) is used to guarantee that the average BER $\bar{P}$ is less than a prescribed threshold P_T,

$$\bar{P} = \frac{\sum_{\ell=0}^{N_T-1} \sum_{k=0}^{N_F-1} \log_2\left(M_{n,m}\left(\ell,k\right)\right) P_{n,m}(\gamma_{\ell,k})}{\sum_{\ell=0}^{N_T-1} \sum_{k=0}^{N_F-1} \log_2\left(M_{n,m}\left(\ell,k\right)\right)} \leq P_T \tag{4.7}$$

where $P_{n,m}(\gamma_{\ell,k})$ is the instantaneous BER for a given $[n,m] \in \mathcal{M}$ and $\gamma_{\ell,k}$. In typical bit loading problems, $\bar{P}$ is computed with the assumption of perfect CSI knowledge at the transmitter. In spectral efficiency analysis, the accuracy of the algorithm, SNR and spectral efficiency are correlated. For example, two blind CE algorithms with different accuracy would actually have different spectral efficiency values.

As $\eta^{(i)}$, $i \in \{1,2\}$ in (4.4) is conditioned on $\mathbf{B}$, the conditioning can be removed by averaging over all possible values of $\mathbf{B}$,

$$\eta^{(i)} = \sum_{\mathbf{B}} \eta^{(i)}|\mathbf{B} \ \Pr\left(\mathbf{B}\right), \quad i \in \{1,2\}. \tag{4.8}$$

However, the set $\mathbf{B}$ is selected based on the optimization problem given in (4.6a), which can be solved only using particular searching algorithms, and thus, it will be infeasible to compute $\Pr(\mathbf{B})$. To simplify the discussion, we consider a sub-optimal solution of $\mathbf{B}$ where the modulation order for each subcarrier is assigned independently of all other subcarriers based on its SNR, this approach will be denoted as the basic allocation (BA) algorithm. In such scenarios, the BER constraint is satisfied when the BER for each subcarrier is satisfied. Thus, the modulation order is selected such that

$$\Pr(b_{\ell,k} = q) = \Pr(\zeta_{q-1} < \gamma_{\ell,k} < \zeta_{q+1}) \tag{4.9}$$

where ζ_q^- and ζ_q^+ are the minimum and maximum SNR values required to chose a modulation order $M = 2^q$, which satisfies the BER constraint. Since the SNR ranges are contiguous and non-overlapping, then (4.9) can be written as $\Pr(b_{\ell,k} = q) = \Pr(\gamma_{\ell,k} \in \Gamma_q)$. Therefore,

$$\Pr(\mathbf{B} = \mathbf{q}) = \Pr\Big(\gamma_{0,0} \in \Gamma_{q_{0,0}}, \ \ldots, \ \gamma_{0,N_F-1} \in \Gamma_{q_{0,N_F-1}},$$

$$\gamma_{1,0} \in \Gamma_{q_{1,0}}, \ \ldots, \ \gamma_{1,N_F-1} \in \Gamma_{q_{1,N_F-1}}$$

$$\vdots$$

$$\gamma_{N_T-1,0} \in \Gamma_{q_{0,0}}, \ \ldots, \ \gamma_{N_T-1,N_F-1} \in \Gamma_{q_{N_T-1,N_F-1}}\Big). \tag{4.10}$$

Since the instantaneous SNRs are correlated in time and frequency domain, the probability $\Pr(\mathbf{B} = \mathbf{q})$ can be expressed as

$$\Pr(\mathbf{B} = \mathbf{q}) =$$

$$\int_{\zeta_{q_{0,0}}^-}^{\zeta_{q_{0,0}}^+} \cdots \int_{\zeta_{q_{N_T-1,N_F-1}}^-}^{\zeta_{q_{N_T-1,N_F-1}}^+} P\left(\gamma_{0,0}, \ \ldots, \ \gamma_{N_T-1,N_F-1}\right)$$

$$d\gamma_{N_T-1,N_F-1}...d\gamma_{0,0} \tag{4.11}$$

where $P\left(\gamma_{0,0}, ..., \gamma_{N_T-1,N_F-1}\right)$ is the joint probability distribution function (PDF) of the SNRs within the block. Since the marginal PDF of $\gamma_{\ell,k}$ is exponential, then the joint PDF $P\left(\gamma_{0,0}, ..., \gamma_{N_T-1,N_F-1}\right)$ is a multivariate exponential, which has no closed-form representation and can only be expressed by $N_F \times N_T$ fold integral. Therefore, evaluating (4.11) for a particular $\mathbf{q}$ requires $2N_F \times N_T$ fold integral. Which is prohibitively complex to be evaluated, even numerically.

4.3   Numerical Results
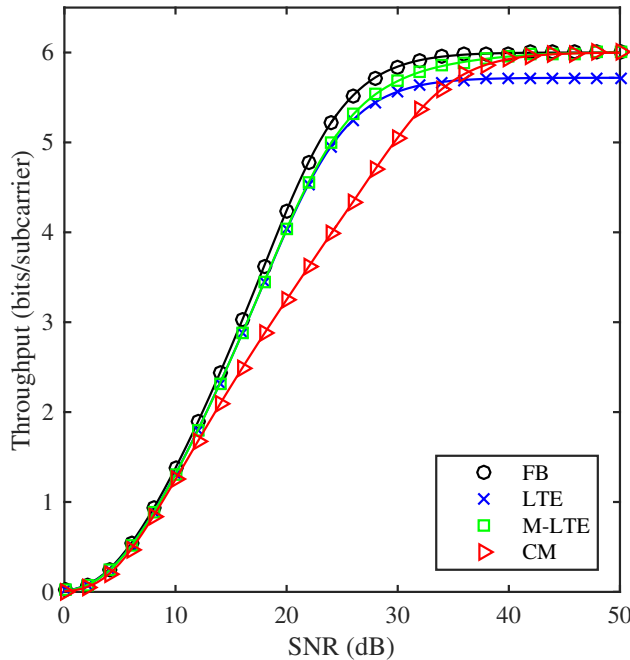
Figure 4.1: Average throughput per subcarrier using IA algorithm for $P_T = 10^{-3}$.

In this section, simulation results are presented to evaluate the average throughput per subcarrier and the relative spectral efficiency. The channel is assumed to be frequency-selective quasi-static with Rayleigh fading, where the channel remains fixed within one OFDM symbol, but changes randomly over consecutive symbols. The channel model considered in this work is the typical urban (TUx) multipath fading
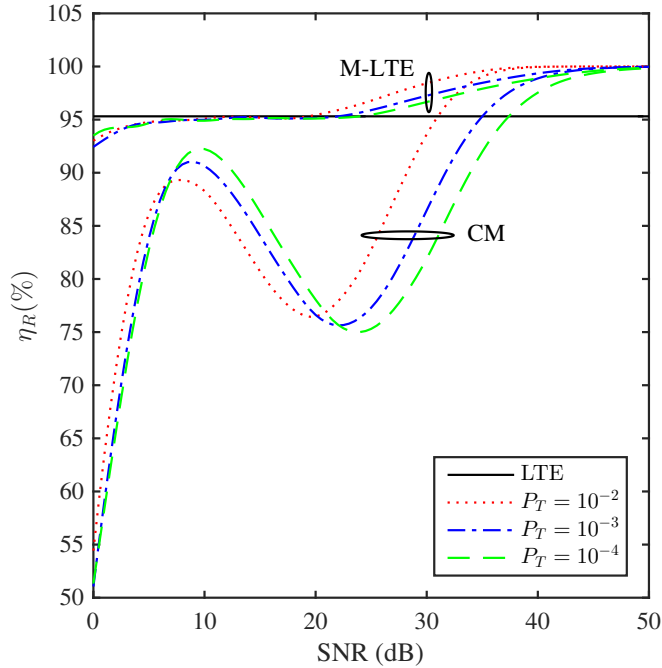
Figure 4.2: Relative spectral efficiency using IA algorithm for an OFDM system with different modulation constraints.

channel model [45] that consists of 9 taps with normalized delays of $[0, 1, ..., 8]$ and average normalized taps' gains of $[2.69, 1.74, 2.89, 1.17, 0.23, 0.58, 0.36, 0.26, 0.08]/10$.

The spectral efficiency for four different systems is considered, which are the fully blind (FB), constant modulus (CM), LTE and the modified LTE (M-LTE) [3], [34]. The FB system is similar to LTE except that no pilots are used. The CM has no pilots, but the modulation is limited to PSK. The M-LTE is similar LTE except that pilot symbols are replaced by unipolar ASK and one of the subcarriers adjacent to the pilot should have PSK symbols. Moreover, the modulation order M can be set to one to satisfy the BER requirements. In all considered systems, the appropriate modulation order is selected such that the average BER is less than P_T. The modulation orders for all subcarriers are computed using the Incremental Allocation (IA) Algorithm [36]. The spectral efficiency of the FB system is considered as $\eta^{(2)}$ when η_R is computed,
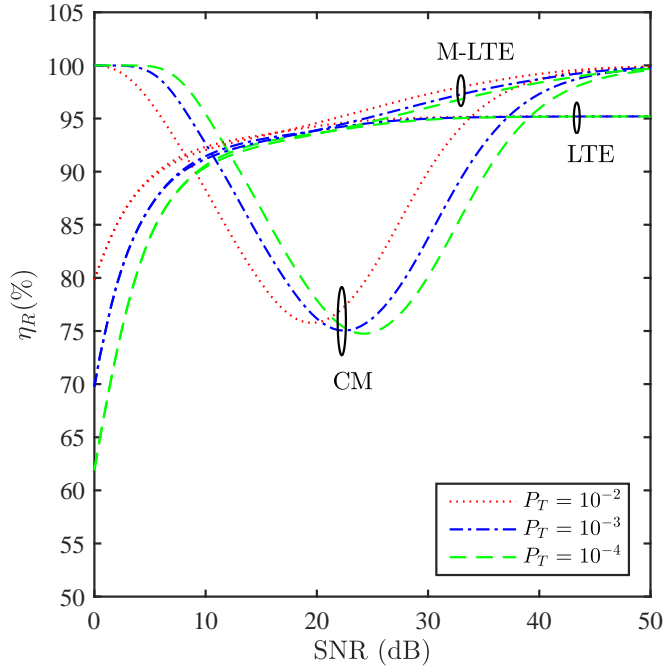
Figure 4.3: Relative spectral efficiency using BA algorithm for an OFDM system with different modulation constraints.

because FB has the maximum spectral efficiency.

Fig. 4.1 presents the average throughput per subcarrier, for the FB, CM, LTE, and M-LTE [3] systems. As it can be noted from the figure, the FB system outperforms all other systems since it does not require pilots, and it has no modulation-type constraint. Unlike what is usually assumed, the LTE outperforms CM systems for a wide range of SNRs. Therefore, sacrificing a few subcarriers as pilots and selecting the modulation type freely for other subcarriers results in higher throughput as compared to the case where all subcarriers carry information, but have the CM constraint. The M-LTE throughput is equivalent to LTE at low SNRs, but it shows higher throughput at high SNRs.

The relative spectral efficiency η_R of the CM, LTE, and M-LTE systems is presented in Fig. 4.2 for BER thresholds $P_T = 10^{-2}$, 10^{-3} and 10^{-4}. As it can be noted from the figure, the LTE system has a constant spectral efficiency of about

95%, where the 5% loss is caused by the pilots. Surprisingly, the figure shows that LTE outperforms the blind M-LTE and CM system for low to medium SNRs. The figure also shows that the spectral efficiency depends on SNR and P_T because both parameters affect the selection of the modulation orders for the different systems. For example, in the range of low SNRs, we note that η_R for the CM system is increasing as a function of SNR, which is due to the fact that majority of the FB system subcarriers at this range of SNRs are modulated using $M = 1$, 2 or 4, which is similar to the CM case. In the mid-rage SNRs, more subcarriers in the FB system will start to use 16-QAM, while the CM is mostly limited to $M \leq 8$, and hence, η_R decreases. Finally, at high SNRs, the FB will be mostly using 64-QAM, which is the maximum allowed modulation order, and hence $\eta^{(1)}$ of the CM will eventually approach $\eta^{(2)}$ of the FB system. Similar to the CM case, the LTE outperforms the M-LTE at low SNRs. However, the difference is negligible. At high SNRs, the M-LTE outperforms the LTE noticeably.

On the other hand, Fig. 4.3 depicts the relative spectral efficiency evaluated using BA algorithm discussed in Subsection 4.2.2 for different BER thresholds. As the figure indicates, the CM system has a spectral efficiency of about 100% around SNR $= 0$ dB. In fact, BA algorithm forces all subcarriers to use BPSK and QPSK modulation in the CM and FB systems at low SNRs in order to satisfy the BER P_T constraint for each subcarrier, while nulling the ASK symbols in the M-LTE, as they do not satisfy constraint with the minimum modulation order. As a result, the relative spectral efficiency of the M-LTE is similar to the LTE system at low SNRs. However, at moderate and high values of SNRs, the BA algorithm converges to the same allocation obtained by IA algorithm, which demonstrates the significance of optimizing the bit allocation scheme to the system's relative spectral efficiency.

4.4   Chapter Summary

In this chapter, the concept of spectral efficiency of blind CE/SY techniques was revisited, where we proposed a new fair and reliable approach to compute and compare the spectral efficiency of various blind and non-blind communications systems. The new approach considers the fact that different subcarriers in OFDM systems may be modulated using different modulation types and orders to satisfy QoS requirements. Moreover, the proposed approach considers the modulation type constraint on the overall system spectral efficiency. The obtained results showed that the modulation type constraint has a significant impact on the system spectral efficiency, which can make the spectral efficiency of pilot-aided systems higher than that of blind systems with modulation type constraint.

Bibliography

[1] A. Saci, A. Shami and A. Al-Dweik, "Cross-layer spectral efficiency of adaptive communications systems with QoS constraints," *IEEE 86th Veh. Technol. Conf. (VTC-Fall)*, Toronto, ON, 2017.

[2] https://www.ericsson.com/res/docs/2015/ericsson-mobility-report-june-2015.pdf.

[3] A. Al-Dweik, A. Hazmi, S. Younis, B. Sharif, and C. Tsimenidis, "Carrier frequency offset estimation for OFDM systems over mobile radio channels," *IEEE Trans. Veh. Technol.*, vol. 59, pp. 974-979, Feb. 2010.

[4] X. Zeng and A. Ghrayeb, "A blind carrier frequency offset estimation scheme for OFDM systems with constant modulus signaling," *IEEE Trans. Commun.*, vol. 56, no. 7, pp. 1032-1037, July 2008.

[5] A. Al-Dweik, A. Hazmi, S. Younis, B. Sharif, and C. Tsimenidis, "Blind iterative frequency offset estimator for OFDM systems," *IET Commun.*, vol. 4, issue 16, pp. 2008-2019, Nov. 2010.

[6] J. Oh, J. Kim and J. Lim, "Blind carrier frequency offset estimation for OFDM systems with constant modulus constellations," *IEEE Commun. Lett.*, vol. 15, no. 9, pp. 971-973, Sep. 2011.

[7] W. Zhang and Q. Yin, "Blind carrier frequency offset estimation for MIMO-OFDM with constant modulus constellations via rank reduction criterion," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6809-6815, Aug. 2016.

[8] A. Jayaprakash and G. Reddy, "Covariance fitting based blind carrier frequency offset estimation method for OFDM systems," *IEEE Trans. Veh. Technol.*, IEEE early access, doi: 10.1109/TVT.2016.2542181, 2016.

[9] A. Al-Dweik, S. Younis, A. Hazmi, B. Sharif and C. Tisimendis, "An efficient OFDM symbol timing estimator using power difference measurements," *IEEE Trans. Veh. Technol.*, vol. 61, no. 2, pp. 509-520, Feb. 2012.

[10] M. Necker and G. Stuber, "Totally blind channel estimation for OFDM on fast varying mobile radio channels," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1514-1525, Sep. 2004.

[11] E. Moulines, P. Duhamel, J. F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Process.,* vol. 43, no. 2, pp. 516-525, Feb. 1995.

[12] C.-C. Tu and B. Champagne, "Subspace blind MIMO-OFDM channel estimation with short averaging periods: Performance analysis," *Proc. IEEE Wireless Commun. Netw. Conf.*, Las Vegas, NV, Apr. 2008, pp. 24-29.

[13] C.-C. Tu and B. Champagne, "Subspace-based blind channel estimation for MIMO-OFDM systems with reduced time averaging," *IEEE Trans. Veh. Technol.,* vol. 59, no. 3, pp. 1539-1544, Mar. 2010.

[14] F. Gao,Y. Zeng, A.Nallanathan, and T. S.Ng, "Robust subspace blind channel estimation for cyclic prefixed MIMO OFDM systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 2, pp. 378-388, Feb. 2008.

[15] F. Gao and A. Nallanathan, "Blind channel estimation for MIMO OFDM systems via nonredundant linear precoding," *IEEE Trans. Signal Process.,* vol. 55, no. 2, pp. 784-789, Feb. 2007.

[16] H. Muarkami, "Blind estimation of a fractionally sampled FIR channel for OFDM transmission using residue polynomials," *IEEE Trans. Signal Process.,* vol. 54, no. 1, pp. 225-234, Jan. 2006.

[17] S. Banani and R. G. Vaughan, "OFDM with iterative blind channel estimation," *IEEE Trans. Veh. Technol.,* vol. 59, no. 9, Nov. 2010.

[18] C. Li and S. Roy, "Subspace-based blind channel estimation for OFDM by exploiting virtual carriers," *IEEE Trans. Wireless Commun.,* vol. 2, no. 1, pp. 141-150, Jan. 2003.

[19] N. Sarmadi, S. Shahbazpanahi, and A. Greshman, "Blind channel estimation in orthogonally coded MIMO-OFDM systems," *IEEE Trans. Signal Process.,* vol. 57, no. 6, pp. 2354-2364, Jun. 2009.

[20] W. Ma, B. Vo, T. Davidson, and P. Ching, "Blind ML detection of orthogonal space–time block codes: Efficient high-performance implementations," *IEEE Trans. Signal Process.,* vol. 54, no. 2, pp. 738-751, Jan. 2006.

[21] P. Stoica and G. Ganesan, "Space-time block codes: Trained, blind and semi-blind detection," *in Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing* (ICASSP), USA, 2002, pp. 1609-1612.

[22] T. Al-Naffouri and A. Quadeer, "Cyclic prefix based enhanced data recovery in OFDM," *IEEE Trans. Signal Process.,* vol. 58, no. 6, pp. 3406–3410, Jun. 2010.

[23] Y. Li, C. Georghiades, and G. Huang, "Iterative maximum likelihood sequence estimation for space–time coded systems," *IEEE Trans. Commun.,* vol. 49, no. 6, pp. 948-951, June 2001.

[24] T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "Maximum-likelihood detection of orthogonal space–time block coded OFDM in unknown block fading channels," *IEEE Trans. Signal Process.,* vol. 56, no. 4, pp. 1637–1649, Apr. 2008.

[25] L. Li, B. Li, H. Li, C. Chen, "Domain optimal bit allocation algorithm for high efficiency video coding," *IEEE Trans. Circuits Syst. Video Technol, IEEE Xplore early Access,* doi: 10.1109/TCSVT.2016.2598672, 2016.

[26] W. Hachicha, M. Kaaniche, A. Beghdadi and F. Cheikh, "Efficient inter-view bit allocation methods for stereo image coding," *IEEE Trans. Multimedia,* vol. 17, no. 6, pp. 765-777, June 2015.

[27] M. Kalil, A. Shami, A. Al-Dweik, and S. Muhaidat, "Low-complexity power-efficient schedulers for LTE uplink with delay-sensitive traffic," *IEEE Trans. Veh. Technol,* vol. 64, no. 10, pp. 4551-4564, Oct. 2015.

[28] T. Al-Naffouri, A. Dahman, M. Sohail, W. Xu and B. Hassibi, "Low-complexity blind equalization for OFDM systems with general constellations," *IEEE Trans. Signal Process.,* vol. 60, no. 12, pp. 6395-6407, Dec. 2012.

[29] M. Kalil, A. Shami and A. Al-Dweik, "QoS-aware power-efficient scheduler for LTE uplink," *IEEE Trans. Mobile Comput.,* vol. 14, no. 8, pp. 1672-1685, Aug. 2015.

[30] N. Abo Aly, A. Al-Dweik and M. Al-Mualla, "Adaptive OFDM system with limited feedback using truncated channel impulse response," *in Proc. 15th IEEE Int. Symposium on Signal Processing and Information Technology* (ISSPIT), UAE, 2015, pp. 197-202.

[31] Radio broadcasting systems; digital audio broadcasting (DAB) to mobile, portable and fixed receivers, ETS Standard 300 401, 1995.

[32] LTE; Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description, 3GPP Standard TS 36.300, 2011.

[33] Digital video broadcasting (DVB); frame structure channel coding and modulation for a second generation digital transmission system for cable systems (DVB-C2), ETSI Standard EN 302 769, 2010.

[34] A. Al-Dweik and Y. Iraqi, "Error probability analysis and applications of amplitude-coherent detection in flat rayleigh fading channels", *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 2235-2244, May 2016.

[35] ETSI TR 125 943 V9.0.0 (2010-02), Universal Mobile Telecommunications System (UMTS) Deployment Aspects, 3GPP TR 25.943 version 9.0.0 Release 9.

[36] A. Wyglinski, F. Labeau, and P. Kabal, "Bit loading with BER-constraint for multicarrier systems," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1383-1387, July 2005.

Chapter 5

Direct Data Detection of OFDM Signals Over Wireless Channels

5.1 Preamble

[1]Orthogonal frequency division multiplexing (OFDM) is widely adopted in several wired and wireless communication standards, such as worldwide interoperability for microwave access (WiMAX) technologies [2], Long Term Evolution-Advanced (LTE-A) standard [3], Digital Video Broadcasting (DVB), Terrestrial (DVB-T) and Hand-held (DVB-H) [4], optical wireless communications (OWC) [5], [6], and recently, it has been adopted for the fifth-generation (5G) wireless networks [7]. The channel is typically modeled as frequency-selective for WiMax and LTE-A, flat for OWC in the presence of atmospheric turbulence [5], [6]. Therefore, OFDM has become the lead above other modulation schemes at present and in the near future [8].

One of the main advantages of OFDM is that each subcarrier experiences flat fading even though the overall signal spectrum suffers from frequency-selective fading. Moreover, incorporating the concept of a cyclic prefix (CP), which is formed

---

[1]A version of this chapter has been submitted to IEEE Transactions on Vehicular Technology. The manuscript of paper in available in the arXiv [1].

by copying a part of the OFDM symbol of and pre-append it to the transmitted OFDM block, prevents intersymbol interference (ISI) if the CP length is larger than the maximum delay spread of the channel. Consequently, a low-complexity single-tap equalizer can be utilized to eliminate the impact of the multipath fading channel. Under such circumstances, the OFDM demodulation process can be performed once the fading parameters at each subcarrier, commonly denoted as channel state information (CSI), are estimated.

In general, channel estimation can be classified into blind [9]-[14], and pilot-aided techniques [15]-[21]. Blind channel estimation techniques are spectrally efficient because they do not require any overhead to estimate the CSI, nevertheless, such techniques have not yet been adopted in practical OFDM systems. Conversely, pilot-based CSI estimation is preferred for practical systems, because typically it is more robust and less complex. In pilot-based CSI estimation, the pilot symbols are embedded within the subcarriers of the transmitted OFDM signal in time and frequency domain; hence, the pilots form a two dimensional (2-D) grid [3]. The channel response at the pilot symbols can be obtained using the least-squares (LS) frequency domain estimation, and the channel parameters at other subcarriers can be obtained using various interpolation techniques [22]. Optimal interpolation requires a 2-D Wiener filter that exploits the time and frequency correlation of the channel, however, it is substantially complex to implement [23], [24]. The complexity can be reduced by decomposing the 2-D interpolation process into two cascaded 1-D processes, and then, using less computationally-involved interpolation schemes [25], [26]. Low complexity interpolation, however, is usually accompanied by error rate performance degradation [26]. It is also worth noting that most practical OFDM-based systems utilize a fixed grid pattern structure [3].

Once the channel parameters are obtained for all subcarriers, the received samples at the output of the fast Fourier transform (FFT) are equalized to compensate for the

channel fading. Fortunately, the equalization for OFDM is performed in the frequency domain using single-tap equalizers. The equalizer output samples, which are denoted as the decision variables, will be applied to a maximum likelihood detector (MLD) to regenerate the information symbols.

In addition to the direct approach, several techniques have been proposed in the literature to estimate the CSI or detect the data symbols indirectly, by exploiting the correlation among the channel coefficients. For example, the per-survivor processing (PSP) approach has been widely used to approximate the maximum likelihood sequence estimator (MLSE) for coded and uncoded sequences [27], [28], [29]. The PSP utilizes the Viterbi algorithm (VA) to recursively estimate the CSI without interpolation using the least mean squares (LMS) algorithm. Although the PSP provides superior performance when the channel is flat over the entire sequence, its performance degrades severely if this condition is not satisfied, even when the LMS step size is adaptive [28]. Multiple symbol differential detection (MSDD) can be also used for sequence estimation without explicit channel estimation. In such systems, the information is embedded in the phase difference between adjacent symbols, and hence, differential encoding is needed. Although differential detection is only 3 dB worse than coherent detection in flat fading channels, its performance may deteriorate significantly in frequency-selective channels [30], [31]. Consequently, Wu and Kam [32] proposed a generalized likelihood ratio test (GLRT) receiver whose performance without CSI is comparable to the coherent detector. Although the GLRT receiver is more robust than differential detectors in frequency-selective channels, its performance is significantly worse than coherent detectors.

The estimator-correlator (EC) cross-correlates the received signal with an estimate of the channel output signal corresponding to each possible transmitted signal [33], [34]. The signal at channel output is estimated with a minimum mean square error (MMSE) estimator from the knowledge of the received signal and the second

order statistics of the channel and noise. The channel estimation (CE) may provide BER that is about 1 dB from the ML coherent detector in flat fading channels but at the expense of a large number of pilots. Moreover, the BER performance of EC detectors is generally poor in frequency-selective channels where the CE BER is significantly worse than the ML coherent detector [34]. Decision-directed techniques can also be used to avoid conventional channel estimation. For example, the authors in [11] proposed a hybrid frame structure that enables blind decision-directed channel estimation. Although the proposed system manages to offer reliable channel estimates and BER in various channel conditions, the system structure follows the typical coherent detector design where equalization and symbol detection are required.

5.1.1 Motivation and Key Contributions

Unlike conventional OFDM detectors, this work presents a new detector to regenerate the information symbols directly from the received samples at the FFT output, which is denoted as the direct data detector (D^3). By using the D^3, there is no need to perform channel estimation, interpolation, equalization, or symbol decision operations. The D^3 exploits the fact that channel coefficients over adjacent subcarriers are highly correlated and approximately equal. Consequently, the D^3 is derived by minimizing the difference between channel coefficients of adjacent subcarriers. The main limitation of the D^3 is that it suffers from a phase ambiguity problem, which can be solved using pilot symbols, which are part of a transmission frame in most practical standards [2], [3]. To the best of the authors' knowledge, there is no work reported in the published literature that uses the proposed principle.

   The D^3 performance is evaluated in terms of complexity, computational power, and bit error rate (BER), where analytic expressions are derived for several channel models and system configurations. The D^3 BER is compared to other widely used detectors such as the maximum likelihood (ML) coherent detector [35] with perfect

and imperfect CSI, multiple symbol differential detector (MSDD) [30], the ML sequence detector (MLSD) with no CSI [32], and the per-survivor processing detector [27]. The obtained results show that the D^3 is more robust than all the other considered detectors in various cases of interest, particularly in frequency-selective channels at moderate and high SNRs. Moreover, the computational power comparison shows that the D^3 requires less than 35% of the computational power required by the ML coherent detector.

Notation: In what follows, unless otherwise specified, uppercase boldface and blackboard letters such as $\mathbf{H}$ and $\mathbb{H}$, will denote $N \times N$ matrices, whereas lowercase boldface letters such as $\mathbf{x}$ will denote row or column vectors with N elements. Uppercase, lowercase, or bold letters with a tilde such as $\tilde{d}$ will denote trial values, and symbols with a hat, such as $\hat{\mathbf{x}}$, will denote the estimate of $\mathbf{x}$. Letters with apostrophe such as $\acute{v}$ are used to denote the next value, i.e., $\acute{v} \triangleq v + 1$. Furthermore, $\mathrm{E}\left[\cdot\right]$ denotes the expectation operation.

5.2 Signal and Channel Models

Consider an OFDM system with N subcarriers modulated by a sequence of N complex data symbols $\mathbf{d} = [d_0,\ d_1,\ ....,\ d_{N-1}]^T$. The data symbols are selected uniformly from a general constellation such as M-ary phase shift keying (MPSK) or quadrature amplitude modulation (QAM). In conventional pilot-aided OFDM systems [36], N_P of the subcarriers are allocated for pilot symbols, which can be used for channel estimation and synchronization purposes. The modulation process in OFDM can be implemented efficiently using an N-point inverse FFT (IFFT) algorithm, where its output during the ℓth OFDM block can be written as $\mathbf{x}(\ell) = \mathbf{F}^H \mathbf{d}(\ell)$ where $\mathbf{F}$ is the normalized $N \times N$ FFT matrix, and hence, $\mathbf{F}^H$ is the IFFT matrix. To simplify the notation, the block index ℓ is dropped for the remaining parts of this chapter unless

it is necessary to include it. Then, a CP of length N_{CP} samples, no less than the channel maximum delay spread $(\mathcal{D}_{\mathrm{h}})$, is appended to compose the OFDM symbol with a total length $N_{\mathrm{t}} = N + N_{\mathrm{CP}}$ samples and duration of T_{t}s.

At the receiver front-end, the received signal is down-converted to baseband and sampled at a rate $T_{\mathrm{s}} = T_{\mathrm{t}}/N_{\mathrm{t}}$. In this work, the channel is assumed to be composed of $\mathcal{D}_{\mathrm{h}} + 1$ independent multipath components each of which has a gain $h_m \sim \mathcal{CN}\left(0, 2\sigma_{h_m}^2\right)$ and delay $m \times T_{\mathrm{s}}$, where $m \in \{0,\ 1,...,\ \mathcal{D}_{\mathrm{h}}\}$. A quasi-static channel is assumed throughout this work, and thus, the channel taps are considered constant over one OFDM symbol, but they may change over two consecutive symbols. Therefore, the received sequence after dropping the CP samples and applying the FFT can be expressed as,

$$\mathbf{r} = \mathbf{H}\mathbf{d} + \mathbf{w} \tag{5.1}$$

where $\{\mathbf{r}, \mathbf{w}\} \in \mathbb{C}^{N \times 1}$, $w_v \sim \mathcal{CN}\left(0,\ 2\sigma_w^2\right)$ is the additive white Gaussian noise (AWGN) vector and $\mathbf{H}$ denotes the channel frequency response (CFR)

$$\mathbf{H} = \mathrm{diag}\left\{[H_0,\ H_1, \ldots,\ H_{N-1}]\right\}. \tag{5.2}$$

By noting that $\mathbf{r}|_{\mathbf{H},\mathbf{d}} \sim \mathcal{CN}\left(\mathbf{H}\mathbf{d},\ 2\sigma_w^2 \mathbf{I}_N\right)$ where $\mathbf{I}_N$ is an $N \times N$ identity matrix, then it is straightforward to show that the MLD can be expressed as

$$\hat{\mathbf{d}} = \arg\ \min_{\tilde{\mathbf{d}}}\ \left\|\mathbf{r} - \mathbf{H}\tilde{\mathbf{d}}\right\|^2 \tag{5.3}$$

where $\|\cdot\|$ denotes the Euclidean norm, and $\tilde{\mathbf{d}} = \left[\tilde{d}_0,\ \tilde{d}_1, \ldots,\ \tilde{d}_{N1}\right]^T$ denotes the trial values of $\mathbf{d}$. As can be noted from (5.3), the MLD requires the knowledge of $\mathbf{H}$. Moreover, because (5.3) describes the detection of more than one symbol, it is typically denoted as maximum likelihood sequence detector (MLSD). If the elements of $\mathbf{d}$ are independent, the MLSD can be replaced by a symbol-by-symbol MLD

$$\hat{d}_v = \arg \min_{\tilde{d}_v} \left| r_v - H_v \tilde{d}_v \right|^2 . \tag{5.4}$$

Since perfect knowledge of $\mathbf{H}$ is infeasible, an estimated version of $\mathbf{H}$, denoted as $\hat{\mathbf{H}}$, can be used in (5.3) and (5.4) instead of $\mathbf{H}$. Another possible approach to implement the detector is to equalize $\mathbf{r}$, and then use a symbol-by-symbol MLD. Because the considered system is assumed to have no ISI or intercarrier interference (ICI), then a single-tap frequency-domain zero-forcing equalizer can be used. Therefore, the equalized received sequence can be expressed as,

$$\check{\mathbf{r}} = \left[\hat{\mathbf{H}}^H \hat{\mathbf{H}} \right]^{-1} \hat{\mathbf{H}}^H \mathbf{r} \tag{5.5}$$

and

$$\hat{d}_v = \arg \min_{\tilde{d}_v} \left| \check{r}_v - \tilde{d}_v \right|^2 , \ \forall v. \tag{5.6}$$

It is interesting to note that solving (5.3) does not necessarily require the explicit knowledge of $\mathbf{H}$ under some special circumstances. For example, Wu and Kam [32] noticed that in flat fading channels, i.e., $H_v = H \ \forall v$, it is possible to detect the data symbols using the following MLSD,

$$\hat{\mathbf{d}} = \arg \max_{\tilde{\mathbf{d}}} \frac{\left| \tilde{\mathbf{d}}^H \mathbf{r} \right|^2}{\| \tilde{\mathbf{d}} \|} . \tag{5.7}$$

Although the detector described in (5.7) is efficient in the sense that it does not require the knowledge of $\mathbf{H}$, its BER is very sensitive to the channel variations.

5.3   Proposed D^3 System Model

One of the distinctive features of OFDM is that its channel coefficients over adjacent subcarriers in the frequency domain are highly correlated and approximately equal. The correlation coefficient between two adjacent subcarriers can be defined as

$$\begin{aligned}
\varrho_f &\triangleq \mathrm{E}\left[H_v H_{\acute{v}}^*\right] \\
&= \mathrm{E}\left[\sum_{n=0}^{\mathcal{D}_{\mathrm{h}}} h_n e^{-j2\pi\frac{nv}{N}} \sum_{m=0}^{\mathcal{D}_{\mathrm{h}}} h_m^* e^{j2\pi\frac{m\acute{v}}{N}}\right] = \sum_{m=0}^{\mathcal{D}_{\mathrm{h}}} \sigma_{h_m}^2 e^{j2\pi\frac{m}{N}}
\end{aligned} \tag{5.8}$$

where $\sigma_{h_m}^2 = \mathrm{E}\left[|h_m|^2\right]$. The difference between two adjacent channel coefficients is

$$\Delta_f = \mathrm{E}\left[H_v - H_{\acute{v}}\right] = \mathrm{E}\left[\sum_{m=0}^{\mathcal{D}_{\mathrm{h}}} h_n e^{-j2\pi\frac{mv}{N}}\left(1 - e^{-j2\pi\frac{m}{N}}\right)\right] \tag{5.9}$$

For large values of N, it is straightforward to show that $\varrho_f \to 1$ and $\Delta_f \to 0$. Similar to the frequency domain, the time domain correlation defined according to the Jakes' model can be computed as [37],

$$\varrho_t = \mathrm{E}\left[H_v^\ell \left(H_{\acute{v}}^{\acute{\ell}}\right)^*\right] = J_0\left(2\pi f_d T_{\mathrm{t}}\right) \tag{5.10}$$

where $J_0\left(\cdot\right)$ is the Bessel function of the first kind and 0 order, f_d is the maximum Doppler frequency. For large values of N, $2\pi f_d T_{\mathrm{s}} \ll 1$, and hence $J_0\left(2\pi f_d T_{\mathrm{s}}\right) \approx 1$, and thus $\varrho_t \approx 1$. Using the same argument, the difference in the time domain $\Delta_t \triangleq \mathrm{E}\left[H_v^\ell - H_{\acute{v}}^{\acute{\ell}}\right] \approx 0$. Although the proposed system can be applied in the time domain, frequency domain, or both, the focus of this work is the frequency domain.

Based on the aforementioned properties of OFDM, a simple approach to extract the information symbols from the received sequence $\mathbf{r}$ can be designed by minimizing the difference of the channel coefficients between adjacent subcarriers, which can be expressed as

$$\hat{\mathbf{d}} = \arg\min_{\tilde{\mathbf{d}}} \sum_{v=0}^{N-2} \left|\frac{r_v}{\tilde{d}_v} - \frac{r_{\acute{v}}}{\tilde{d}_{\acute{v}}}\right|^2. \tag{5.11}$$

As can be noted from (5.11), the estimated data sequence $\hat{\mathbf{d}}$ can be obtained without the knowledge of $\mathbf{H}$. Moreover, there is no requirement for the channel coefficients over the considered sequence to be equal, and hence, the D^3 should perform

fairly well even in frequency-selective fading channels. Nevertheless, it can be noted that (5.11) does not have a unique solution because $\mathbf{d}$ and $-\mathbf{d}$ can minimize (5.11). To resolve the phase ambiguity problem, one or more pilot symbols can be used as a part of the sequence $\mathbf{d}$. In such scenarios, the performance of the D^3 will be affected indirectly by the frequency selectivity of the channel because the capability of the pilot to resolve the phase ambiguity depends on its fading coefficient. Another advantage of using pilot symbols is that it will not be necessary to detect the N symbols simultaneously. Instead, it will be sufficient to detect $\mathcal{K}$ symbols at a time, which can be exploited to simplify the system design and analysis.

Using the same approach of the frequency domain, the D^3 can be designed to work in the time domain as well by minimizing the channel coefficients over two consecutive subcarriers, i.e., two subcarriers with the same index over two consecutive OFDM symbols, which is also applicable to single carrier systems. It can be also designed to work in both time and frequency domains, where the detector can be described as

$$\hat{\mathbf{D}}_{\mathcal{L},\mathcal{K}} = \arg \min_{\tilde{\mathbf{D}}_{\mathcal{L},\mathcal{K}}} J\left(\tilde{\mathbf{D}}_{\mathcal{L},\mathcal{K}}\right) \tag{5.12}$$

where $\mathbf{D}_{\mathcal{L},\mathcal{K}}$ is an $\mathcal{L} \times \mathcal{K}$ data matrix, $\mathcal{L}$ and $\mathcal{K}$ are the time and frequency detection window size, and the objective function $J\left(\tilde{\mathbf{D}}\right)$ is given by

$$J\left(\tilde{\mathbf{D}}_{\mathcal{L},\mathcal{K}}\right) = \sum_{\ell=0}^{\mathcal{L}-1}\sum_{v=0}^{\mathcal{K}-2}\left|\frac{r_v^\ell}{\tilde{d}_v^\ell} - \frac{r_{\acute{v}}^\ell}{\tilde{d}_{\acute{v}}^\ell}\right|^2 + \left|\frac{r_v^\ell}{\tilde{d}_v^\ell} - \frac{r_v^{\acute{\ell}}}{\tilde{d}_v^\ell}\right|^2. \tag{5.13}$$

For example, if the detection window size is chosen to be the LTE resource block, then, $\mathcal{L} = 14$ and $\mathcal{K} = 12$. Moreover, the system presented in (5.13) can be extended to the multi-branch receiver scenarios, single-input multiple-output (SIMO) as,

$$\hat{\mathbf{D}} = \arg \min_{\tilde{\mathbf{d}}} \sum_{n=1}^{\mathcal{N}}\sum_{\ell=0}^{\mathcal{L}-1}\sum_{v=0}^{\mathcal{K}-2}\left|\frac{r_v^{\ell,n}}{\tilde{d}_v} - \frac{r_{\acute{v}}^{\ell,n}}{\tilde{d}_{\acute{v}}}\right|^2 + \left|\frac{r_v^{\ell,n}}{\tilde{d}_v^\ell} - \frac{r_v^{\acute{\ell},n}}{\tilde{d}_v^\ell}\right|^2 \tag{5.14}$$

where $\mathcal{N}$ is the number of receiving antennas.

5.4   Efficient Implementation of D^3

It can be noted from (5.12) and (5.13) that solving for $\hat{\mathbf{D}}$, given that N_P pilot symbols are used, requires an $M^{\mathcal{KL}-N_P}$ trials if brute force search is adopted, which is prohibitively complex, and thus, reducing the computational complexity is crucial. Towards this goal, the two dimensional (2-D) resource block (RB) can be divided into a number of one-dimensional (1-D) segments in time and frequency domains in order to reduce the complexity from order $\mathcal{O}\left(M^{\mathcal{K}\times\mathcal{L}-N_P}\right)$ to $\mathcal{O}\left(M\times(\mathcal{KL}-N_P)\right)$. In order words, the time complexity evolves exponentially as the detection size increases in the 2-D block, while it grows linearly in the cascaded 1-D block, which is significant complexity reduction. Fig. 5.1 shows an example of decomposing the 2-D LTE-A RB into several 1-D segments over time and frequency. As can be noted from the figure, the RB consists of 168 subcarriers among which 8 subcarriers are pilots. It is worth noting that there are some rows and columns in the RB that do not have pilots, and thus, the detection of the entire block can be performed as described in Subsection 5.4.2.

5.4.1   The Viterbi Algorithm (VA)

By noting that the expression in (5.11) corresponds to the sum of correlated terms, which can be modeled as a first-order Markov process, then MLSD techniques such as the VA can be used to implement the D^3 efficiently. For example, the trellis diagram of the VA with binary phase shift keying (BPSK) is shown in Fig. 5.2, and can be implemented as follows:

1. Initialize the path metrics $\left\{\Gamma_0^U, \acute{\Gamma}_0^U, \Gamma_0^L, \acute{\Gamma}_0^L\right\} = 0$, where U and L denote the upper and lower branches, respectively. Since BPSK is used, the number of states is 2.

2. Initialize the counter, $c = 0$.

103
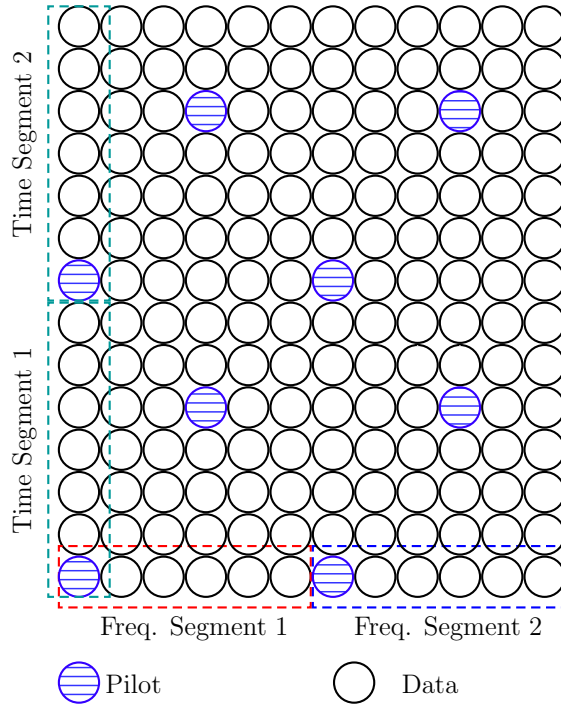
Figure 5.1: Example of a 1-D segmentation over the frequency domain for an LTE-A resource block.

3. Compute the branch metric $J^c_{m,n} = \left| \frac{rc}{m} - \frac{r_{\acute{c}}}{n} \right|^2$, where m is current symbol index, $m = 0 \rightarrow \tilde{d} = -1$, and $m = 1 \rightarrow \tilde{d} = 1$, and n is the next symbol index using the same mapping as m.

4. Compute the path metrics using the following rules,

$$\Gamma^U_{\acute{c}} = \min\left[\Gamma^U_c, \acute{\Gamma}^U_c\right] + J^c_{00} \qquad \Gamma^L_{\acute{c}} = \min\left[\Gamma^L_c, \acute{\Gamma}^L_c\right] + J^c_{01}$$
$$\acute{\Gamma}^U_{\acute{c}} = \min\left[\Gamma^U_c, \acute{\Gamma}^U_c\right] + J^c_{10} \qquad \acute{\Gamma}^L_{\acute{c}} = \min\left[\Gamma^L_c, \acute{\Gamma}^L_c\right] + J^c_{11}$$

5. Track the surviving paths, 2 paths in the case of BPSK.

6. Increase the counter, $c = c + 1$.

7. if $c = \mathcal{K}$, the algorithm ends. Otherwise, go to step 3.

It is worth mentioning that placing a pilot symbol at the edge of a segment terminates the trellis. To simplify the discussion, assume that the pilot value is -1, and thus we compute only $J_{0,0}$ and $J_{1,0}$. Consequently, long data sequences can
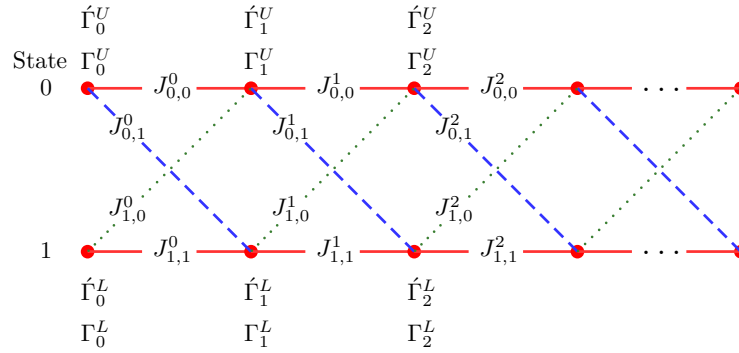
Figure 5.2: Trellis diagram of the D^3 detector for BPSK.

be divided into smaller segments bounded by pilots, which can reduce the delay by performing the detection over the sub-segments in parallel without sacrificing the error rate performance.

5.4.2 Resource Block Detection

As can be noted from Fig. 5.1, the segmentation process can be applied directly to any row or column given that has one or more pilots. Nevertheless, there are some rows and columns that do not have pilots. In such scenarios, the detection, for example, can be performed in two steps as follows:

1. Detect all rows (frequency domain subcarriers) with pilots, i.e., rows 1, 5, 8 and 12.

2. As a result of the first step, each column (time domain subcarrier) has either pilots, data symbols whose values are known as a result of the detection in the first step, or both, as in the case of columns 1, 4, 7 and 10. Therefore, all remaining subcarriers can be detected using the symbols detected in the first step.

It is worth noting that the number and distribution of the pilot symbols in the RB impact the error rate performance, power and spectral efficiency of the system. For example, the first frequency segment shown in Fig. 5.1 consists of seven subcarriers,

two of them are allocated for pilots. By defining the throughput, or the spectral efficiency, as the ratio of the number of information symbols to the total number of symbols per segment, then the throughput of the first frequency and time segments in Fig. 5.1 is about 83.3% and 85.7%, respectively. Nevertheless, the system throughput is determined by the total number of pilots and information subcarriers within an RB rather than a segment. By noting that there are only eight pilots among the 168 resource elements , then the throughput loss is about 4.7% and the throughput is about 95.2%. The same argument can be applied to the power efficiency of the system where 4.7% of the power will be allocated to pilots.

5.4.3   System Design with an Error Control Coding

Forward error correction (FEC) coding can be integrated with the D^3 in two ways, based on the decoding process, i.e., hard or soft decision decoding. For the hard decision decoding, the integration of FEC coding is straightforward where the output of the D^3 is applied directly to the hard decision decoder (HDD).

For the soft decision decoding, we can exploit the coded data to enhance the performance of the D^3, and then use the D^3 output to estimate the channel coefficients in a decision-directed manner. The D^3 with coded data can be expressed as

$$\hat{\mathbf{d}} = \arg \min_{\tilde{\mathbf{u}} \in \mathbb{U}} \sum_{v=0}^{N-2} \left| \frac{r_v}{\tilde{u}_v} - \frac{r_{\acute{v}}}{\tilde{u}_{\acute{v}}} \right|^2 \tag{5.15}$$

where $\mathbb{U}$ is the set of all codewords modulated using the same modulation used at the transmitter. Therefore, the trial sequences $\tilde{\mathbf{u}}$ are restricted to particular sequences. For the case of convolutional codes, the detection and decoding processes can be integrated smoothly since both of them are using the VA. Such an approach can be adopted with linear block codes as well because trellis-based decoding can be also applied to block codes [38].
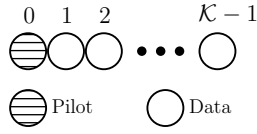
Figure 5.3: Single-sided pilot segment.
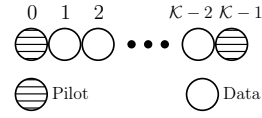
Figure 5.4: Double-sided pilot segment.

5.5   Error Rate Analysis of the D^3

The system BER analysis is presented for several cases according to the pilot and data arrangements. For simplicity, each case is discussed in separate subsections. To make the analysis tractable, we consider BPSK modulation in the analysis while the BER of higher-order modulations is obtained via Monte Carlo simulations.

5.5.1   Single-Sided Pilot

To detect a data segment that contains $\mathcal{K}$ symbols, at least one pilot symbol should be part of the segment in order to resolve the phase ambiguity problem. Consequently, the analysis in this subsection considers the case where there is only one pilot within the $\mathcal{K}$ symbols, as shown in Fig. 5.3. Given that the FFT output vector $\mathbf{r} = [r_0, r_1, \ldots, r_{N-1}]$ is divided into L segments each of which consists of $\mathcal{K}$ symbols, including the pilot symbol, then the frequency domain D^3 detector can be written as,

$$\hat{\mathbf{d}}_l = \arg\min_{\tilde{\mathbf{d}}} \sum_{v=l}^{\mathcal{K}-2+l} \left| \frac{r_v}{\tilde{d}_v} - \frac{r_{\acute{v}}}{\tilde{d}_{\acute{v}}} \right|^2 \quad \mathcal{K} \in \{2, 3, \ldots, N-1\} \tag{5.16}$$

where l denotes the index of the first subcarrier in the segment, and without loss of generality, we consider that $l = 0$. Therefore, by expanding (5.16) we obtain,

$$\hat{\mathbf{d}}_0 = \arg\min_{\tilde{\mathbf{d}}} \left(\frac{r_0}{\tilde{d}_0} - \frac{r_1}{\tilde{d}_1} \right) \left(\frac{r_0}{\tilde{d}_0} - \frac{r_1}{\tilde{d}_1} \right)^* + \cdots + \left(\frac{r_{\mathcal{K}-2}}{\tilde{d}_{\mathcal{K}-2}} - \frac{r_{\mathcal{K}-1}}{\tilde{d}_{\mathcal{K}-1}} \right) \left(\frac{r_{\mathcal{K}-2}}{\tilde{d}_{\mathcal{K}-2}} - \frac{r_{\mathcal{K}-1}}{\tilde{d}_{\mathcal{K}-1}} \right)^*$$

$$\tag{5.17}$$

which can be simplified to,

$$\hat{\mathbf{d}}_0 = \arg\min_{\tilde{\mathbf{d}}} \left|\frac{r_0}{\tilde{d}_0}\right|^2 + \left|\frac{r_1}{\tilde{d}_1}\right|^2 + \cdots + \left|\frac{r_{\mathcal{K}-1}}{\tilde{d}_{\mathcal{K}-1}}\right|^2 - \frac{r_0}{\tilde{d}_0}\frac{r_1}{\tilde{d}_1^*} - \frac{r_0}{\tilde{d}_0^*}\frac{r_1}{\tilde{d}_1} - \cdots$$
$$-\frac{r_{\mathcal{K}-2}}{\tilde{d}_{\mathcal{K}-2}}\frac{r_{\mathcal{K}-1}}{\tilde{d}_{\mathcal{K}-1}^*} - \frac{r_{\mathcal{K}-2}}{\tilde{d}_{\mathcal{K}-2}^*}\frac{r_{\mathcal{K}-1}}{\tilde{d}_{\mathcal{K}-1}}. \quad (5.18)$$

For BPSK, $\left|r_v/\tilde{d}_v\right|^2 = |r_v|^2$, which is a constant term with respect to the maximization process in (5.18), and thus, they can be dropped. Therefore, the detector is reduced to

$$\hat{\mathbf{d}}_0 = \arg\max_{\tilde{\mathbf{d}}_0} \sum_{v=0}^{\mathcal{K}-2} \Re\left\{\frac{r_v r_{\acute{v}}}{\tilde{d}_v \tilde{d}_{\acute{v}}}\right\}. \quad (5.19)$$

Given that the pilot symbol is placed in the first subcarrier and noting that $d_v \in \{-1, 1\}$, then $\tilde{d}_0 = 1$ and $\hat{\mathbf{d}}_0$ can be written as

$$\hat{\mathbf{d}}_0 = \arg\max_{\tilde{d}_0 \notin \tilde{\mathbf{d}}_0} \frac{1}{\tilde{d}_1}\Re\left\{r_0 r_1\right\} + \sum_{v=1}^{\mathcal{K}-2} \frac{1}{\tilde{d}_v \tilde{d}_{\acute{v}}}\Re\left\{r_v r_{\acute{v}}\right\}. \quad (5.20)$$

The sequence error probability (P_S), conditioned on the channel frequency response over the $\mathcal{K}$ symbols ($\mathbf{H}_0$) and the transmitted data sequence $\mathbf{d}_0$ can be defined as,

$$P_S|_{\mathbf{H}_0,\mathbf{d}_0} \triangleq \Pr\left(\hat{\mathbf{d}}_0 \neq \mathbf{d}_0\right)\Big|_{\mathbf{H}_0,\mathbf{d}_0} \quad (5.21)$$

which can be also written in terms of the conditional probability of correct detection P_C as,

$$P_C|_{\mathbf{H}_0,\mathbf{d}_0} = 1 - \Pr\left(\hat{\mathbf{d}}_0 = \mathbf{d}_0\right)|_{\mathbf{H}_0,\mathbf{d}_0}. \quad (5.22)$$

Without loss of generality, we assume that $\mathbf{d}_0 = [1, 1, \ldots, 1] \triangleq \mathbf{1}$. Therefore,

$$P_C|_{\mathbf{H}_0,\mathbf{1}} = \Pr\left(\sum_{v=0}^{\mathcal{K}-2} \Re\left\{r_v r_{\acute{v}}\right\} = \max_{\tilde{\mathbf{d}}_0}\left\{\sum_{v=0}^{\mathcal{K}-2} \frac{\Re\left\{r_v r_{\acute{v}}\right\}}{\tilde{d}_v \tilde{d}_{\acute{v}}}\right\}\right). \quad (5.23)$$

Since $\mathbf{d}_0$ has $\mathcal{K}-1$ data symbols, then there are $2^{\mathcal{K}-1}$ trial sequences, $\tilde{\mathbf{d}}_0^{(0)}$, $\tilde{\mathbf{d}}_0^{(1)}$,...,

$\tilde{\mathbf{d}}_0^{(\psi)}$, where $\psi = 2^{\mathcal{K}-1}-1$, and $\tilde{\mathbf{d}}_0^{(\psi)}=[1, 1,\ldots, 1]$. The first symbol in every sequence is

set to 1, which is the pilot symbol. By defining $\sum_{v=0}^{\mathcal{K}-2} \frac{\Re\{r_v r_{\acute{v}}\}}{\tilde{d}_v \tilde{d}_{\acute{v}}} \triangleq A_n$, where $\tilde{d}_v \tilde{d}_{\acute{v}} \in \tilde{\mathbf{d}}_0^{(n)}$,

then (5.23) can be written as,

$$P_C|_{\mathbf{H}_0,\mathbf{1}} = \Pr\left(A_\psi > A_{\psi-1}, A_{\psi-2}, \ldots, A_0\right) \tag{5.24}$$

which, as depicted in Appendix B, can be simplified to

$$P_C|_{\mathbf{H}_0,\mathbf{1}} = \prod_{v=0}^{\mathcal{K}-2} \Pr\left(\Re\{r_v r_{\acute{v}}\} > 0\right). \tag{5.25}$$

To evaluate $P_C|_{\mathbf{H}_0,\mathbf{1}}$ given in (5.25), it is necessary to compute $\Pr\left(\Re\{r_v r_{\acute{v}}\} > 0\right)$,

which can be written as

$$\Pr\left(\Re\{r_v r_{\acute{v}}\} > 0\right) = \Pr\left(\underbrace{r_v^I r_{\acute{v}}^I - r_v^Q r_{\acute{v}}^Q}_{r_{v,\acute{v}}^{\text{SP}}} > 0\right). \tag{5.26}$$

Given that $\mathbf{d}_0=[1, 1,\ldots, 1]$, then $r_v^I = \Re\{r_v\} = H_v^I + w_v^I$ and $r_v^Q = \Im\{r_v\} = H_v^Q + w_v^Q$. Therefore, r_v^I, r_v^Q, $r_{\acute{v}}^I$ and $r_{\acute{v}}^Q$ are independent conditionally Gaussian random variables with averages H_v^I, H_v^Q, $H_{\acute{v}}^I$ and $H_{\acute{v}}^Q$, respectively, and the variance for all elements is σ_w^2. To derive the PDF of $r_{v,\acute{v}}^{\text{SP}}$, the PDFs of $r_v^I r_{\acute{v}}^I$ and $r_v^Q r_{\acute{v}}^Q$ should be evaluated, where each of which corresponds to the product of two Gaussian random variables. Although the product of two Gaussian variables is not usually Gaussian, the limit of the moment-generating function of the product has Gaussian distribution. Therefore, the product of two variables $X \sim \mathcal{N}(\mu_x, \sigma_x^2)$ and $Y \sim \mathcal{N}(\mu_y, \sigma_y^2)$ tends to be $\mathcal{N}(\mu_x\mu_y, \mu_x^2\sigma_y^2+\mu_y^2\sigma_x^2)$ as the ratios μ_x/σ_x and μ_y/σ_y increase [39]. By noting that in in (5.26) $\mathrm{E}\left[r_y^x\right] = H_y^x$, $x \in \{I,Q\}$ and $y \in \{v,\acute{v}\}$ and $\sigma_{r_y^x} = \sigma_w$, thus $\mathrm{E}\left[r_y^x\right]/\sigma_{r_y^x} \gg 1$ $\forall\{x,y\}$. Moreover, because the PDF of the sum or difference of two Gaussian random variables is also Gaussian, then, $r_{v,\acute{v}}^{\text{SP}} \sim \mathcal{N}\left(\bar{\mu}_{\text{SP}}, \bar{\sigma}_{\text{SP}}^2\right)$ where $\bar{\mu}_{\text{SP}} = H_v^I H_{\acute{v}}^I + H_v^Q H_{\acute{v}}^Q$ and $\bar{\sigma}_{\text{SP}}^2 = \sigma_w^2\left(|H_v|^2 + |H_{\acute{v}}|^2 + \sigma_w^2\right)$. Consequently,

$$P_C|_{\mathbf{H}_0,\mathbf{1}} = \prod_{v=0}^{\mathcal{K}-2} \Pr\left(r_{v,\acute{v}}^{\text{SP}} > 0\right) = \prod_{v=0}^{\mathcal{K}-2} \left[1 - Q\left(\sqrt{\frac{2\bar{\mu}_{\text{SP}}}{\bar{\sigma}_{\text{SP}}^2}}\right)\right] \tag{5.27}$$

and

$$P_S|_{\mathbf{H}_0,\mathbf{1}} = 1 - \prod_{v=0}^{\mathcal{K}-2} \left[1 - Q\left(\sqrt{\frac{2\bar{\mu}_{\text{SP}}}{\bar{\sigma}_{\text{SP}}^2}}\right)\right] \tag{5.28}$$

where $Q(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{t^2}{2}\right) dt$. Since H_v^I and H_v^Q are independent, then, the condition on $\mathbf{H}_0$ in (5.28) can be removed by averaging P_S over the PDF of $\mathbf{H}_0^I$ and $\mathbf{H}_0^Q$ as,

$$\text{SEP}|_{\mathbf{d}=1} = \underbrace{\int_{-\infty}^\infty \int_{-\infty}^\infty \cdots \int_{-\infty}^\infty}_{2\mathcal{K} \text{ fold}} \text{SEP}|_{\mathbf{H}_0,\mathbf{d}=1} f_{\mathbf{H}_0^I}\left(H_0^I, H_1^I, \ldots, H_{\mathcal{K}-1}^I\right) \times$$
$$f_{\mathbf{H}_0^Q}\left(H_0^Q, H_1^Q, \ldots, H_{\mathcal{K}-1}^Q\right) dH_0^I dH_1^I \ldots dH_{\mathcal{K}-1}^I dH_0^Q dH_1^Q \ldots dH_{\mathcal{K}-1}^Q. \tag{5.29}$$

Because the random variables H_i^I and H_i^Q $\forall i$ in (5.29) are real and Gaussian, their PDFs are multivariate Gaussian distributions [35],

$$f_{\mathbf{X}}(X_0, X_1, \ldots, X_{\mathcal{K}-1}) = \frac{\exp\left(-\frac{1}{2}(\mathbf{X} - \boldsymbol{\mu})^{\text{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu})\right)}{\sqrt{(2\pi)^{\mathcal{K}} |\boldsymbol{\Sigma}|}} \tag{5.30}$$

where $\boldsymbol{\mu}$ is the mean vector, which is defined as,

$$\boldsymbol{\mu} = \text{E}[\mathbf{X}] = [\text{E}[X_1], \text{E}[X_2], \ldots, \text{E}[X_{\mathcal{K}-1}]]^T \tag{5.31}$$

and $\boldsymbol{\Sigma}$ is the covariance matrix, $\boldsymbol{\Sigma} = \text{E}\left[(\mathbf{X} - \mu)(\mathbf{X} - \mu)^T\right]$.

Due to the difficulty of evaluating $2\mathcal{K}$ integrals, we consider the special case of flat fading, which implies that $H_v = H_{\acute{v}} \triangleq H$ and $\left(H^I\right)^2 + \left(H^Q\right)^2 \triangleq \alpha^2$, where α is the channel fading envelope, $\alpha = |H|$. Therefore, the SEP expression in (5.28) becomes,

$$P_S|_{\alpha,\mathbf{1}} = 1 - \left[1 - Q\left(\sqrt{\frac{\alpha^2}{\sigma_w^2(\alpha^2 + \sigma_w^2)}}\right)\right]^{\mathcal{K}-1}. \tag{5.32}$$

Recalling the Binomial Theorem, we get

$$(a+b)^n = \sum_{v=0}^{n} \binom{n}{v} a^{n-v} b^v, \quad \binom{n}{v} \triangleq \frac{n!}{(n-v)! v!}. \tag{5.33}$$

Then the SEP formula in (5.32) using the Binomial Theorem in (5.33) can be written as,

$$P_S|_{\alpha,\mathbf{1}} = 1 - \sum_{v=0}^{\mathcal{K}-1} \binom{\mathcal{K}-1}{v} (-1)^v \left[Q\left(\sqrt{\frac{\alpha^2}{\sigma_w^2 (\alpha^2 + \sigma_w^2)}} \right) \right]^v. \tag{5.34}$$

The conditioning on α can be removed by averaging over the PDF of α, which is Rayleigh. Therefore,

$$f(\alpha) = \frac{\alpha}{\sigma_H^2} e^{-\frac{\alpha^2}{2\sigma_H^2}}. \tag{5.35}$$

And hence,

$$P_S|_{\mathbf{1}} = \int_0^{\infty} P_S|_{\alpha,\mathbf{1}} \, f(\alpha) \, d\alpha. \tag{5.36}$$

Because the expression in (5.32) contains high order of Q-function $Q^n(x)$, evaluating the integral analytically becomes intractable for $\mathcal{K} > 2$. For the special case of $\mathcal{K} = 2$, P_S can be evaluated by substituting (5.34) and (5.35) into (5.36) and evaluating the integral yields the following simple expression,

$$P_S|_{\mathbf{1}} = \frac{1}{2(\bar{\gamma}_s + 1)}, \quad \bar{\gamma}_s \triangleq \frac{\mathrm{E}\left[|d_v|^2\right] \mathrm{E}\left[|H|^2\right]}{2\sigma_w^2} \tag{5.37}$$

where $\bar{\gamma}_s$ is the average signal-to-noise ratio (SNR). Moreover, because all data sequences have an equal probability of error, then $P_S|_{\mathbf{1}} = P_S$, which also equivalent to the bit error rate (BER). It is interesting to note that (5.37) is similar to the BER of the differential binary phase shift keying (DBPSK) [35]. However, the two techniques are essentially different as D^3 does not require differential encoding, has no constraints on the shape of the signal constellation, and performs well even in

frequency-selective fading channels.

To evaluate P_S for $\mathcal{K} > 2$, we use an approximation for $Q(x)$ in [40], which is given by

$$Q(x) \approx \frac{1}{\sqrt{2\pi(x^2+1)}} e^{-\frac{1}{2}x^2}, \; x \in [0, \infty). \tag{5.38}$$

Therefore, by substituting (5.38) into the conditional SEP (5.34) and averaging over the Rayleigh PDF (5.35), the evaluation of the SEP becomes straightforward. For example, evaluating the integral for $\mathcal{K} = 3$ gives,

$$P_S|_\mathbf{1} = \frac{\zeta_1}{\pi} \operatorname{Ei}(1, \zeta_1+1) e^{\zeta_1+1}, \quad \zeta_1 \triangleq \frac{1}{2\bar{\gamma}_s}\left(\frac{1}{\bar{\gamma}_s}+1\right) \tag{5.39}$$

where $\operatorname{Ei}(x)$ is the exponential integral (EI), $\operatorname{Ei}(x) \triangleq -\int_{-x}^{\infty} \frac{e^{-t}}{t} dt$. Similarly, P_S for $\mathcal{K} = 7$ can be evaluated to,

$$P_S|_\mathbf{1} = \frac{\zeta_2}{64\pi^3} \left[e^{\zeta+3}(2\zeta_2+6)^2 \operatorname{Ei}(1, \zeta_2+3) - 4(\zeta_2+1)\right], \quad \zeta_2 \triangleq \frac{1}{2\bar{\gamma}_s}\left(\frac{1}{4\bar{\gamma}_s}+1\right). \tag{5.40}$$

Although the SEP is a very useful indicator for the system error probability performance, the BER is actually more informative. For a sequence that contains $\mathcal{K}_D$ information bits, the BER can be expressed as $P_B = \frac{1}{\Lambda}P_S$, where Λ denotes the average number of bit errors given a sequence error, which can be defined as

$$\Lambda = \sum_{m=1}^{\mathcal{K}_D} m\Pr(m). \tag{5.41}$$

Because the SEP is independent of the transmitted data sequence, then, without loss of generality, we assume that the transmitted data sequence is $\mathbf{d}_0^{(0)}$. Therefore,

$$\Lambda = \sum_{m=1}^{\mathcal{K}_D} m\Pr\left(\left\|\hat{\mathbf{d}}_0\right\|^2 = m\right) \tag{5.42}$$

where $\left\|\hat{\mathbf{d}}_0\right\|^2$, in this case, corresponds to the Hamming weight of the detected

sequence $\hat{\mathbf{d}}_0$, which can be expressed as

$$\Pr\left(\left\|\hat{\mathbf{d}}_0\right\|^2 = m\right) = \Pr\left(\mathbf{d}_0^{(0)} \to \bigcup_i \mathbf{d}_0^{(i)}\right), \ \left\|\mathbf{d}_0^{(i)}\right\|^2 = m \tag{5.43}$$

where $\mathbf{d}_0^{(0)} \to \mathbf{d}_0^{(i)}$ denotes the pairwise error probability (PEP). By noting that $\Pr\left(\mathbf{d}_0^{(0)} \to \mathbf{d}_0^{(i)}\right) \neq \Pr\left(\mathbf{d}_0^{(0)} \to \mathbf{d}_0^{(j)}\right) \forall i \neq j$, then deriving the PEP for all cases of interest is intractable. As an alternative, a simple approximation is derived.

For a sequence that consists of $\mathcal{K}_D$ information bits, the BER is bounded by

$$\frac{1}{\mathcal{K}_D} P_S \leq P_B \leq P_S. \tag{5.44}$$

In practical systems, the number of bits in the detected sequence is generally not large, which implies that the upper and lower bounds in (5.44) are relatively tight, and hence, the BER can be approximated as the middle point between the two bounds as,

$$P_B \approx \frac{P_S}{0.5\left(1 + \mathcal{K}_D\right)}. \tag{5.45}$$

The analysis of the general $1 \times \mathcal{N}$ SIMO system is a straightforward extension of the single-input single-output (SISO) case. To simplify the analysis, we consider the flat channel case where the conditional SEP can be written as,

$$P_S|_\alpha = 1 - \left[1 - Q\left(\sqrt{\frac{\sum_{i=1}^{\mathcal{N}} \alpha_i^2}{\sigma_w^2 \left(\mathcal{N}\sigma_w^2 + \sum_{i=1}^{\mathcal{N}} \alpha_i^2\right)}}\right)\right]^{\mathcal{K}-1}. \tag{5.46}$$

Given that all the receiving branches are independent, the fading envelopes will have Rayleigh distribution $\alpha_i \sim \mathcal{R}\left(2\sigma_H^2\right) \forall i$, and thus, $\sum_{i=1}^{\mathcal{N}} \alpha_i^2 \triangleq a$ will have Gamma distribution, $a \sim \mathcal{G}\left(\mathcal{N}, 2\sigma_H^2\right)$,

$$f\left(a\right) = \left(2\sigma_H^2\right)^{\mathcal{N}} e^{-2\sigma_H^2 a} \frac{a^{\mathcal{N}-1}}{\Gamma\left(\mathcal{N}\right)}. \tag{5.47}$$

Therefore, the unconditional SEP can be evaluated as,

$$P_S = \int_0^\infty P_S|_\alpha \, f_A(a) \, da. \tag{5.48}$$

For the special case of $\mathcal{N} = 2$, $\mathcal{K} = 2$, P_S can be evaluated as,

$$P_S = \frac{1}{2} + Q\left(\frac{\varkappa}{\sqrt{\bar{\gamma}_s}}\right)\left[2\bar{\gamma}_s\left(\frac{\bar{\gamma}_s}{\sqrt{2}} + 2\right) - e^{\varkappa^2}\right] - \bar{\gamma}_s\frac{\varkappa}{\sqrt{2\pi}} \tag{5.49}$$

where $\varkappa \triangleq \sqrt{2 + \bar{\gamma}_s}$. Computing the closed-form formulas for other values of $\mathcal{N}$ and $\mathcal{K}$ can be evaluated following the same approach used in the SISO case.

5.5.2   Double-Sided Pilot

Embedding more pilots in the detection segment can improve the detector's performance. Consequently, it worth investigating the effect of embedding more pilots in the SEP analysis. More specifically, we consider double-sided segment, $\tilde{d}_0 = 1$, $\tilde{d}_{\mathcal{K}-1} = 1$, as illustrated in Fig. 5.4. In this case, the detector can be expressed as,

$$\hat{\mathbf{d}}_0 = \arg\max_{\tilde{\mathbf{d}}_0} \frac{\Re\{r_0 r_1\}}{\tilde{d}_1} + \frac{\Re\{r_{\mathcal{K}-2} r_{\mathcal{K}-1}\}}{\tilde{d}_{\mathcal{K}-2}} + \sum_{v=1}^{\mathcal{K}-3} \frac{\Re\{r_v r_{\acute{v}}\}}{\tilde{d}_v \tilde{d}_{\acute{v}}}, \quad \mathcal{K} \in \{3, 4, \ldots, N-1\}. \tag{5.50}$$

From the definition in (5.50), the probability of receiving the correct sequence can be derived based on the reduced number of trials as compared to (5.20). Therefore,

$$P_C|_{\mathbf{H}_0,\mathbf{1}} = \Pr\left(\left(\Re\{r_0 r_1\} + \Re\{r_{\mathcal{K}-2} r_{\mathcal{K}-1}\}\right) \cap \right.$$
$$\left. \Re\{r_1 r_2\} \cap \Re\{r_2 r_3\} \cap \cdots \cap \Re\{r_{\mathcal{K}-4} r_{\mathcal{K}-3}\} > 0\right) \tag{5.51}$$

which, similar to the single-sided case, can be written as,

$$P_C|_{\mathbf{H}_0,\mathbf{1}} = \Pr\left(\left[\prod_{v=0}^{\mathcal{K}-3} \Pr\left(\Re\{r_v r_{\acute{v}}\}\right) + \prod_{v=1}^{\mathcal{K}-2} \Pr\left(\Re\{r_v r_{\acute{v}}\}\right)\right] > 0\right). \tag{5.52}$$

Therefore,

$$P_S|_{\mathbf{H}_{0,1}} = 1 - \left[1 - Q\left(\sqrt{\frac{2\sqrt{2}\bar{\mu}_{\mathrm{SP}}}{\bar{\sigma}_{\mathrm{SP}}^2}} \right) \right] \times \prod_{v=1}^{\mathcal{K}-3} \left[1 - Q\left(\sqrt{\frac{2\bar{\mu}_{\mathrm{SP}}}{\bar{\sigma}_{\mathrm{SP}}^2}} \right) \right]. \tag{5.53}$$

For flat fading channels, the SEP expression in (5.53) can be simplified by following the same procedure in Subsection 5.5.1, for the special case of $\mathcal{K} = 3$, the SEP becomes,

$$P_S = \left(\frac{\Upsilon}{2} - \sqrt{2} \right) \frac{1}{\Upsilon}, \quad \Upsilon \triangleq \sqrt{8\bar{\gamma}_s + \sqrt{2}\left(4 + \frac{1}{\bar{\gamma}_s} \right)}. \tag{5.54}$$

For $\mathcal{K} > 3$, the approximation of $Q^n(x)$, as illustrated in Subsection 5.5.1, can be used in (5.53) to average over the PDF in (5.35). For example, the case $\mathcal{K} = 4$ can be evaluated as,

$$P_S = \frac{1}{8\pi\bar{\gamma}_s} (\Omega_1 - 1) e^{\Omega_1} \mathrm{Ei}(1, \Omega_1), \quad \Omega_1 \triangleq 1 + \frac{\sqrt{2}}{4\bar{\gamma}_s}\left(1 + \frac{1}{4\bar{\gamma}_s} \right). \tag{5.55}$$

For $\mathcal{K} = 6$,

$$P_S = \frac{\Omega_1 - 1}{4\pi^2} \left[1 - \left[(\Omega_1 - 1) e^{\Omega_2} + 2 \right] \mathrm{Ei}(1, \Omega_2) \right], \quad \Omega_2 \triangleq 2 + \frac{\sqrt{2}}{\bar{\gamma}_s}\left(8 + \frac{1}{32\bar{\gamma}_s} \right) \tag{5.56}$$

For the double-sided pilot, $P_B = P_S$ for the case of $\mathcal{K} = 3$, while it can be computed using (5.45) for $\mathcal{K} > 3$.

5.6 Complexity Analysis

The computational complexity is evaluated as the total number of primitive operations needed to perform the detection. The operations that will be used are the number of real additions (R_A), real multiplications (R_M), and real divisions (R_D) required to produce the set of detected symbols $\hat{\mathbf{d}}$ for each technique. It worth noting that one complex multiplication (C_M) is equivalent to four R_M and three R_A opera-

tions, while one complex addition (C_A) requires two R_A. To simplify the analysis, we first assume that constant modulus (CM) constellations such as MPSK is used, then, we evaluate the complexity for higher-order modulation such as quadrature amplitude modulation (QAM) modulation.

5.6.1   Complexity of Conventional OFDM Detectors

The complexity of the conventional OFDM receiver that consists of the following main steps with the corresponding computational complexities:

1. Channel estimation of the pilot symbols, which computes $\hat{H}_k$ at all pilot subcarriers. Assuming that the pilot symbol d_k is selected from a CM constellation, then $\hat{H}_k = r_k d_k^*$ and hence, N_P complex multiplications are required. Therefore, $R_A^{(1)} = 4N_P$ and $R_M^{(1)} = 4N_P$.

2. Interpolation, which is used to estimate the channel at the non-pilot subcarriers. The complexity of the interpolation process depends on the interpolation algorithm used. For comparison purposes, we assume that linear interpolation is used, which is the least complex interpolation algorithm. The linear interpolation requires one complex multiplication and two complex additions per interpolated sample. Therefore, the number of complex multiplications required is $N - N_P$ and the number of complex additions is $2(N - N_P)$. And hence, $R_A^{(2)} = 7(N - N_P)$ and $R_M^{(2)} = 4(N - N_P)$.

3. Equalization, a single-tap equalizer requires $N - N_P$ complex division to compute the decision variables $\check{r}_k = \frac{r_k}{\hat{H}_k} = r_k \frac{\hat{H}_k^*}{\left|\hat{H}_k^*\right|^2}$. Therefore, one complex division requires two complex multiplications and one real division. Therefore, $R_A^{(3)} = 6(N - N_P)$, $R_M^{(3)} = 8(N - N_P)$ and $R_D^{(3)} = (N - N_P)$.

4. Detection, assuming symbol-by-symbol minimum distance detection, the detector can be expressed as $\hat{d}_k = \arg\min_{\tilde{d}_i} J\left(\tilde{d}_i\right)$, $\forall i \in \{0, 1, \dots, M - 1\}$ where

$J\left(\tilde{d}_i\right) = \left|\check{r}_k - \tilde{d}_i\right|^2$. Assuming CM modulation is used, expanding the cost function and dropping the constant terms we can write $J\left(\tilde{d}_k\right) = -\check{r}_k\tilde{d}_k^* - \check{r}_k^*\tilde{d}_k$. We can also drop the minus sign from the cost function, and thus, the objective becomes maximizing the cost function $\hat{d}_k = \arg\min_{\tilde{d}_i} J\left(\tilde{d}_i\right)$. Since the two terms are complex conjugate pair, then $-\check{r}_k\tilde{d}_k^* - \check{r}_k^*\tilde{d}_k = 2\Re\left\{\check{r}_k\tilde{d}_k^*\right\}$, and thus we can write the detected symbols as,

$$\hat{d}_k = \arg\max_{\tilde{d}_k}\left(\Re\left\{\check{r}_k\right\}\Re\left\{\tilde{d}_k^*\right\} - \Im\left\{\check{r}_k\right\}\Im\left\{\tilde{d}_k^*\right\}\right) \tag{5.57}$$

Therefore, the number of real multiplications required for each information symbol is $2M$, and the number of additions is M. Therefore, $R_A^{(4)} = (N - N_P)M$ and $R_M^{(4)} = 2(N - N_P)M$.

Finally, the total computational complexity per OFDM symbol can be obtained by adding the complexities of the individual steps $1 \to 4$, as:

$$R_A^{CM} = \sum_{i=1}^{4} R_A^{(i)} = (13 + M)N - (10 + M)N_P \tag{5.58}$$

$$R_M^{CM} = \sum_{i=1}^{4} R_M^{(i)} = 2N(6 + M) - 2N_P(4 + M) \tag{5.59}$$

$$R_D^{CM} = \sum_{i=1}^{4} R_D^{(i)} = N - N_P. \tag{5.60}$$

For higher modulation orders, such as QAM, the complexity of the conventional OFDM receivers considering addition division operations is computed following the same steps $1 \to 4$ above, and found to be as:

$$R_A^{QAM} = \sum_{i=1}^{4} R_A^{(i)} = 6N_P + (13 + 2M)(N - N_P) \tag{5.61}$$

$$R_M^{QAM} = \sum_{i=1}^{4} R_M^{(i)} = 8N_P + (12 + 4M)(N - N_P) \tag{5.62}$$

$$R_D^{QAM} = \sum_{i=1}^{4} R_D^{(i)} = N_P + 2M(N - N_P) \tag{5.63}$$

5.6.2   Complexity of the D^3

The complexity of the D^3 based on the VA is mostly determined by the branch and path metrics calculation. The branch metrics can be computed as

$$J_{m,n}^c = \frac{|r_c|^2}{\left|\tilde{d}_m\right|^2} - \frac{r_c r_{\acute{c}}^*}{\tilde{d}_m \tilde{d}_n^*} - \frac{r_c^* r_{\acute{c}}}{\tilde{d}_m^* \tilde{d}_n} + \frac{|r_c|^2}{\left|\tilde{d}_n\right|^2}. \tag{5.64}$$

For CM constellation, the first and last terms are constants, and hence, can be dropped. Therefore,

$$J_{m,n}^c = -\frac{r_c r_{\acute{c}}^*}{\tilde{d}_m \tilde{d}_n^*} + \frac{r_c^* r_{\acute{c}}}{\tilde{d}_m^* \tilde{d}_n}. \tag{5.65}$$

By noting that the two terms in (5.65) are the complex conjugate pair, then

$$J_{m,n}^c = -2\Re \left\{ \frac{r_c r_{\acute{c}}^*}{\tilde{d}_m \tilde{d}_n^*} \right\}. \tag{5.66}$$

From the expression in (5.66), the constant "−2" can be dropped from the cost function, however, the problem with be flipped to a maximization problem. Therefore, by expanding (5.66), we get,

$$J_{m,n}^c = \Re \left\{ \frac{\Re\{r_c\}\Re\{r_{\acute{c}}^*\} - \Im\{r_c\}\Im\{r_{\acute{c}}^*\} + j\left[-\Re\{r_c\}\Im\{r_{\acute{c}}^*\} + \Im\{r_c\}\Im\{r_{\acute{c}}^*\}\right]}{\Re\{\tilde{d}_m \tilde{d}_n^*\} + j\Im\{\tilde{d}_m \tilde{d}_n^*\}} \right\}. \tag{5.67}$$

By defining $\tilde{d}_m \tilde{d}_n^* \triangleq \tilde{u}_{m,n}$, and using complex numbers identities, we get (5.68),

$$J_{m,n}^c = \frac{[\Re\{r_c\}\Re\{r_{\hat{c}}^*\} + \Im\{r_c\}\Im\{r_{\hat{c}}^*\}]\Re\{\tilde{u}_{m,n}\} - [-\Re\{r_c\}\Im\{r_{\hat{c}}^*\} + \Im\{r_c\}\Im\{r_{\hat{c}}^*\}]\Im\{\tilde{u}_{m,n}\}}{\Re\{\tilde{u}_{m,n}\}^2 + \Im\{\tilde{u}_{m,n}\}^2}.$$
$$(5.68)$$

For CM, $\Re\{\tilde{u}_{m,n}\}^2 + \Im\{\tilde{u}_{m,n}\}^2$ is constant, and hence, it can be dropped from the cost function, which implies that no division operations are required.

To compute $J_{m,n}^c$, it is worth noting that the two terms in brackets are independent of $\{m, n\}$, and hence, they are computed only once for each value of c. Therefore, the complexity at each step in the trellis can be computed as $R_A = 3 \times 2^M$, $R_M = 4 + 2 \times 2^M$ and $R_D = 0$, where 2^M is the number of branches at each step in the trellis. However, if the trellis starts or ends by a pilot, then only M computations are required. By noting that the number of full steps is $N - 2N_P - 1$, and the number of steps that require M computations is $2(N_P - 1)$, then the total computations of the branch metrics (BM) are:

$$R_A^{BM} = \left(3 \times 2^M\right)(N - 2N_P - 1) + 2(3 \times M)(N_P - 1)$$
$$R_M^{BM} = \left(4 + 2^{M+1}\right)(N - 2N_P - 1) + 2(N_P - 1)(4 + 2M)$$
$$R_D^{BM} = 0$$

The path metrics (PM) require $R_A^{PM} = (N - 2N_P - 1) + M(N_P - 1)$ real addition. Therefore, the total complexity is:

$$R_A^{CM} = (N - 2N_P - 1)\left(5 \times 2^M\right) + 7M(N_P - 1) \tag{5.69}$$
$$R_M^{CM} = (N - 2N_P - 1)\left(4 + 2^{M+1}\right) + 2(N_P - 1)(4 + 2M) \tag{5.70}$$
$$R_D^{CM} = 0 \tag{5.71}$$

For QAM modulation, the most general case for the branch metrics of the D^3 will be used as,

Table 5.1: Computational complexity comparison using different values of N, $N_P = N/4$, for BPSK.

N	128	256	512	1024	2048
η_{R_A}	0.58	1.07	1.21	1.27	1.31
η_{R_M}	0.77	0.72	0.68	0.64	0.61
R_D	96	192	384	768	1536
η_P	0.20	0.21	0.22	0.26	0.31

$$J_{m,n}^c = \left| \frac{r_c}{\tilde{d}_m} - \frac{r_{\acute{c}}}{\tilde{d}_n} \right|^2. \tag{5.72}$$

The branch metric in (5.72) requires one complex addition, $C_A = 1$, one complex multiplication, $C_M = 1$, and two complex divisions, $C_D = 2$, per branch metrics. Therefore, the total path metric complexity is:

$$R_A^{QAM} = 5MN_P + 10M\left(N - N_P\right) \tag{5.73}$$
$$R_M^{QAM} = 4MN_P + 8M\left(N - N_P\right) \tag{5.74}$$
$$R_D^{QAM} = 2MN_P + 4M\left(N - N_P\right) \tag{5.75}$$

To compare the complexity of the D^3, we use the conventional detector using LS channel estimation, linear interpolation, zero-forcing (ZF) equalization, and MLD, denoted as coherent-L, as a benchmark due to its low complexity. The relative complexity is denoted by η, which corresponds to the ratio of the D^3 complexity to the conventional detector, i.e., η_{R_A} denotes the ratio of real additions and η_{R_M} corresponds to the ratio of real multiplications. As depicted in Table 5.1, R_A for D^3 less than coherent-L only using BPSK for $N = 128$, and then it becomes larger for all the other considered values of N. For R_M, D^3 is always less than the coherent-L, particularly for high values of N, where it becomes 0.61 for $N = 2048$. It is worth noting that R_D in the table corresponds to the number of divisions in the conventional OFDM since the D^3 does not require any division operations. For a more informative comparison between the two systems, we use the computational power analysis presented in [41],

Table 5.2: Computational complexity comparison using different values of N, $N_P = N/4$, for 16-QAM and 64-QAM.

M	16		64	
N	512	2048	512	2048
η_{R_A}	1.25	1.25	1.64	1.64
η_{R_M}	0.52	0.47	0.64	0.62
η_{R_D}	0.98	0.98	0.99	0.99
η_P	0.94	0.84	0.91	0.80

where the total power for each detector is estimated based on the total number of operations. Table 5.1 shows the relative computational power η_P, which shows that the D^3 detector requires only 0.2 of the power required by the coherent-L detector for $N = 128$ and 0.31% for $N = 2048$.

It is also worth considering the complexity analysis for higher modulation orders that require division operations such as 16-QAM and 64-QAM since they widely used in modern wireless broadband systems [2], [3]. Table 5.2 shows the rations of real multiplications, multiplications, divisions, and lastly the ration of the overall computational power for 16-QAM and 64-QAM considering $N = 512$ and $N = 2048$. Unlike the CM modulus case, the D^3 requires division operations, where it is very comparable to conventional OFDM receivers in terms of the division computational resources. Although, the total number of computational addition resources needed is higher in D^3 by $25\% - 65\%$, Nevertheless, the overall computational resources in D^3 is less than the conventional OFDM reveries by $\%6 - 20\%$ due to the significant saving in the multiplication operations of the D^3.

Besides, it is worth noting that linear interpolation has lower complexity as compared to more accurate interpolation schemes such as the spline interpolation [42], [43], which comes at the expense of the error rate performance. Therefore, the results presented in Table 5.1 can be generally considered as upper bounds on the relative complexity of the D^3, when more accurate interpolation schemes are used, the relative

Table 5.3: Computational complexity comparison using hard and soft VA for different values of K, $N = 2048$.

K	3	4	5	6	7
Soft	0.96	0.97	0.97	0.98	0.99
Hard	0.24	0.26	0.28	0.33	0.41

complexity will drop even further as compared to the results in Table 5.1.

5.6.3   Complexity with Error Correction Coding

To evaluate the impact of the complexity reduction of the D^3 in the presence of FEC coding, convolutional codes are considered with soft and hard decision decoding using the VA. BPSK is the modulation considered for the complexity evaluation and the code rate is assumed to be $1/2$. For decoding of convolutional codes, the soft VA requires $n \times 2^K$ addition or subtractions and multiplications per decoded bit, where $1/n$ is the code rate and K is the constraint length [44]. Therefore, for $1/2$ code rate, $R_A = R_M = 2^{K+1}$. Given that each OFDM symbol has N coded bits and $N/2$ information bits, the complexity per OFDM symbol becomes $R_A = R_M = N \times 2^K$. For the hard VA, $N \times 2^K$ XOR operations are required for the branch metric computation, while $N \times 2^{K-1}$ additions are required for the path metric computations. Because the XOR operation is a bit operation, it's complexity is much less than the addition. Assuming that addition is using an 8-bit representation, then the complexity of an addition operation is about eight times the XOR. Therefore, R_A, in this case, can be approximated as $N\left(2^K + 2^{K-2}\right)$.

As can be noted from Table 5.3, the complexity reduction when soft VA is used less significant as compared to the hard VA. Such a result is obtained because the soft VA requires the CSI to compute the reliability factors, which requires $N - N_P$ division operations when the D^3 is used. For hard decoding, the advantage of the D^3 is significant even for high constraint length values.

5.7 Numerical Results

This section presents the performance of the D^3 detector in terms of BER for several operating scenarios. The system model follows the LTE-A physical layer (PHY) specifications [3], where the adopted OFDM symbol has $N = 512$, $N_{CP} = 64$, the sampling frequency $f_s = 7.68$ MHz, the subcarrier spacing $\Delta f = 15$ kHz, and the pilot grid follows that of Fig. 5.1. The total OFDM symbol period is 75 μ sec, and the CP period is 4.69 μ sec. The channel models used are the flat Rayleigh fading channel, the typical urban (TUx) multipath fading model [45] that consists of 6 taps with normalized delays of $[0, 2, 3, 9, 13, 29]$ and average taps gains are $[0.2, 0.398, 0.2, 0.1, 0.063, 0.039]$, which corresponds to a severe frequency-selective channel. The TUx model is also used to model a moderate frequency-selective channel where the number of taps in the channel is 9 with normalized delays of $[0, 1, \ldots, 8]$ samples, and the average taps gains are $[0.269, 0.174, 0.289, 0.117, 0.023, 0.058, 0.036, 0.026, 0.008]$. The channel taps gains are assumed to be independent and Rayleigh distributed. The Monte Carlo simulation results included in this work are obtained by generating 10^6 OFDM symbols per simulation run. Throughout this section, the ML coherent detector with perfect CSI will be denoted as coherent, while the coherent with linear and spline interpolation will be denoted as coherent-L and coherent-S, respectively. Moreover, the results are presented for the SISO system, $\mathcal{N} = 1$, unless it is mentioned otherwise. The SNR in the obtained results is defined as the ratio of the average received signal power to the average noise power regardless of the number of pilots. Such an approach is followed because the proposed system in this work is evaluated in the context of the LTE RB, which has a fixed structure. For more general comparisons, the power and spectral efficiency of all considered systems should be identical. It is worth noting that the considered channel model is widely used in the literature for 4G, and hence, it was adopted in the work. In addition, the new channel models have been studied for 5G from [46, Table 7.7.1-5. CDL-E., page 66] are similar to the

adopted channel for several cases of interests.

Fig. 5.5 shows the BER of the single-sided (SS) and double-sided (DS) D^3 over flat fading channels for $\mathcal{K} = 2, 6$ and $3, 7$, respectively, and using BPSK. The number of data symbols $\mathcal{K}_D = \mathcal{K} - 1$ for the SS and $\mathcal{K}_D = \mathcal{K} - 2$ for the DS because there are two pilot symbols at both ends of the data segment for the DS case. The results in the figure for the SS show that $\mathcal{K}$ has a noticeable impact on the BER where the difference between the $\mathcal{K} = 2$ and 6 cases is about 1.6 dB at BER of 10^{-3}. For the DS segment, the BER has the same trends of the SS except that it becomes closer to the coherent case because having more pilots reduces the probability of sequence inversion due to the phase ambiguity problem. The figure shows that the approximated and simulation results match very well for all cases, which confirms the accuracy of the derived approximations.

The effect of the frequency selectivity is illustrated in Fig. 5.6 for the SS and DS configurations using $\mathcal{K}_D = 1$. As can be noted from the figure, frequency-selective channels introduce error floors at high SNRs, which is due to the difference between adjacent channel values caused by the channel frequency selectivity. Furthermore, the figure shows a close match between the simulation and the derived approximations. The approximation results are presented only for $\mathcal{K} = 2$ because evaluating the BER for $\mathcal{K} > 2$ becomes computationally prohibitive. For example, evaluating the integral (5.29) for the $\mathcal{K} = 3$ requires solving a 6-fold integral. The results for the frequency-selective channels are quite different from the flat fading cases. In particular, the BER performance drastically changes when the DS pilot segment is used. Moreover, the impact of the frequency selectivity is significant, particularly for the SS pilot case.

Fig. 5.7 shows the BER of the 1×2 SIMO D^3 over flat fading channels for SS and DS pilot segments. It can be noted from the figure that the maximum ratio combiner (MRC) BER with perfect CSI outperforms the DS and SS systems by about 2 and 3 dB, respectively. Moreover, the figure shows that the MLSD [32] and the D^3 have
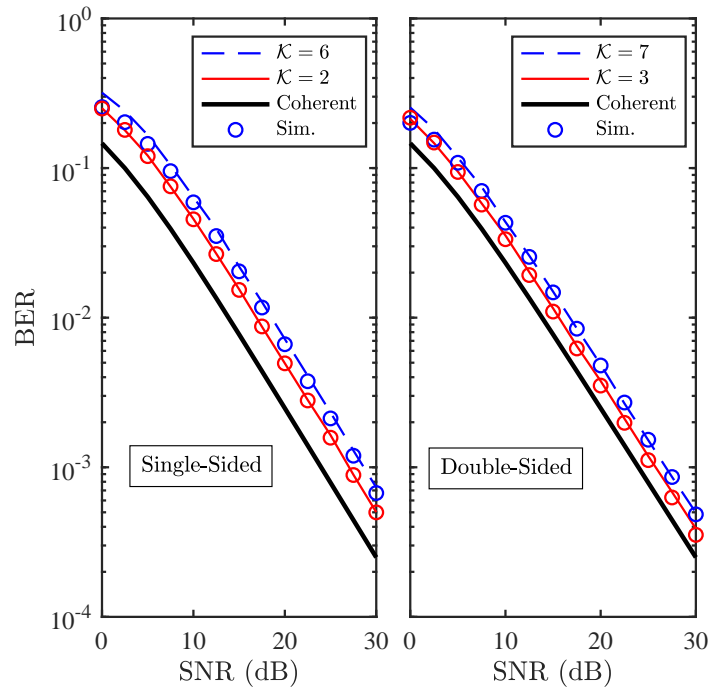
Figure 5.5: BER using SS and DS pilots for different values of $\mathcal{K}$ over flat fading channels using BPSK, $\mathcal{N} = 1$.
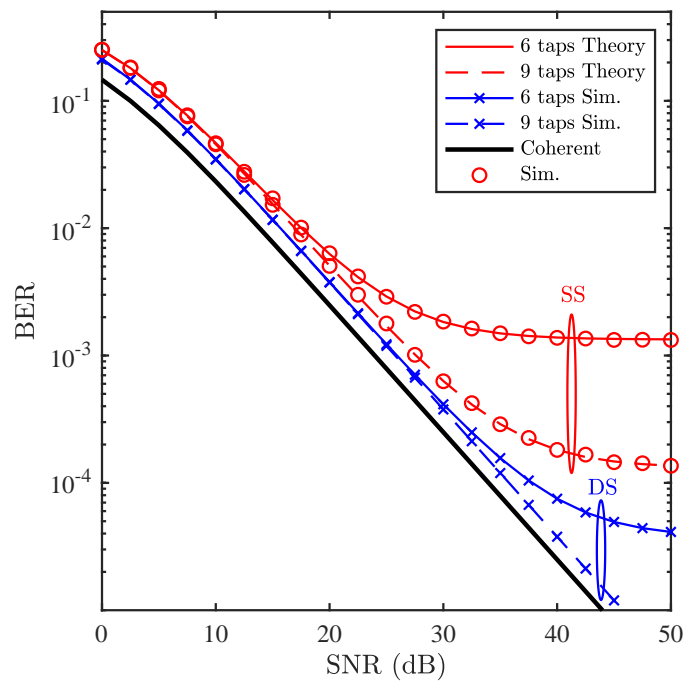
Figure 5.6: BER in frequency-selective channels using BPSK, $\mathcal{K}_D = 1$ and $\mathcal{N} = 1$.
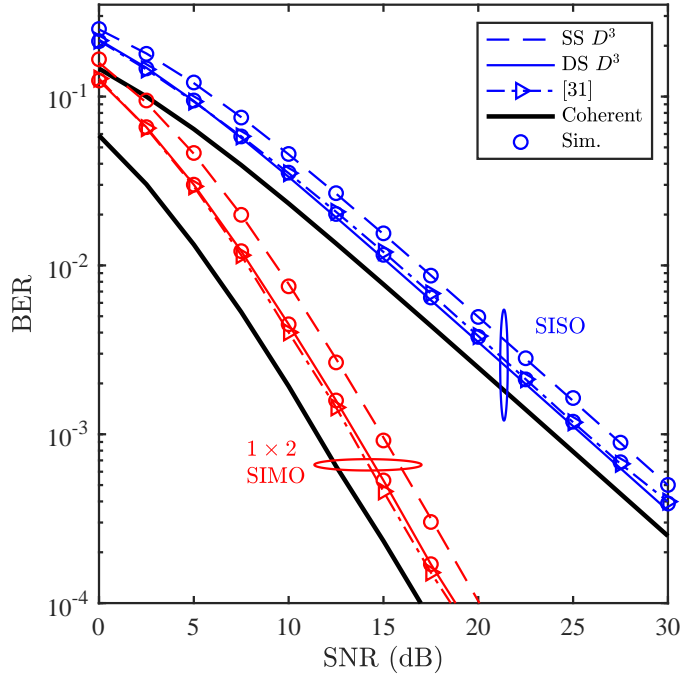
Figure 5.7: BER of D^3 and MLSD [32] SISO and SIMO using SS and DS pilots, flat fading, BPSK, $\mathcal{N} = 1, 2$, and $\mathcal{K}_D = 1$.

equivalent BER for the SISO and SIMO scenarios. The figure also shows the BER of the 1×2 SIMO systems as compared to the SISO case.

Figs. 5.8 shows the BER of the SISO and 1 × 2 SIMO MLSD, coherent, coherent-S and coherent-L systems over frequency-selective channels. For both SISO and SIMO, the BER of all the considered techniques converges at low SNRs because the AWGN dominates the BER in the low SNR range. For moderate and high SNRs, the D^3 outperforms all the other considered techniques except for the coherent, where the difference is about 3.5 and 2.75 dB at BER of 10^{-3} for the SISO and SIMO systems, respectively.

Fig. 5.9 compares the BER of the D^3, PSP [27], MLSD [32], MSDD [30], and the coherent detector over the 6-taps channel using BPSK. As can be noted from the figure, the D^3 noticeably outperforms all other detectors for SNR $\gtrsim$ 15 dB, which indicates that the D^3 is more robust to the frequency selectivity of the channel.
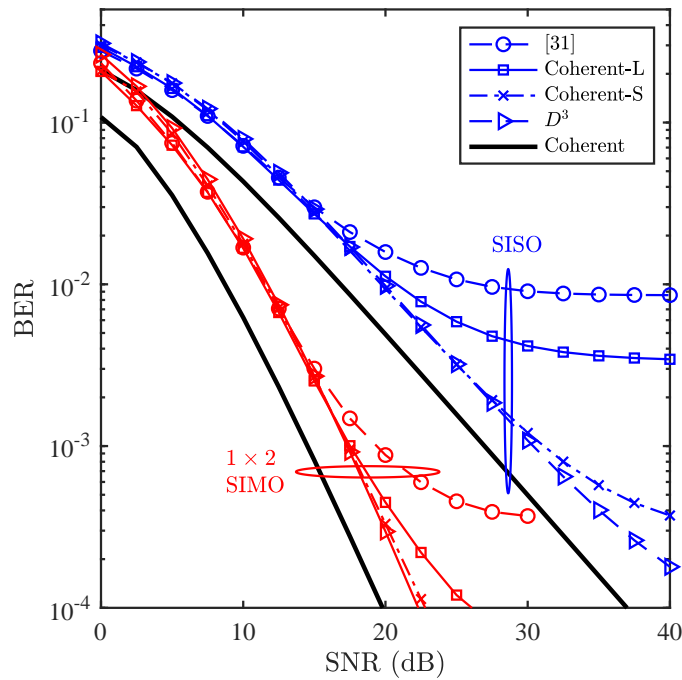
Figure 5.8: BER of the SISO D^3 and MLSD [32] over the 6-taps frequency-selective channel using QPSK, $\mathcal{K}_D = 1$, $\mathcal{N} = 1, 2$.
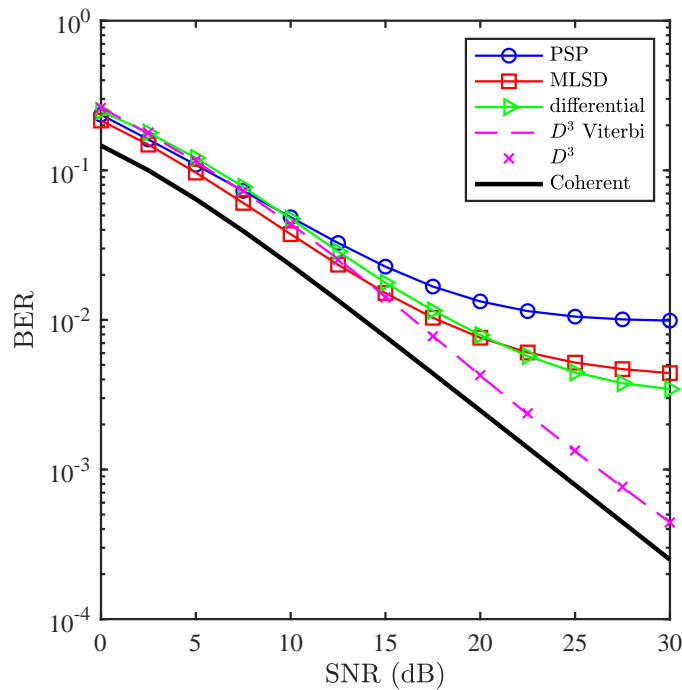
Figure 5.9: BER of the D^3 for $\mathcal{K} = 7$ DS using BPSK compared to several other sequence detectors over 6-taps frequency-selective channel.
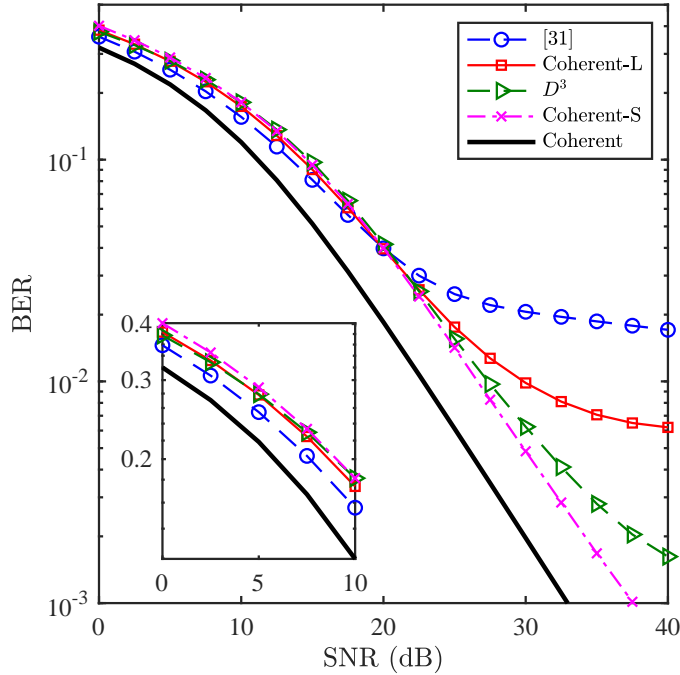
Figure 5.10: BER of the D^3 for $\mathcal{K} = 7$ DS over the 6-taps frequency-selective channel using 16-QAM, compared with MLSD [32].

Moreover, the figure shows the D^3BER using VA which, as expected, is identical to the BER obtained using (5.11). It is worth noting that all the systems considered in the figure are implemented using the DS segment where $\mathcal{K} = 7$, and thus, they are evaluated under similar throughput conditions. However, the BER sensitivity of each technique to the number of pilot symbols could be different from other techniques, which implies that some of these techniques might be able to provide roughly the same BER but using fewer pilot symbols. The same argument applies to the power efficiency as well, because the power allocated per information bit becomes different for various systems. However, because the LTE RB is used as the basis for testing all systems, then the current comparison can be considered generally fair. In the worst case scenario, i.e., considering that all other systems are fully blind, then the throughput power loss is only 4.7% as described in Subsection 5.4.2, which has a negligible effect on the BER.

Fig. 5.10 shows the BER for the D^3, MLSD [32], coherent, coherent-L and coherent-S using 16-QAM. As can be noted from the figure, the MLSD slightly out-performs the D^3 at low SNRs, and the coherent-S outperforms the D^3 at high SNRs. However, the coherent-S has generally much higher complexity.

Fig. 5.11 shows the simulated BER of the D^3 system when it is used to detect a complete RB as described in Subsection 5.4.2. The channel model is similar to the 6-taps used described above, and the channel gain variation over consecutive OFDM symbols is generated using the Jakes's model, where the maximum Doppler frequency $f_d = \frac{V}{c} f_c$, where V is the speed of the vehicle, c is the speed of light, $c = 3 \times 10^8$ m/s, and the carrier $f_c = 1.9$ GHz. The channel is considered quasi-static, i.e., the channel remains constant over the OFDM symbol period, but changes over consecutive symbols. As the figure indicates, the D^3 is more immune to channel mobility at 50 km/h as compared to pilot-based systems as it did not have an error floor. For the high mobility case, $V = 300$ km/h, the D^3 BER exhibited an error floor at about 6×10^{-4}, which is much lower than the error floor of the coherent detector with linear and spline interpolation.

Fig. 5.12 shows the simulated BER of the D^3 using convolutional codes with hard decision decoding, using the widely used (171, 131) convolutional code with a block length of 256 bits, and a 512×512 channel block interleaver. Moreover, the results without interleaver are considered, which corresponds to the case of slow fading channels with very long coherence time. As it can be noted from the figure, the BER of the D^3 and coherent-L are comparable for the considered range of SNR when the block interleaver is used. On the contrary, with no interleaving, the D^3 offers about 5 dB advantage at 10^{-6}. Both detectors are approximately 3 dB away from the coherent detector with perfect CSI. It is wor
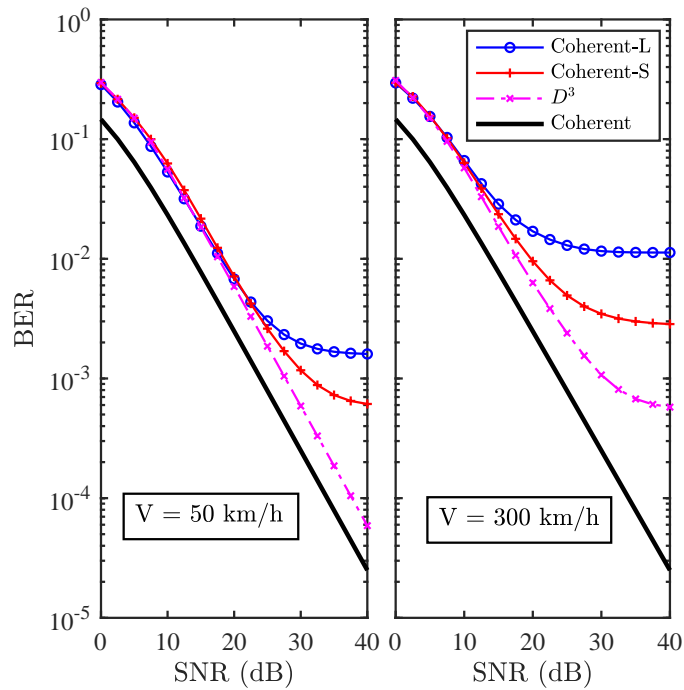
Figure 5.11: BER for the SISO D^3, coherent-L, and coherent detector for a complete LTE RB using the 6-taps channel and BPSK for different mobility values.
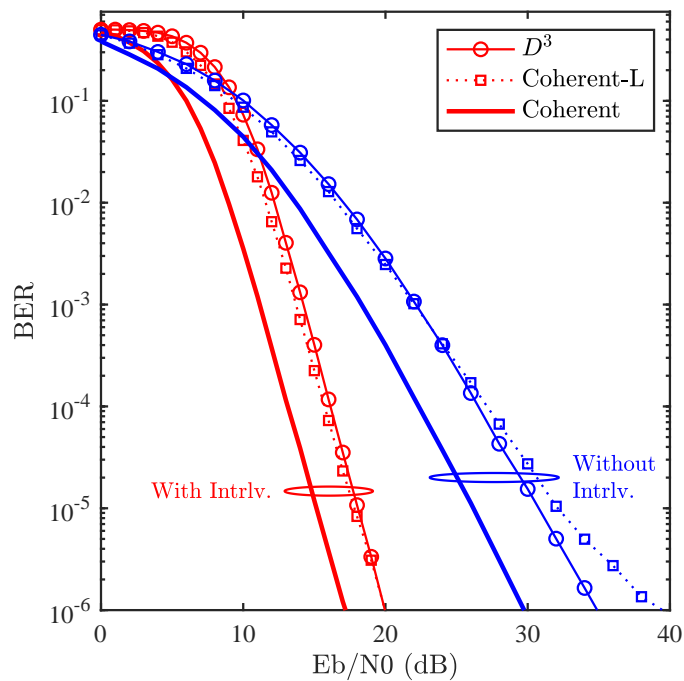
Figure 5.12: Coded BER for the SISO D^3, coherent-L, and coherent detector for $\mathcal{K} = 7$ DS over the 6-taps frequency-selective channel using BPSK.

5.8 Conclusion and Future Work

This work proposed a new receiver design for OFDM-based broadband communication systems. The new receiver performs the detection process directly from the FFT output symbols without the need of experiencing the conventional steps of channel estimation, interpolation, and equalization, which led to a considerable complexity reduction. Moreover, the D^3 system can be deployed efficiently using the VA. The proposed system was analyzed theoretically where simple closed-form expressions were derived for the BER in several cases of interest. The analytical and simulation results show that the D^3 BER outperforms the coherent pilot-based receiver in various channel conditions, particularly in frequency-selective channels where the D^3 demonstrated high robustness.

Although the D^3 may perform well even in severe fading conditions, it is crucial to evaluate its sensitivity to various practical imperfections. Thus, we will consider in our future work the performance of the D^3 in the presence of various system imperfections such as phase noise, synchronization errors and IQ imbalance. Moreover, we will evaluate the D^3 performance in mobile fading channels, where the channel variation may introduce intercarrier interference. In addition, the D^3 has been considered in this work for SIMO systems, it can be also applied to multiple-input multiple-output (MIMO) systems. Nevertheless, the system design and performance analysis require a dedicated work, and hence, the MIMO case will be considered in our future work.

Bibliography

[1] A. Saci, A. Al-Dweik and A. Shami, "Direct Data Detection of OFDM Signals Over Wireless Channels," Dec. 2019, Available: arXiv, https://arxiv.org/submit/2958501.

[2] IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Broadband Wireless Access Systems Amendment 3: Advanced Air Interface, IEEE Std. 802.16m, 2011.

[3] LTE; Evolved Universal Terrestrial Radio Access (E-ULTRA), LTE physical layer, 3GPP TS 36.300, 2011.

[4] T. Hwang, C. Yang, G. Wu, S. Li, and G. Y. Li, "OFDM and its wireless applications: A survey," IEEE Trans. Veh. Technol., vol. 58, no. 4, pp. 1673–1694, May 2009.

[5] D. Tsonev, *et al.*, "A 3-Gb/s single-LED OFDM-based wireless VLC link using a gallium nitride μLED," *IEEE Photon. Technol. Lett.*, vol. 26, no. 7, pp. 637-40, Apr. 2014.

[6] S. Dissanayake, J. Armstrong, "Comparison of ACO-OFDM, DCO-OFDM and ADO-OFDM in IM/DD systems," *J. Lightw. Technol.*, vol. 31, no. 7, pp. 1063-72, Apr. 2013

[7] P. Guan et al., "5G field trials: OFDM-based waveforms and mixed numerologies," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1234-1243, June 2017.

[8] M. Agiwal, A. Roy and N. Saxena, "Next generation 5G wireless networks: a comprehensive survey," IEEE Commun. Surveys & Tutorials, vol. 18, no. 3, pp. 1617-1655, thirdquarter 2016.

[9] Weile Zhang, Qinye Yin, Wenjie Wang, and Feifei Gao, "One-shot blind CFO and channel estimation for OFDM with multi-antenna receiver," *IEEE Trans. Signal Process.*, vol. 62, no. 15, pp. 3799-3808, Aug. 2014.

[10] Song Noh, Youngchul Sung, Michael Zoltowski, "A new precoder design for blind channel estimation in MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 12, pp. 7011-7024, Dec. 2014.

[11] A. Saci, A. Al-Dweik, A. Shami, and Y. Iraqi, "One-shot blind channel estimation for OFDM systems over frequency-selective fading channels," *IEEE Trans. Commun.*, vol. 65, no. 12, pp. 5445-5458, Dec. 2017.

[12] A. Saci, A. Al-Dweik and A. Shami, "Blind channel estimation using cooperative subcarriers for OFDM systems," *IEEE Int. Conf. Commun.* (ICC), Kansas City, USA, May 2018.

[13] X. Zhang and D. Xu, "Blind channel estimation for multiple antenna OFDM system subject to unknown carrier frequency offset," *J. of Sys. Eng. and Electron.*, vol. 25, no. 5, pp. 721-727, Oct. 2014.

[14] A. Mezghani and A. L. Swindlehurst, "Blind estimation of sparse broadband massive MIMO channels with ideal and one-bit ADCs," *IEEE Trans. Signal Process.*, vol. 66, no. 11, pp. 2972-2983, June 2018.

[15] Hongting Zhang and Hsiao-Chun Wu, "Robust pilot detection techniques for channel estimation and symbol detection in OFDM systems," *IEEE Signal Process. Lett.*, vol. 22, no. 6, pp. 733-737, June 2015.

[16] R. Shaked, N. Shlezinger and R. Dabora, "Joint estimation of carrier frequency offset and channel impulse response for linear periodic channels," *IEEE Trans. Commun.*, vol. 66, no. 1, pp. 302-319, Jan. 2018.

[17] Y. Wang, G. Liu, F. Han and H. Qu, "Channel estimation and equalization for SS-OOFDM system with high mobility," *IEEE Wireless Commun. Lett.*, vol. 23, no. 1, pp. 92-95, Jan. 2019.

[18] Chenhao Qi, Guosen Yue, Lenan Wu, and A. Nallanathan, "Pilot design for sparse channel estimation in OFDM-based cognitive radio systems," *IEEE Trans. on Veh. Technol.*, vol. 63, no. 2, pp. 982-987, Feb. 2014.

[19] G. Liu, L. Zeng, H. Li, L. Xu, and Z. Wang, "Adaptive complex interpolator for channel estimation in pilot-aided OFDM system," *J. Commun. Networks*, vol. 15, no. 5, pp. 496-503, Oct. 2013.

[20] Jung-Chieh Chen, Chao-Kai Wen, and Pangan Ting, "An efficient pilot design scheme for sparse channel estimation in OFDM systems," *IEEE Commun. Lett.*, vol. 17, no. 7, pp. 1352-1355, July 2013.

[21] T. Lee, D. Sim, B. Seo and C. Lee, "Channel estimation scheme in oversampled frequency domain for FBMC-QAM systems based on prototype filter set," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 728-739, Jan. 2019.

[22] P. Tan and N. Beaulieu, "Effect of channel estimation error on bit error probability in OFDM systems over Rayleigh and Ricean fading channels," *IEEE Trans. Commun.*, vol. 56, no. 4, pp. 675-685., Apr. 2008.

[23] S. Tomasin and M. Butussi, "Analysis of interpolated channel estimation for mobile OFDM systems," *IEEE Trans. Commun.*, vol. 58, no. 5, pp. 1578-1588, May 2010.

[24] P. Hoeher, S. Kaiser and P. Robertson, "Two-dimensional pilot-symbol-aided channel estimation by Wiener filtering," *In Proc IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, vol. 3, Munich, 1997, pp. 1845-1848.

[25] F. D'Agostini, S. Carboni, M. De Castro, F. De Castro, and D. Trindade, "Adaptive concurrent equalization applied to multicarrier OFDM systems," *IEEE Trans. Broadcast*, vol. 54, no. 3, pp. 441-447, Sep. 2008.

[26] M. Henkel, C. Schilling and W. Schroer, "Comparison of channel estimation methods for pilot aided OFDM systems," in *Proc. IEEE VTC. Spring*, Dublin, 2007, pp. 1435-1439.

[27] R. Raheli, A. Polydoros and C-K Tzou, "Per-survivor processing: a general approach to MLSE in uncertain environments," *IEEE Trans. Commun.*, vol. 43, no. 2, pp. 354-364, Feb. 1995.

[28] Z. Zhu and H. Sadjadpour, "An adaptive per-survivor processing algorithm," *IEEE Trans. Commun.*, vol. 50, no. 11, pp. 1716-1718, Nov. 2002.

[29] M. Luise, R. Reggiannini and G. M. Vitetta, "Blind equalization/detection for OFDM signals over frequency-selective channels," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1568-1578, Oct. 1998.

[30] D. Divsalar and M. K. Simon, "Multiple-symbol differential detection of MPSK," *IEEE Trans. Commun.*, vol. 38, no. 3, pp. 300-308, Mar. 1990.

[31] L. Zhang, Z. Hong, Y. Wu, R. Boudreau and L. Thibault, "A novel differential detection for differential OFDM systems with high mobility," *IEEE Trans. Broadcast.*, vol. 62, no. 2, pp. 398-408, June 2016.

[32] M. Wu and P. Y. Kam, "Performance analysis and computational complexity comparison of sequence detection receivers with no explicit channel estimation," *IEEE Trans. on Veh. Technol.*, vol. 59, no. 5, pp. 2625-2631, Jun 2010.

[33] D. W. Ricker and A. J. Cutezo, "A model-based estimator-correlator (EC) structure," IEEE Trans. Signal Process., vol. 48, no. 10, pp. 2733-2742, Oct. 2000.

[34] M. Matinmikko and A. Mammela, "Estimator-correlator receiver in fading channels for signals with pilot symbols," in *Proc. IEEE 15th Ann. Int. Symp. Pers. Indoor Mobial Radio Commun. (PIMRC)*, Barcelona, 2004, pp. 2278-2282 vol. 3.

[35] J. Proakis and M. Salehi, *Digital communications*, 5th ed. New York: McGraw-Hill, 2008.

[36] IEEE Standard for Information technology Telecommunications and information exchange between systems local and metropolitan area networks, specific requirements, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, Amendment 4: Enhancements for Very High Throughput for Operation in Bands below 6 GHz, Dec. 2013.

[37] W. C. Jakes, *Microwave Mobile Communications*, 2nd ed. Wiley, 1994.

[38] S. Lin, T. Kasami, T. Fujiwara, and M. Fossorier, "Trellises and trellis-based decoding algorithms for linear block codes", *Springer Science & Business Media*, vol. 443. 2012.

[39] Seijas-Macias, Antonio, and A. Oliveira. "An approach to distribution of the product of two normal variables," *Discussiones Mathematicae Probability and Statistics*, vol. 32, no. 1-2, pp. 87-99, 2012.

[40] Borjesson, P., and C-E. Sundberg. "Simple approximations of the error function $Q(x)$ for communications applications," *IEEE Trans. Commun.*, vol. 27, no. 3, pp. 639-643, March 1979.

[41] M. Tariq, A. Al-Dweik, B. Mohammad, H. Saleh and T. Stouraitis, "Computational power analysis of wireless communications systems using operation-level power measurements," in *Proc. ICECTA*, Ras Al Khaimah, 2017, pp. 1-6.

[42] D. Petrinovic, "Causal cubic splines: formulations, interpolation properties and implementations," *IEEE Trans. Signal Process.*, vol. 56, no. 11, pp. 5442-5453, Nov. 2008.

[43] D. Lamb, L. F. O. Chamon and V. H. Nascimento, "Efficient filtering structure for spline interpolation and decimation," *IET Electron. Lett.*, vol. 52, no. 1, pp. 39-41, Aug. 1 2016.

[44] P. Wu, "On the complexity of turbo decoding algorithms," *IEEE VTS 53rd Veh. Technol. Conf.*, Spring 2001, Rhodes, Greece, 2001, pp. 1439-1443.

[45] ETSI TR 125 943 V9.0.0 (2010-02), Universal Mobile Telecommunications System (UMTS) Deployment Aspects, 3GPP TR 25.943, Release 9.

[46] 3GPP TR 38.901, 2017. Study on channel model for frequencies from 0.5 to 100 GHz.

Chapter 6

Autocorrelation Integrated Gaussian Based Anomaly Detection using Vibration Sensory Data

6.1 Preamble

[1]Over the past decade, the deployment of Internet of Things (IoT) devices and connected sensors have experienced an upsurge due to the vast number of applications that rely on these technologies. Most IoT devices generate live data streams that can be used for detecting particular phenomena or isolated systems through indirect inference. Recent reports predicts that the number of connected devices will exceed 75 billion by 2025 [1]. To accommodate to such massive number of devices, IoT and cloud computing are key technologies that enable efficient monitoring, scalability and maintenance of the network [2]. With the deployment of the fifth-generation (5G)

---

[1]A version of this chapter has been submitted to the IEEE Sensors Journal.

wireless networks, industrial 4.0 technologies, which is empowered by the industrial IoT (IIoT) have enabled a range of new application such as edge computing [3, 4], smart grids [5], and smart manufacturing [6]. IIoT allows collecting a large amount of data gathered by various sensors to monitor the progress of the manufacturing process, and predict its evolution [7].

The staggering amount of data live streams generated by a massive number of IoT devices makes it difficult for a human-based monitor to be used. This has facilitated a continuous shift towards the use of artificial intelligence (AI) based technologies such as machine learning, deep learning, and even natural language processing to create better intelligent solutions capable of addressing the critical engineering problems currently facing the industrial sector [8, 9]. Another aspect that assisted in the shift is the ability to offload the computational capabilities to the cloud [10] making adoption and implementation of developed solutions quick and efficient. One of the pivotal machine learning applications in the industrial field is anomaly detection, defined as the process of identifying odd or rare events, or observations that flag suspicious behavior that does not confront with the nature of the typical data trends [11, 12]. The term anomalous data is typically associated with alarms or a flag for popular problems such as credit card fraud, health issues, server crashes, etc. Therefore, anomaly detection can be considered as an essential diagnosing tool for assembly lines of industrial processes.

In the context of steel manufacturing applications, steelmaking furnaces convert molten iron into steel via decarburization using high-pressure oxygen [13]. The injection of oxygen is a controlled process that generates multiple chemical reactions in the liquid steel bath, temporarily increasing its energy and volume. This allows for multiple potentially harmful events to occur, especially when the volume of the liquid bath reaches or exceeds the volume of the resulting in a material overflow. Such undesired events in steelmaking plants are quite costly and more impactful is

the disruption it causes to the flow of the factory while the harmful and wasted raw products are removed, and the operations are being reset back to normal. However, the complex nature of the chemical reactions and the high variability of the raw material leads to increasing the unpredictability of the conversion process [14]. The adverse effects of material overflow are significant, and they include losses in yield due to iron oxides droplets and pellets exiting in the furnace, damage to fume hoods, and even the furnace itself, build-ups of solid ferrous material on furnace structures [15]. Apart from the direct economic costs, molten material spill-out leads to higher health and safety costs, and in most cases to environmental fines.

While the mechanism that controls such events is understood, existing preventative countermeasures have resulted in less than ideal outcomes. However, anomaly detection algorithms can be considered as the most effective methods because it allows for high unimpeded productivity while providing timely alerts when the process is deviating from its normal evolution. Towards this goal, several vibration sensors are deployed around the steel furnace main body and remain active to monitor the melting process and identify the occurrences of anomalous events.

6.1.1   Related Work

Anomaly detection has gained tremendous research in diversified application such as fraud detection [16], intrusion detection [17], and predictive fault tolerance [18]. Anomaly detection algorithms are typically classified based on the input format into either point or contextual [19]. Point anomalies occur on a spot check based sampling within the data sample. On the other hand, contextual anomaly detection aims to detect anomalies based on the entire or a window of the data samples. In both types, several machine learning algorithms have been proposed in recent literature to develop anomaly detection algorithms [19], [20]. Machine learning techniques, such as support vector machine (SVM), k-nearest neighbors (k-NN) algorithm, and neural networks

(NN), perform the anomaly detection task by relying on previous experimental data that are usually labeled as being normal "0" or anomalous "1" labeled [20]. However, since anomalous events occur less frequently than their normal counterparts, the trained machine learning algorithm results are typically poor-performing capability in predicting the anomaly class due to the high imbalance in training data [20].

A different family of anomaly detection algorithms are based on reconstructing the input data, which is referred to as autoencoders [21, 22, 23]. Autoencoders are built by training a NN on the normal data and then reconstruct it. In other words, the input data will be used as both an input and an output to the NN, and hence, depending on the reconstruction error of the input, it can be used efficiently in several anomaly detection applications. Nevertheless, autoencoders are not able to identify the temporal correlation of input data. Other methods are proposed in literature such as recurrent NN (RNN) [24], and long short term memory (LSTM) networks [25], which are capable of learning from sequence and time-series data. Although they can be effective in detecting anomalies, their architectures are computationally complex, and cannot be a feasible anomaly detection solution for IoT sensor devices with limited computational and memory resources.

Alternatively, anomaly detection algorithms that are based on the principal component analysis (PCA) have been considered in the literature [26], [27]. The main concept of PCA-based algorithms is based on singular value decomposition (SVD) process, which is widely used in many engineering applications such as control systems, signal processing and wireless systems applications for dimensionality reduction [26]. The PCA can be configured as a detector by decomposing the weighted energy collected by the first component and its reconstruction error. The architecture in [26] is deployed on a cloud cluster to achieve for the system monitoring. Also, it is worth mentioning that the work in [27] finds the optimal subset of features using the genetic algorithm (GA) prior to deploy the PCA model in order to reduce the computational

complexity. Nevertheless, the PCA-based models are detecting anomalies based on the sample level, and consequently, not suitable for contextual time-series data.

Another class of anomaly detection algorithms are based on finding the distribution models for normal samples, and hence, detect anomalies that do not fit the models [28]. This class of algorithms is commonly known in the literature as density-based algorithms. If labels are given in the dataset, the probability density function (PDF) is built using the normal samples, and then consider the scattered samples of the PDF as anomalies. The normal process can be modeled as a Gaussian PDF [29], Poisson distribution [30], or hidden Markov chain [31]. Density-based methods are effective in detecting anomalies in low dimensional datasets [28], but fail to capture the time variation between samples, which are not effective enough in high-dimensionality datasets [28].

6.1.2   Contributions

In this chapter, a novel low-complexity anomaly detection algorithm is designed by exploiting the temporal signals collected from multiple sensors. The algorithm is based on modeling the healthy processes as a Gaussian distribution, and hence, the anomalous processes can be distinguished based on a threshold set for the Gaussian PDF. Also, it captures the temporal data obtained by creating new autocorrelation features. The main contributions of this chapter can be summarized as follows:

1. Design of a low-complexity and accurate automated anomaly detection algorithm. Based on the joint Gaussian modeling of the sensor vibration signals from the historical data.

2. The proposed algorithm can detect temporal fluctuations of the vibration signals by constructing new features that carry autocorrelation information. Therefore, the algorithm is denoted as autocorrelation integrated gaussian (AIG)

algorithm.

3. Propose an objective and informative metrics for anomaly detection in a time-series data that reflects a better insight into the overall performance of the anomaly detection of temporal data.

4. Formulate and solve an optimization framework for the hyper-parameter tuning of the proposed algorithm. The obtained results show a reliable and timely estimation of the anomalous process against the true anomaly labels while keeping the processing time complexity both in training and execution phases substantially low.

6.2 Sensor Data Description and Feature Engineering

The dataset used for developing and testing the model presented in this work is a proprietary and confidential dataset obtained from a privately owned steelmaking plant. The dataset is comprised of the captured measurements of several identical IoT-based vibration sensors, each of which operates at a particular bandwidth inside a steel furnace process during the steelmaking process.

6.2.1 Sensor Data Description

The individual sensors' readings captured before being considered as input ready are filtered to reduce the measurement noise. Two reference vibration signals are considered for the analysis stage, here referred as long-range (LR) and short-range (SR) signals, where LR signals are averaged over a shorter period than the SR signals. Moreover, LR and SR are measured over 4 frequency bands, and hence, it is referred to $HR_i, LR_i, i \in [1, \ldots, 4]$. More details about the characteristics of the signals are
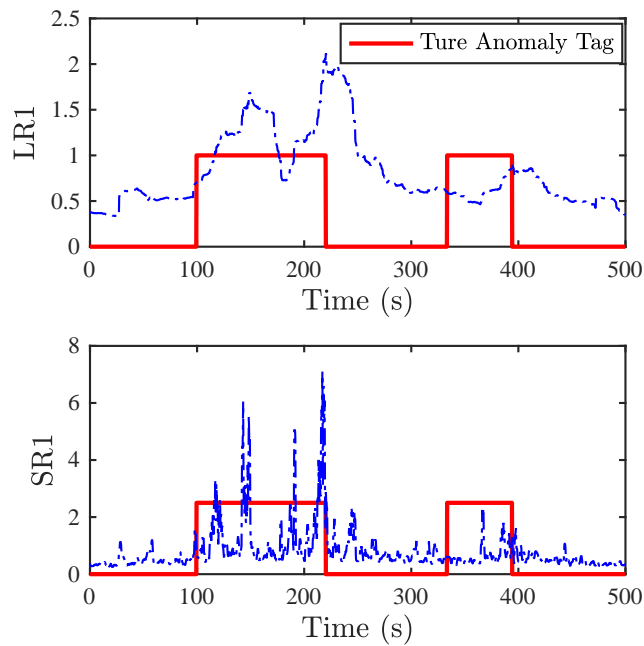
Figure 6.1: Example of vibration sensor measurements.

the proprietary information of the company and therefore, cannot be disclosed in this article. The IoT sensor measurements are time-dependable as results of the variation during the process, and then, the measurements are recorded periodically with the sampling frequency f_s. Finally, the true anomaly tags are given in the datasets. Fig. 6.1 shows the most two anomaly distinctive sensor measurements indicating all anomalous regions as it will be described in Section 6.2.2. As can be noted from Fig. 6.1, the LR signals exhibit a clearly smoother copy of the HR readings, and both can be seen as a filtered version of the actual vibration signals.

6.2.2   Feature Importance Ranking

Several importance feature ranking techniques can be invoked to rank the importance of each vibration measurement with respect to their relation to the occurrence of anomalies. As a design choice, we chose to implement Lasso regression as it is well known to be a robust technique for ranking and sorting the significance of features

against a specific target variable, which suits our current objective, the anomaly flag quite readily. Lasso regression adds an L1 penalty term to the conventional cost function as [32],

$$J(\theta) = \sum_{i=1}^{m} \left(y_i - \sum_j x_{ij}\theta_j \right)^2 + \lambda \sum_j |\theta_j| \qquad (6.1)$$

where m is the number of examples in the dataset, λ is the penalty tuning parameter, and θ_j is the weight of feature j. After applying the method of Gradient Descent (GD) to the cost function in (6.1), the least significant features will be nulled first during the run of the GD algorithm. For instance, Fig. 6.2 shows that SR3, LR2 are the least important ones, while LR1 and SR1 are leading the most important features for determining the anomaly detection process. Consequently, LR1 and SR1 will be used as main signals throughout this work to reduce the feature space complexity and the noise from other measurements.
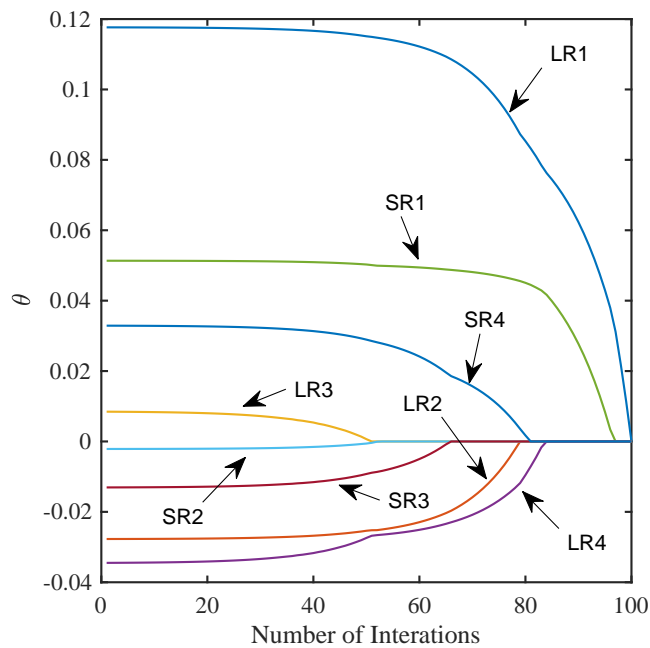
Figure 6.2: Feature importance ranking using Lasso regression.

145

6.2.3   Data Preprocessing

To accurately model the process using Gaussian distribution, the data collected from the vibration sensors should follow the normal distribution. Therefore, the Box-Cox transformation is used to transform non-Gaussian data into a Gaussian distribution. The transformation can be implemented as [33]:

$$\breve{x}_i = \begin{cases} \frac{x_i^\lambda - 1}{\lambda} & \text{if } \lambda \neq 0. \\ \log{[x_i]} & \text{if } \lambda = 0. \end{cases} \tag{6.2}$$

The value of λ is chosen to maximize the log-likelihood function (LLF).

As can be observed from Fig. 6.3, the features LR1 and SR1 have right-skewed distributions since all values of both features are always positive.
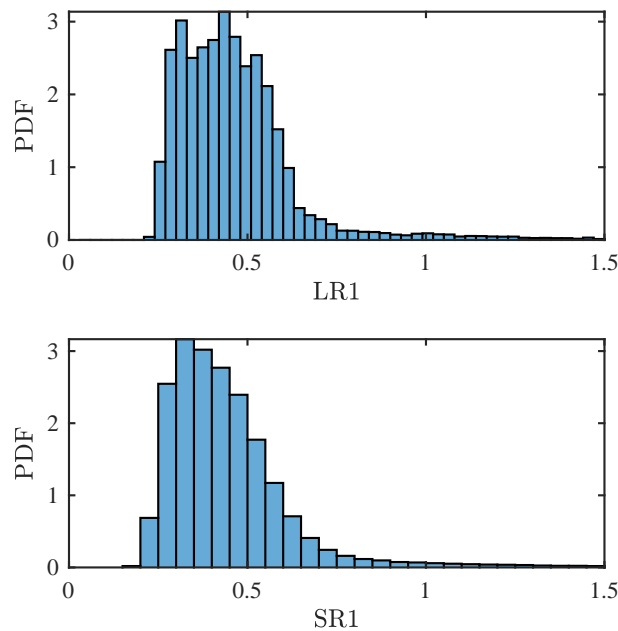
Figure 6.3: PDF of the LR1 and SR1 before Box-Cox transformation.

After the transformation, the distribution of the transformed data and the approximated Gaussian PDF is shown in Fig. 6.4. As can be noted, the transformed data fit better the Gaussian distribution, which will eventually lead to more accurate modeling.
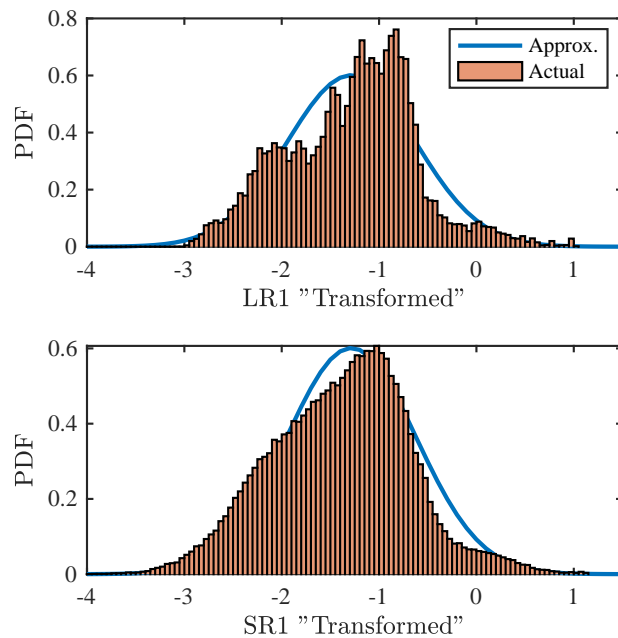
Figure 6.4: PDF of the LR1 and SR1 after Box-Cox transformation.

6.3   Density-based Anomaly Detection Algorithms

6.3.1   Gaussian-Based Algorithm Description

Since the sensors' measurements have been reshaped into Gaussian-distributed data, the anomaly detection algorithm can be finally be modeled using the following straight forward steps:

1. A subset of features $\mathcal{X}$ are chosen, whose $x_i \in \{1, 2, \ldots, n\}$ are deemed to be indicative of the anomalous examples, i.e., unusual values when anomalous examples occur. If all vibration measurements are used, then $n = 8$.

2. Estimate the PDF of each feature $x_i \in \mathcal{X}$ from the training samples, normal samples, $\{x_1, x_2, \ldots, x_n\}$. It is worth noting that the central limit theorem can be invoked, and hence, the distribution of each feature can be assumed to be Gaussian with mean μ and co-variance Σ, which can be estimated as

$$\mu = [\mathrm{E}\,[x_1]\,, \mathrm{E}\,[x_2]\,, \ldots, \mathrm{E}\,[x_n]]^T \tag{6.3}$$

where

$$\mathrm{E}\,[x_i] = \frac{1}{m}\sum_{k=1}^{m} x_i^{(k)} \tag{6.4}$$

and

$$\Sigma = (\mathcal{X} - \mu)(\mathcal{X} - \mu)^T \tag{6.5}$$

whose elements are

$$\Sigma_{i,j} = (x_i - \mu_i)(x_j - \mu_j) \tag{6.6}$$

where

$$\mathcal{X} = [x_1, x_2, \ldots, x_n]^T \tag{6.7}$$

3. Finding the joint Gaussian PDF using the parameters computed in step 2 as [34],

$$f_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n) = \frac{1}{\sqrt{(2\pi)^n\,|\Sigma|}} \times$$
$$\exp\left(-\frac{1}{2}(\mathcal{X} - \mu)^T \Sigma^{-1} (\mathcal{X} - \mu)\right). \tag{6.8}$$

If only one feature is used, then

$$f(x_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}}\exp\left(\frac{-(x_i - \mu_i)^2}{2\sigma_i^2}\right) \tag{6.9}$$

where

$$\mu_i = \frac{1}{m} \sum_{k=1}^{m} x_i^{(k)} \tag{6.10}$$

$$\sigma_i^2 = \frac{1}{m} \sum_{k=1}^{m} \left(x_i^{(k)} - \mu_i \right)^2. \tag{6.11}$$

4. For a new sample $\mathcal{X}_{new} = [x_1, x_2, \ldots, x_n]$, compute the probability,

$$\check{y} = (\Pr\{\mathcal{X}_{new}\} < \epsilon) \tag{6.12}$$

where

$$\Pr\{\mathcal{X}_{new}\} = f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n) \tag{6.13}$$

if true, flag it as an anomaly $\check{y} = 1$, otherwise, consider it as normal $\check{y} = 0$.

5. Buffer the output results obtained in Step 4 to a particular sliding window with a delay $\mathcal{D}_w$, then perform the decision based on a decision policy Π, where $\Pi \in [1, 2, \ldots, \mathcal{D}_w]$, such that the final alarm output is as follows:

$$\hat{y} = \left[\check{y}^{(n)} + \check{y}^{(n-1)} + \cdots + \check{y}^{(n-\mathcal{D}_w)} > \Pi\right] \tag{6.14}$$

$$= \sum_{i=1}^{\mathcal{D}_w} \check{y}^{(n-\mathcal{D}_w)} > \Pi. \tag{6.15}$$

This step is used to smooth the alarm decisions and maintain the time dependencies between the samples.

The perfect threshold value ϵ_{opt} is the one that can maximize the detection rate and minimize false alarms. More formally,

$$\epsilon_{opt} = \arg\max_{\tilde{\epsilon}} J(\theta, \tilde{\epsilon}) \tag{6.16}$$

where $J(\theta, \tilde{\epsilon})$ is a performance metric. To better illustrate how the model works, Fig. 6.5 shows the contour plot of the two features LR1 and SR1 jointly as a function of the anomaly tag. As it can be observed from the figure, the majority of the normal examples reside around the mean of their joint Gaussian PDF, while the outliers are located at the edges of the joint density. Therefore, by setting a suitable threshold, we will be able to depreciate normal and anomalous distributions based on the feature measurements. The optimization part will be discussed in Section 6.5. A summary of the Gaussian-based algorithm is outlined in Algorithm 1.
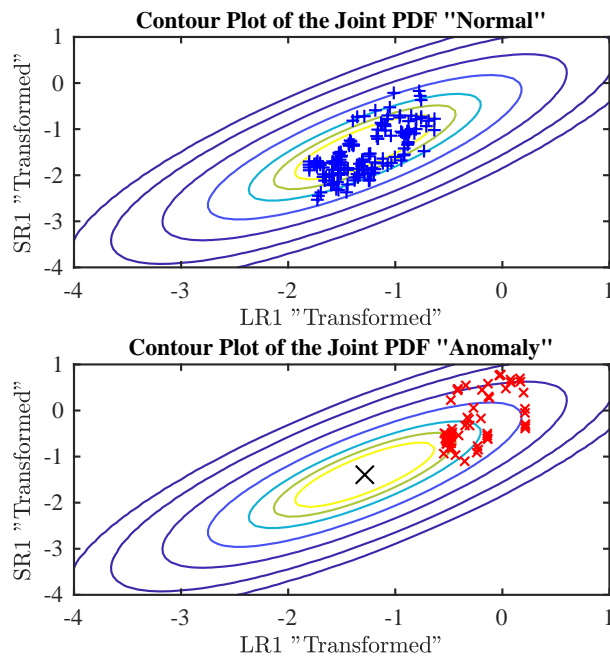
Figure 6.5: Contour plot of the joint distribution of LR1 and SR1 of the normal and anomaly examples.

Input: $\mathcal{X} = [x_1, x_2, \ldots, x_n]$
Output: Alarm Output, $\hat{y} \in [0, 1]$

1 PDF est. (mean):

$\mu = \left[\frac{1}{m} \sum_{k=1}^{m} x_1^{(k)}, \frac{1}{m} \sum_{k=1}^{m} x_2^{(k)}, \ldots, \frac{1}{m} \sum_{k=1}^{m} x_n^{(k)} \right]^T$ **2** PDF est. (co-variance):
$\Sigma = (\mathcal{X} - \mu) (\mathcal{X} - \mu)^T$

3 Evaluate the prob. of each sample:

$\breve{y}_i = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\left(-\frac{1}{2} (\mathcal{X} - \mu)^T \Sigma^{-1} (\mathcal{X} - \mu) \right)$ **4** Determine the anomaly flag per
sample:

if $\breve{y}_i < \epsilon$ **then**
| $\mathring{y} = 1$
end
else
| $\mathring{y} = 0$
end

5 Buffer the anomaly flags for w
$\mathring{\mathbf{y}} = [\mathring{y}_1, \mathring{y}_2, \ldots, \mathring{y}_w]$

6 Determine the flag w
if $\sum \mathring{\mathbf{y}} > \Pi$ **then**
| $\hat{\mathbf{y}} = \vec{1}_{1 \times w}$
end
else
| $\hat{\mathbf{y}} = \vec{0}_{1 \times w}$
end

Algorithm 1: Gaussian-Based Anomaly Detection Algorithm.

6.3.2   Autocorrelation Integrated Gaussian-based (AIG) Algorithm

The main drawbacks of the algorithm described in Section 6.3.1 is that it cannot carry any information that captures the temporal variation and correlation between the data samples during the model building phase. In other words, any decision made at sample t_n does not depend on the decision at sample t_{n-1}, which is a clear shortcoming of the model presented in Section 6.3.1, and consequently, affecting the detection capabilities. To overcome this problem, new features are created based on the autocorrelation function (ACF). More specifically, from each measurement signal, $x_i \in \mathcal{X}$ $i \in [1, 2, \ldots, n]$, we can create the autocorrelation matrix defined as:

$$R_{\mathbf{X_i X_i}} \triangleq \mathbf{X}_i \mathbf{X}_i^{\mathrm{T}} \tag{6.17}$$

where $(\cdot)^T$ denotes the transpose operation, and $\mathbf{X_i}$ is a measurement signal that is collected as a function of time over a window w such that:

$$\mathbf{X}_i = \left[x_i^{(n)}, x_i^{(n-1)}, \ldots, x_i^{(n-w)} \right]^T \tag{6.18}$$

Then, we can expand (6.17) as,

$$R_{\mathbf{X_i X_i}} \triangleq \begin{bmatrix} \overbrace{x_i^{(n)} x_i^{(n)}}^{1^{st}\text{feature}} & \overbrace{x_i^{(n)} x_i^{(n-1)}}^{2^{nd}\text{feature}} & \ldots & \overbrace{x_i^{(n)} x_i^{(n-w)}}^{w^{th}\text{feature}} \\ x_i^{(n-1)} x_i^{(n)} & x_i^{(n-1)} x_i^{(n-1)} & \ldots & x_i^{(n-1)} x_i^{(n-w)} \\ \vdots & \vdots & \ddots & \vdots \\ x_i^{(n-w)} x_i^{(n)} & x_i^{(n-w)} x_i^{(n-1)} & \ldots & x_i^{(n-w)} x_i^{(n-w)} \end{bmatrix} \tag{6.19}$$

Once the matrix is obtained, then we can use each column as a feature in the algorithm described in Section 6.3.1 following the steps. It is worth mentioning the dimensionality of the feature space has increased by a factor of w as compared to the conventional Gaussian based approach. For simplicity, the pseudocode of the AIG algorithm is given in Algorithm 2.

Input: $\mathcal{X} = [x_1, x_2, \ldots, x_n]$

Output: Alarm Output, $\hat{y} \in [0, 1]$

1 Buffer w samples for each input feature:
$$\mathbf{X}_i = \left[x_i^{(1)}, x_i^{(2)}, \ldots, x_i^{(w)} \right], i \in [1, 2, \ldots, n]$$

2 Create auto-correlation features from each input feature:
$$R_{\mathbf{X}_i \mathbf{X}_i} = \mathbf{X}_i \mathbf{X}_i^{\mathrm{T}}$$

3 PDF est. (mean):
$$\mu = \left[\tfrac{1}{m} \sum_{k=1}^{m} x_1^{(k)}, \tfrac{1}{m} \sum_{k=1}^{m} x_2^{(k)}, \ldots, \tfrac{1}{m} \sum_{k=1}^{m} x_{n \times w}^{(k)} \right]^T$$ **4** PDF est. (co-variance):
$$\Sigma = (\mathcal{X} - \mu)(\mathcal{X} - \mu)^T$$

5 Evaluate the prob. of each sample:
$$\breve{y}_i = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\left(-\tfrac{1}{2} (\mathcal{X} - \mu)^T \Sigma^{-1} (\mathcal{X} - \mu) \right)$$ **6** Determine the anomaly flag per sample:

if $\breve{y}_i < \epsilon$ **then**
| $\mathring{y} = 1$
end
else
| $\mathring{y} = 0$
end

7 Buffer the anomaly flags for w
$$\mathring{\mathbf{y}} = [\mathring{y}_1, \mathring{y}_2, \ldots, \mathring{y}_w]$$

8 Determine the flag w
if $\sum \mathring{\mathbf{y}} > \Pi$ **then**
| $\hat{\mathbf{y}} = \vec{1}_{1 \times w}$
end
else
| $\hat{\mathbf{y}} = \vec{0}_{1 \times w}$
end

Algorithm 2: AIG Algorithm.

6.3.3 Evaluation Metrics

In the context of anomaly detection in time-series data, conventional evaluation metrics such as accuracy, F1-score, and area under the curve (AUC) are not comprehensive enough to act as an accurate indicator of the detection model's performance. This is due to the fact that the aforementioned metrics format of measuring on sample-by-sample bases, while the aim of this work is to develop an algorithm for contextual time-series data. Therefore, we relied on a different set of measures capable of assessing the performance of the time-series anomaly detection algorithms:

1. **Number of correctly detected anomalies**: Any alarm that is being generated between during the anomaly process will count as a correctly detected anomaly.

2. **Number of missed alarms**: if no alarm is generated during the anomaly period, it will be considered a missed alarm.

3. **Number of false alarms**: any flagged alarm when there is no anomaly is a false alarm.

4. **True detection rate**: the ratio between the number of correctly detected anomalies and the total number of anomalies.

5. **Average delay**: the mean of the time delay between the start time of the anomaly and start time of the generated alarms.

The perfect alarm based on these metrics is the one that maximizes the number of correctly detected anomalies and the true detection rate while minimizing the number of missed alarms, number of false alarms, and the average delay.

6.4 Time Complexity and Memory Requirements

There is no unified metric to evaluate the computational complexity of anomaly detection algorithms. However, we can analyze the complexity of the proposed algorithm in Section. 6.3.2 by breaking down the algorithm major steps, and evaluate the computational requirements for each step. To simplify the discussion, the algorithm requires the following computations:

1. **Computing the auto-correlation matrix $R_{\mathbf{X_i X_i}} \triangleq \mathbf{X}_i \mathbf{X}_i^{\mathrm{T}}$:** Given that $\mathbf{X}_i \in \mathbb{R}^{w \times 1}$ as w being the window size. Therefore, $\mathbf{X}_i \mathbf{X}_i^{\mathrm{T}}$ can be seen as a matrix multiplication process with a complexity of $\mathcal{O}\left(w^2\right)$ for a single-window

size. Consequently, it requires $\mathcal{O}\left(m_w \times w^2\right)$ to span the entire dataset, where m_w is the number of windows in the training dataset. In terms of memory requirements, we need to store a matrix of size $m \times n$ temporary only for using it in the subsequent steps.

2. **Calculating the mean of each feature as in (6.3):** This operation requires $\mathcal{O}\left(m \times n\right),$ where m is the number of the data points and n is the number of features used. Also, we need only to store these values permanently in a one-dimensional array of size $1 \times n$.

3. **Evaluating the co-variance matrix:** This step is more complex than step (2) since it involves matrix multiplication. Noting that $\mathbf{X}_i \in \mathbb{R}^{m \times 1}$, then similar to step (1) given that we ignore the mean subtraction, this step needs $\mathcal{O}\left(m^2\right)$. The matrix needs to be stored temporary in an $n \times n$ array.

4. **Co-variance matrix inversion:** This operation is necessary to compute the joint Gaussian pdf (6.8). The complexity of the inversion operation depends on the implementation algorithm used. Assuming that Gauss-Jordan Elimination is adopted, then this step requires $\mathcal{O}\left(n^3\right)$. Unlike the previous step, this matrix needs to be stored *"permanently"* in an $n \times n$ array.

5. **Co-variance matrix determinant:** This step is also required to compute (6.8). Assuming that the Laplace expansion algorithm is used in this step, the computational cost of this step is $\mathcal{O}\left(n!\right)$. This value is going to be stored only in a single memory location permanently.

Although steps $(1) \rightarrow (5)$ may involve some expensive computational operations, more precisely steps (4) and (5), they are only required for the training phase. In other words, these processes can be implemented offline and store the outcomes of step (2), (4) and (5) in specific memory locations, which can not affect any delay-sensitive applications.

For real-time deployment, we need the following:

1. **Buffer:** w samples in temporary memory location

2. **Compute the auto-correlation matrix:** As we saw earlier, it only requires $\mathcal{O}(w^2)$, and storing the $w \times w$ matrix in a temporary storage location

3. **Evaluating the probability in (6.8):** the requires $\mathcal{O}(n^3)$ complexity

4. **Alarm decision:** This step needs only a simple logic gate.

As we can observe from the analysis, the major computational requirements of the algorithm are solely implemented offline, and the memory requirements remain minimal with the only need for it is to store a relatively small singular matrix. On the other hand, the real-time deployment needs are only buffering few samples and then perform straightforward matrix multiplication which is relatively cheap to evaluate for most practical values of w and n.

6.5   Optimizing the Performance of the Proposed Alarm

In order to enhance the efficiency of the proposed detection algorithms in Sections 6.3.1 and 6.3.2, optimizing the proposed algorithm is critical for improving the performance with respect to the cost functions. To properly evaluate the performance of the proposed algorithm, fair and quantified metrics should be proposed. These metrics are based on evaluation metrics in Subsection 6.3.3. Therefore, we can formulate the problem as:

- **Cost functions**

$$J_1\left(\epsilon, w, \Pi\right) = \sum \left[\vec{1} \oplus y\right] \cap \hat{y} \tag{6.20}$$

$$J_2\left(\epsilon, w, \Pi\right) = T_s \cdot \mathrm{E}\left\{\left[\vec{0} \oplus y\right] \cap \hat{y}\right\} \tag{6.21}$$

where (6.20) is the number of false alarms and (6.21) is the average delay as defined is Subsection 6.3.3.

- **Objective function**

$$\min_{\mathbf{x}} \mathbf{F}\left(\mathbf{x}\right) = \left[J_1, J_2\right], \forall \mathbf{x} \in \{\epsilon, w, \Pi\} \tag{6.22}$$

- **Subject to:**

$$0 < \epsilon \leq 1 \tag{6.23}$$

$$1 \leq w \leq w_{max} \tag{6.24}$$

$$1 \leq \Pi \leq w \tag{6.25}$$

As can be noted from the cost functions in (6.20)-(6.21), and from the underlying decision variables of each cost function, the optimization problem is a multiobjective mixed-integer non-linear programming (MINLP). In the context of this work, GA can be invoked to find the set of optimal points over the objective space. It is worth mentioning that GA is widely used in solving both constrained and unconstrained optimization problems that are based on natural selection, the process that drives biological evolution. The algorithm starts with an initial generation of candidate solutions that are tested against the objectives functions [35]. Consequently, the subsequent generation evolves from the first process through selection, crossover, and mutation.

6.6   Numerical Results

In this section, we validate the AIG algorithm performance on the dataset described in Section 6.2. The results are benchmarked against the true anomaly tags, which are given in the dataset, and compared with SVM and RF algorithms using the evaluation metrics presented in Subsection 6.3.3. The SVM algorithm is trained with a Gaussian kernel, and the RF classifier is trained with 50 trees. The results are based on the most two important signals which are obtained from the feature ranking are SR1 and LR1 measurements. The λ values for the LR1 and SR1 used in the Box-Cox model are $\lambda_{\mathrm{LR1}} = -0.8977$ and $\lambda_{\mathrm{HR1}} = -0.8691$, respectively. In addition, The number of newly constructed time features is set to $w = 10$. The validation of the AIG anomaly algorithm was conducted on an Intel Core i7 machine with 16 GB of RAM.

To start deploying the algorithm efficiently, the optimization problem formulated in Section 6.5 is solved, and the set of the optimal allocations is found as shown in the Pareto front allocation in Fig. 6.6. It is worth mentioning that each point on the Pareto front figure is mapped to a set of decision variables, $\{\epsilon, \mathcal{D}_w, \Pi\}$. Since no point is a dominant solution to another, the choice of the solution will be left to the operator to determine the best compromise between the two objectives: the number of false alarms and the average detection delay.

Fig. 6.7 shows the computing time complexity that is needed for training the model as well as deploying the algorithm for online detection for several new time features w. As Subfig. 6.7a indicates, the time required to complete the training task ranges between 70 to 300 seconds, where $250,000$ samples are used in the training task. It also suggests that adding new time features to enhance the time tracking of the process may cost more computational resources. Nevertheless, this step is required to be performed less frequently and offline. On the hand, Subfig. 6.7b shows the computational time needed to perform the detection "on the fly" per one experiment for various time features w. As can be noted from the figure, the
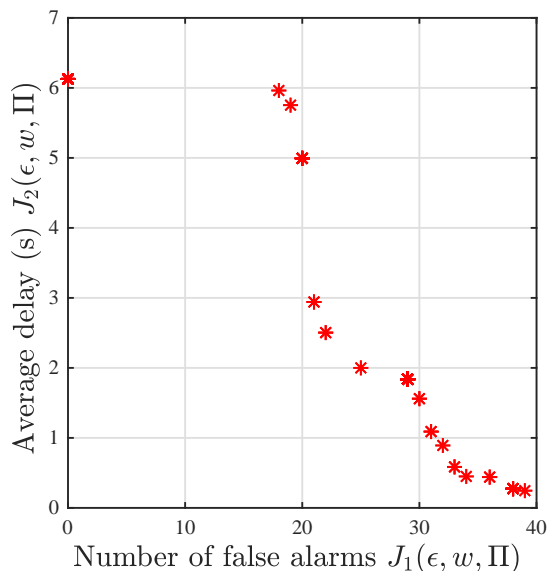
Figure 6.6: Pareto front of the two objective functions.

execution processing time varies from $8 - 55$ milliseconds. It is worth mentioning that the average experiment duration is 500 seconds, which imply that the processing time is negligible as compared to the process duration, which makes the proposed AIG algorithm a practical solution for real-time anomaly detection algorithm.

To validate the performance of the AIG algorithm, first, we chose the hyper-parameters from one of the solutions on the Pareto front with decision variables $\epsilon = 0.0719, \mathcal{D}_w = 106, \Pi = 11$. The validation in Fig. 6.8 is conducted on two experiments, the first with one anomaly period, and the second has two anomaly durations throughout the process. The experiment in Subfig. 6.8a shows that the AIG flags the alarm with a delay that is less than 10 seconds, with no flagged false alarm. On the other side, the AIG algorithm was also able to correctly detect both anomaly regions with an average delay per alarm of 38 seconds. This delay can be compensated by choosing a point that provides less delay on the Pareto front, however, more false alarms will be generated by the algorithm.

Fig. 6.9 shows the time complexity requirements for the training and the execution phase. As the figure indicates, the AIG has a considerable shorter training time
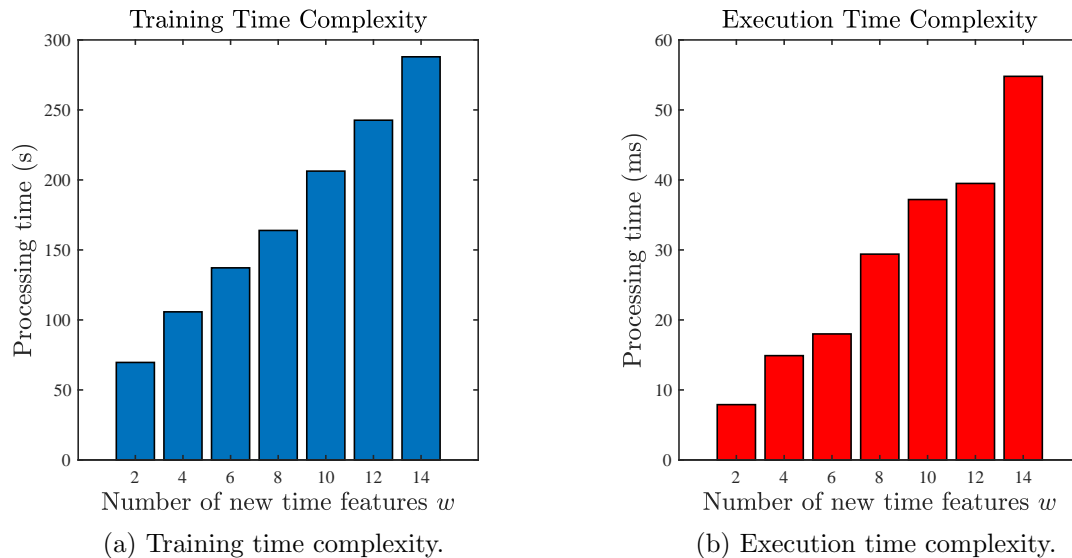
(a) Training time complexity. (b) Execution time complexity.

Figure 6.7: Time complexity of the AIG algorithm as a function of the number of the new time features w.

as compared to the SVM, and about 70% of the time required by the RF algorithm. This advantage is extremely important in the context of the IIoT, since re-training the model parameters is conducted frequently in order to cope with the process evolutions over time. On the other hand, the execution time, as shown in Fig. 6.9, is substantially low for the AIG as compared to RF and SVM, which makes the AIG algorithm a preferred option for real-time deployment.

To compare the detection performance of the AIG to other well established algorithms, the results for the evaluation metrics presented in Section 6.3.3 are presented in Fig. 6.10. As illustrated in the figure, the AIG shows a the has the highest number of truly detecting anomalies while maintaining the minimum number of false alarms. Furthermore, the AIG shows a detection success rate of approximately 95%, and 90%, 82% for the RF and SVM, respectively. In addition, it shows that the AGI missed one alarm as compared to two, four alarms the RF and SVM, respectively. The main cause for performance degradation of the RF and SVM is the class imbalance of the dataset, which leads to poor performance of the algorithms on the test dataset [36]. On the other hand, the Subfigure 6.10e shows that the AIG Average delay is
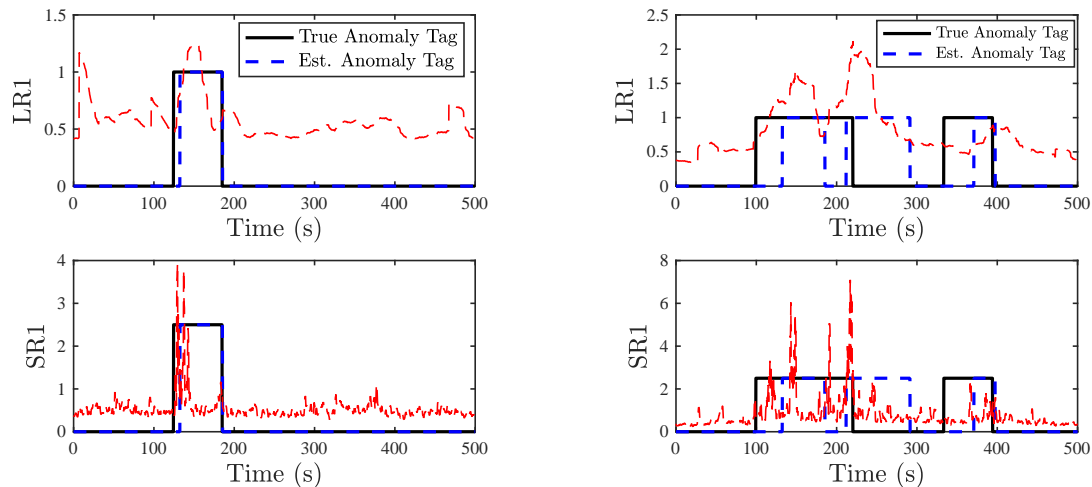
(a) Example of a heat experiment with one anomaly alarm during the process.

(b) Example of a heat experiment with two anomaly alarms during the process.

Figure 6.8: AIG anomaly detection algorithm assessment.

slightly higher than other classifiers by a slight margin. Nevertheless, the algorithm outperforms RF and SVM in 4 out of 5 performance metrics.

6.7 Conclusions

This chapter proposed a low-complexity anomaly detection algorithm for vibration sensor measurements presented in a time-series format. The proposed AIG algorithm is based on the Gaussian modeling of each measurement as well as an ACF to capture any temporal signal variations. The model parameters were tuned to optimize the set of objectives using multiobjective GA optimization. The obtained results were compared with conventional supervised learning algorithms such as SVM and RF. The results showed that the AIG outperformed the SVM and RF in most performance evaluation metrics with the advantage of a significant reduction in time complexity during both the training and execution phases, which highlights the AIG suitability for delay-sensitive applications and limited computational resources devices.

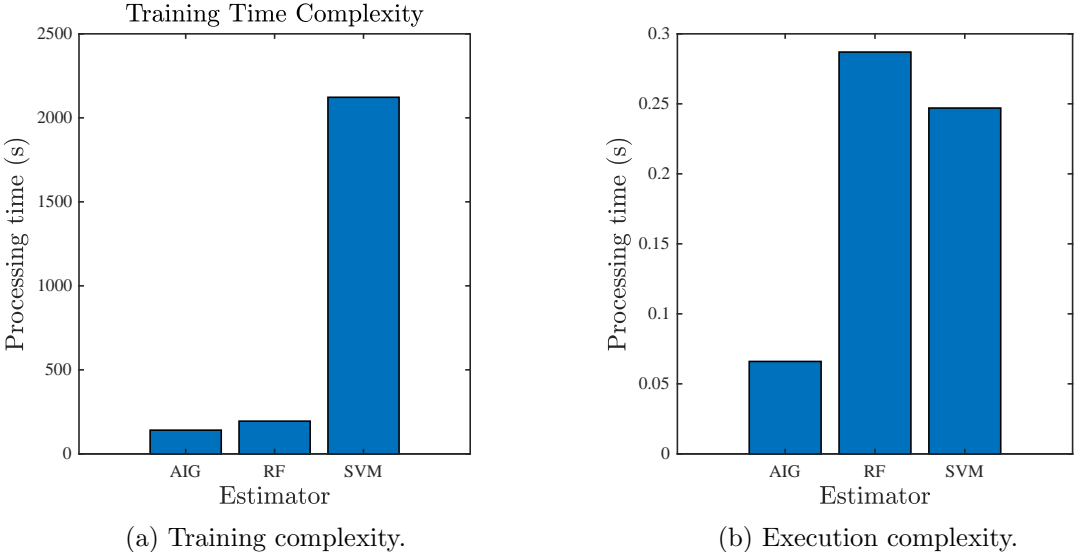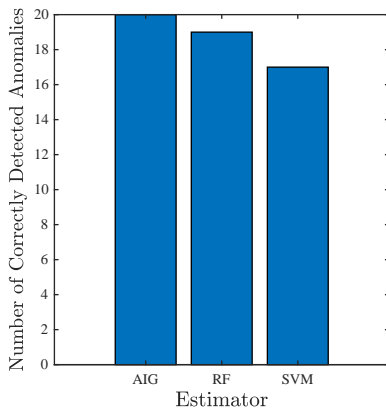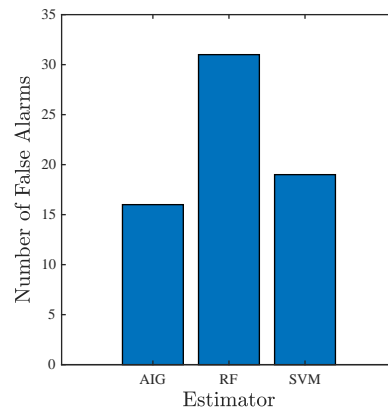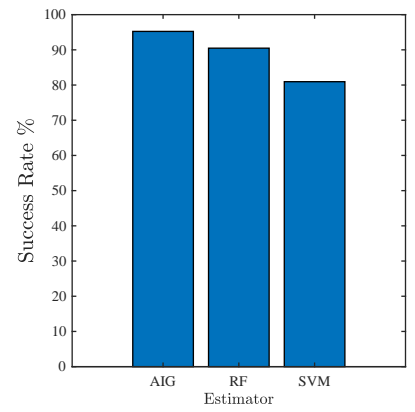(a) Training complexity.

(b) Execution complexity.

Figure 6.9: Time complexity comparison between AIG, RF and SVM.

(a) Number of truly detected alarm.

(b) Number of false alarms.

(c) Success detection rate.
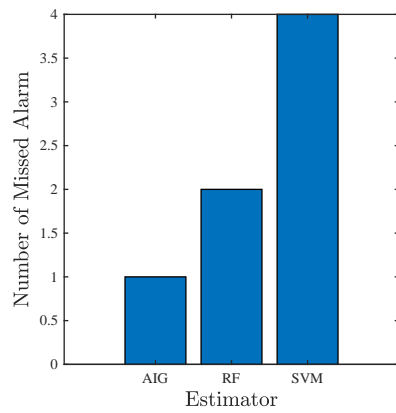
(d) Number of missed alarms.
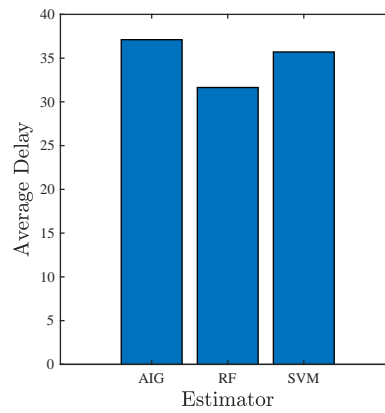
(e) Average delay.

Figure 6.10: Evaluation metrics of the AIG as compared with the RF and SVM.

Bibliography

[1] I. L. BRIEF, "Internet of things," 2019.

[2] Y. Xu and A. Helal, "Scalable cloud–sensor architecture for the internet of things," *IEEE Internet Things J*, vol. 3, no. 3, pp. 285–298, 2015.

[3] T. Wang, H. Ke, X. Zheng, K. Wang, A. K. Sangaiah, and A. Liu, "Big data cleaning based on mobile edge computing in industrial sensor-cloud," *IEEE Trans. Ind. Informat.*, vol. 16, no. 2, pp. 1321–1329, Feb 2020.

[4] Y. Kang, I. Park, J. Rhee, and Y. Lee, "Mongodb-based repository design for iot-generated rfid/sensor big data," *IEEE Sensors J.*, vol. 16, no. 2, pp. 485–497, 2016.

[5] S. Chen, H. Wen, J. Wu, W. Lei, W. Hou, W. Liu, A. Xu, and Y. Jiang, "Internet of things based smart grids supported by intelligent edge computing," *IEEE Access*, vol. 7, pp. 74 089–74 102, 2019.

[6] J. Leng, D. Yan, Q. Liu, K. Xu, J. L. Zhao, R. Shi, L. Wei, D. Zhang, and X. Chen, "Manuchain: Combining permissioned blockchain with a holistic optimization model as bi-level intelligence for smart manufacturing," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 50, no. 1, pp. 182–192, Jan 2020.

[7] T. Hussain, K. Muhammad, J. Del Ser, S. W. Baik, and V. H. C. de Albuquerque, "Intelligent embedded vision for summarization of multi-view videos in iiot," *IEEE Trans. Ind. Informat.*, 2019.

[8] M. S. Mahdavinejad, M. Rezvan, M. Barekatain, P. Adibi, P. Barnaghi, and A. P. Sheth, "Machine learning for internet of things data analysis: A survey," *Digital Commun. and Networks*, vol. 4, no. 3, pp. 161–175, 2018.

[9] M. Kwon, J. Lee, and H. Park, "Intelligent iot connectivity: Deep reinforcement learning approach," *IEEE Sensors J.*, vol. 20, no. 5, pp. 2782–2791, 2020.

[10] T. Yu, X. Wang, and A. Shami, "A novel fog computing enabled temporal data reduction scheme in iot systems," in *2017 IEEE Global Commun. Conf. (GLOBECOM).* IEEE, 2017, pp. 1–5.

[11] M. Injadat, F. Salo, A. B. Nassif, A. Essex, and A. Shami, "Bayesian optimization with machine learning algorithms towards anomaly detection," in *2018 IEEE Global Commun. Conf. (GLOBECOM).* IEEE, 2018, pp. 1–6.

[12] J. Park, "Anomaly detection from the signal of low-cost laser device without the false alarm and the missing," *IEEE Sensors J.*, vol. 18, no. 10, pp. 4275–4285, 2018.

[13] P. Ni, T. Haglund, and M. Ersson, "Study on slopping prevention in the bof steelmaking process," *steel research international*, vol. 88, no. 8, p. 1600399, 2017.

[14] M. Shakirov, A. Boutchenkov, G. Galperine, and B. Schrader, "Prediction and prevention of slopping," in *Proceedings from ISSTech 2003*, 2003.

[15] J. Kafie, B. Babaei, and V. Scipolo, "Bof slop detection - savings potential with real-time slop detection and mitigation technology," in *Proceedings from AISTech 2015*, 2015.

[16] D. Huang, D. Mu, L. Yang, and X. Cai, "Codetect: Financial fraud detection with anomaly feature detection," *IEEE Access*, vol. 6, pp. 19 161–19 174, 2018.

[17] I. A. Khan, D. Pi, Z. U. Khan, Y. Hussain, and A. Nawaz, "Hml-ids: A hybrid-multilevel anomaly prediction approach for intrusion detection in scada systems," *IEEE Access*, vol. 7, pp. 89 507–89 521, 2019.

[18] S. Jin, Z. Zhang, K. Chakrabarty, and X. Gu, "Toward predictive fault tolerance in a core-router system: Anomaly detection using correlation-based time-series analysis," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 37, no. 10, pp. 2111–2124, 2017.

[19] J. Jiang, J. Chen, T. Gu, K.-K. R. Choo, C. Liu, M. Yu, W. Huang, and P. Mohapatra, "Anomaly detection with graph convolutional networks for insider threat and fraud detection," in *MILCOM 2019-2019 IEEE Military Commun. Conference (MILCOM)*. IEEE, 2019, pp. 109–114.

[20] W. Rasheed and T. B. Tang, "Anomaly detection of moderate traumatic brain injury using auto-regularized multi-instance one-class svm," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 1, pp. 83–93, 2020.

[21] S. Rajendran, W. Meert, V. Lenders, and S. Pollin, "Unsupervised wireless spectrum anomaly detection with interpretable features," *IEEE Trans. Cogn. Commun. Netw*, vol. 5, no. 3, pp. 637–647, 2019.

[22] S. Naseer, Y. Saleem, S. Khalid, M. K. Bashir, J. Han, M. M. Iqbal, and K. Han, "Enhanced network anomaly detection based on deep neural networks," *IEEE Access*, vol. 6, pp. 48 231–48 246, 2018.

[23] Y. Lin and J. Wang, "Probabilistic deep autoencoder for power system measurement outlier detection and reconstruction," *IEEE Trans. Smart Grid*, vol. 11, no. 2, pp. 1796–1798, 2020.

[24] A. E. Mekki, A. Bouhoute, and I. Berrada, "Improving driver identification for the next-generation of in-vehicle software systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 7406–7415, 2019.

[25] K. Zhu, Z. Chen, Y. Peng, and L. Zhang, "Mobile edge assisted literal multi-dimensional anomaly detection of in-vehicle network using lstm," *IEEE Trans. Veh. Technol.*, vol. 68, no. 5, pp. 4275–4284, 2019.

[26] W. Yu, T. S. Dillon, F. Mostafa, W. Rahayu, and Y. Liu, "A global manufacturing big data ecosystem for fault detection in predictive maintenance," *IEEE Trans. Ind. Informat.*, 2019.

[27] Q. Jiang, X. Yan, and B. Huang, "Performance-driven distributed pca process monitoring based on fault-relevant variable selection and bayesian inference," *IEEE Trans. Ind. Electron.*, vol. 63, no. 1, pp. 377–386, 2015.

[28] W. Hu, J. Gao, B. Li, O. Wu, J. Du, and S. J. Maybank, "Anomaly detection using local kernel density estimation and context-based regression," *IEEE Trans. Knowl. Data Eng.*, 2018.

[29] M. Desforges, P. Jacob, and J. Cooper, "Applications of probability density estimation to the detection of abnormal conditions in engineering," *Proceedings of the Institution of Mechanical Engineers, Part C: J. of Mechanical Engineering Science*, vol. 212, no. 8, pp. 687–703, 1998.

[30] R. Moghaddass and J. Wang, "A hierarchical framework for smart grid anomaly detection using large-scale smart meter data," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 5820–5830, 2017.

[31] W. Haider, J. Hu, Y. Xie, X. Yu, and Q. Wu, "Detecting anomalous behavior in cloud servers by nested arc hidden semi-markov model with state summarization," *IEEE Trans. Big Data*, 2017.

[32] X. Yang and W. Wen, "Ridge and lasso regression models for cross-version defect prediction," *IEEE Trans. Rel.*, vol. 67, no. 3, pp. 885–896, 2018.

[33] R. Sakia, "The box-cox transformation technique: a review," *J. of the Royal Statistical Society: Series D (The Statistician)*, vol. 41, no. 2, pp. 169–178, 1992.

[34] Proakis, *Digital Commun. 5th Edition.* McGraw Hill, 2007.

[35] H.-L. Liu, L. Chen, K. Deb, and E. D. Goodman, "Investigating the effect of imbalance between convergence and diversity in evolutionary multiobjective algorithms," *IEEE Trans. Evol. Comput.*, vol. 21, no. 3, pp. 408–425, 2016.

[36] S. H. Ebenuwa, M. S. Sharif, M. Alazab, and A. Al-Nemrat, "Variance ranking attributes selection techniques for binary classification problem in imbalance data," *IEEE Access*, vol. 7, pp. 24 649–24 666, 2019.

Chapter 7

Conclusions and Future Directions

In this chapter, the main conclusions of this thesis and future research venues are given below:

7.1　Conclusions

- A novel blind channel estimator was introduced for OFDM systems with single transmit antenna based on a hybrid OFDM symbol structure, where pilot subcarriers in conventional OFDM systems are replaced by MPSK symbols, and the adjacent subcarriers are modulated using MASK. Therefore, the MASK symbol can be considered equivalent to the channel frequency response with respect to the MPSK symbol, and hence, the MPSK symbol can be immediately detected, and then used to estimate the channel in a DD manner.

- The proposed OSBCE can be incorporated effectively and efficiently in practical systems such as LTE-A standard. The proposed estimator requires one OFDM symbol to estimate the CFR, which makes it suitable for mobile channels, where the channel frequently varies in the time domain.

- Closed-form analytical analysis for the SER at MPSK symbols has been derived,

which are used as pilots for the subsequent steps. Also, a closed-form expression for the PDF of the phase error ψ_k^ℓ in Rayleigh fading channels with imperfect channel knowledge at the receiver and an accurate analytical expression for the MSE in DD manner has been derived.

- An extension to OSBCE has been introduced where it exploits the correlation between two consecutive OFDM systems to perform the channel estimation. The system performance was evaluated using LTE downlink specifications and for various speeds of the UE. Monte Carlo simulations, verified by the analytical results, showed that the proposed system supports high mobility speeds up to 150 km/h and produces accurate channel estimates comparable to pilot-aided systems, with improved power and spectral efficiencies. However, significant performance improvement can be achieved by using an adaptive time/frequency configuration.

- A new fair and reliable approach to compute and compare the spectral efficiency of various blind and non-blind communications systems was proposed. The new approach considers the fact that different subcarriers in OFDM systems may be modulated using different modulation types and orders to satisfy QoS requirements. Moreover, the proposed approach considers the modulation type constraint on the overall system spectral efficiency.

- The obtained spectral efficiency results showed that the modulation type constraint has a significant impact on the system spectral efficiency, which can make the spectral efficiency of pilot-aided systems higher than that of blind systems with modulation type constraint.

- A new receiver that performs the detection process directly from the FFT output symbols without the need of experiencing the conventional steps of channel estimation, interpolation, and equalization, which led to a considerable com-

plexity reduction. Moreover, the D^3 system can be deployed efficiently using the VA.

- The D^3 was analyzed theoretically where simple closed-form expressions were derived for the BER in several cases of interest and in terms of complexity analysis. The analytical and simulation results show that the D^3 BER outperforms the coherent pilot-based receiver in various channel conditions, particularly in frequency-selective channels where the D^3 demonstrated high robustness with a remarkable complexity reduction for different modulation types and orders.

- A low-complexity anomaly detection algorithm for vibration time-series sensor measurements was proposed. The algorithm is based on Gaussian modelling of each measurement as well as autocorrelation function to capture the temporal signal variations. Then, the performance of the system was optimized by tuning the hyperparameters of the proposed system against several objective functions. The model parameters can be tuned to optimize the set of objectives using multi-objective genetic optimization. The obtained results show that the AIG is efficient from both sides, the time complexity and the performance capabilities, which makes the AIG suitable for delay-sensitive applications and limited computational resources devices.

7.2   Future Work

- The OSBCE system presented in Chapter 2 did not consider synchronization errors such as carrier frequency offset (CFO). As a future extension of the work, designing a joint channel estimation and a carrier frequency estimation to mitigate the inter-carrier interference is an interesting work to be considered.

- We plan on extending the D^3system presented in Chapter 5 to work in a MIMO

configuration. The challenge is to implement a complexity efficient algorithm in MIMO since the number of cases will increase as the number of antenna increases.

- The D^3 can be extended to assess the performance of the D^3 in the presence of various system imperfections such as phase noise, synchronization errors and IQ imbalance.

- The design of the D^3 to work fully blindly without the need to use pilot symbols for phase ambiguity resolutions is a future extension of the presented D^3 system to improve the spectral efficiency.

- In order to further reduce the system's complexity, the design of a joint channel encoding with data detection of the D^3 system to reduce the system's complexity in a coded-system environment will be considered.

- Implement a hybrid detection algorithm to utilize neural networks and long short term memory (LSTM) networks to work in parallel with the AIG algorithm presented in Chapter 6 in order to further improve the detection capabilities.

- Since processes are non-stochastic process, then, training the algorithm should be done frequently to cope with the process evolutions and variations. Therefore, optimizing the retraining time of AIG is crucial for keeping the performance at desirable levels.

Appendix A

The total power P_T across all subcarriers can be written as

$$P_T = P_{sc} \times N \tag{A.1}$$

where P_{sc} is the power consumed by each subcarrier. Then, in blind systems, the power per information subcarrier $P_{sc,\text{inf}}^{blind}$ can be expressed by

$$P_{sc,\text{inf}}^{blind} = \frac{P_T}{N}. \tag{A.2}$$

Similarly, the power per information subcarrier for the pilot-aided systems $P_{sc,\text{inf}}^{pilot}$ is given by

$$P_{sc,\text{inf}}^{pilot} = \frac{P_T}{N - N_P}. \tag{A.3}$$

The excess power per information subcarrier is the difference between the power per information subcarrier in blind and pilot-aided systems, which can be written as

$$P_{sc,exc} = \frac{P_T}{N - N_P} - \frac{P_T}{N}. \tag{A.4}$$

Therefore, we can write the total excess power as

$$\begin{aligned} P_{T,exc} &= N \times P_{sc,exc} \\ &= N \left(\frac{P_T}{N - N_P} - \frac{P_T}{N} \right). \end{aligned} \tag{A.5}$$

Since the MASK symbols have the worst BER performance, the total excess power will be allocated to the N_P MASK symbols. Thus

$$
\begin{aligned}
P_{\text{MASK},exc} &= \frac{P_{T,exc}}{N_P} \\
&= \frac{N}{N_P} \left(\frac{P_T}{N - N_P} - \frac{P_T}{N} \right) \\
&= \frac{N}{N_P} \left(\frac{N}{N - N_P} - 1 \right) P_{sc,\text{inf}}^{blind}.
\end{aligned}
\tag{A.6}
$$

Consequently, the power per MASK subcarrier can be computed as

$$
P_{\text{MASK}} = P_{sc,\text{inf}}^{blind} + \frac{N}{N_P} \left(\frac{N}{N - N_P} - 1 \right) P_{sc,\text{inf}}^{blind}.
$$

Let $\alpha = N/N_P$, then

$$
\begin{aligned}
P_{\text{MASK}} &= \alpha \left(\frac{N}{N - \frac{N}{\alpha}} - 1 \right) P_{sc,\text{inf}}^{blind} + P_{sc,\text{inf}}^{blind} \\
&= \alpha \left(\frac{1}{1 - \frac{1}{\alpha}} - 1 \right) P_{sc,\text{inf}}^{blind} + P_{sc,\text{inf}}^{blind}
\end{aligned}
\tag{A.7}
$$

Because $\alpha \gg 1$ for most practical systems, then we can write

$$
\lim_{\alpha \to \infty} P_{MASK} = 2 P_{sc,\text{inf}}^{blind}.
\tag{A.8}
$$

Therefore, doubling the MASK symbols' power satisfies (3.25).

Appendix B

By defining the events $A_\psi > A_n \triangleq E_{\psi,n}$, $n \in \{0, 1, \ldots, \psi - 1\}$, then,

$$P_C|_{\mathbf{H}_0,\mathbf{1}} = P \left(\bigcap_{n=0}^{\psi-1} E_{\psi,n} \right) \tag{B.1}$$

Using the chain rule, $P_C|_{\mathbf{H}_0,\mathbf{1}}$ can be written as,

$$P_C|_{\mathbf{H}_0,\mathbf{1}} = \Pr \left(E_{\psi,\psi-1} \Big| \bigcap_{n=0}^{\psi-2} E_{\psi,n} \right) \Pr \left(\bigcap_{n=0}^{\psi-2} E_{\psi,n} \right). \tag{B.2}$$

For $\mathcal{K} = 2$, $\psi = 1$, $\tilde{\mathbf{d}}_0^{(0)} = [1, -1]$, $\tilde{\mathbf{d}}_0^{(1)} = [1,1]$, and thus,

$$
\begin{aligned}
P_C|_{\mathbf{H}_0,\mathbf{1}} &= \Pr(E_{1,0}) \\
&= \Pr(\Re\{r_1 r_2\} > \Re\{-r_1 r_2\}) \\
&= \Pr(\Re\{r_0 r_1\} > 0)
\end{aligned}
\tag{B.3}
$$

For $\mathcal{K} = 3$, $\psi = 4$, $\tilde{\mathbf{d}}_0^{(0)} = [1, 1, -1]$, $\tilde{\mathbf{d}}_0^{(1)} = [1, -1, -1]$, $\tilde{\mathbf{d}}_0^{(2)} = [1, -1, 1]$ and $\tilde{\mathbf{d}}_0^{(3)} = [1, 1,\ldots, 1]$. Using the chain rule

$$
\begin{aligned}
P_C|_{\mathbf{H}_0,\mathbf{1}} &= \Pr(E_{3,2}|E_{3,1}, E_{3,0}) \Pr(E_{3,1}, E_{3,0}) \\
&= \Pr(E_{3,2}|E_{3,1}, E_{3,0}) \Pr(E_{3,1}|E_{3,0}) \Pr(E_{3,0})
\end{aligned}
\tag{B.4}
$$

However, $\Pr(E_{3,0}) = \Pr(A_3 > A_0)$, and thus

Appendix B.

$$
\begin{aligned}
\Pr\left(E_{3,0}\right) &= \Pr\left(\Re\left\{r_0 r_1 + r_1 r_2\right\} > \Re\left\{r_0 r_1 - r_1 r_2\right\}\right) \\
&= \Pr\left(\Re\left\{r_1 r_2\right\} > \Re\left\{-r_1 r_2\right\}\right) \\
&= \Pr\left(\Re\left\{r_1 r_2\right\} > 0\right).
\end{aligned} \tag{B.5}
$$

The second term in (B.4) can be evaluated by noting that the events $E_{3,1}$ and $E_{3,0}$ are independent. Therefore $\Pr\left(E_{3,1}|E_{3,0}\right) = \Pr\left(E_{3,1}\right)$, which can be computed as

$$
\begin{aligned}
\Pr\left(E_{3,1}\right) &= \Pr\left(\Re\left\{r_0 r_1 + r_1 r_2\right\} > \Re\left\{-r_0 r_1 + r_1 r_2\right\}\right) \\
&= \Pr\left(\Re\left\{r_0 r_1\right\} > \Re\left\{-r_0 r_1\right\}\right) \\
&= \Pr\left(\Re\left\{r_0 r_1\right\} > 0\right)
\end{aligned} \tag{B.6}
$$

The first term in (B.4) $\Pr\left(E_{3,2}|E_{3,1}, E_{3,0}\right) = 1$ because if $A_3 > \{A_1, A_0\}$, then $A_3 > A_2$ as well. Consequently,

$$
P_C|_{\mathbf{H}_0,\mathbf{1}} = \Pr\left(\Re\left\{r_0 r_1\right\} > 0\right) \Pr\left(\Re\left\{r_1 r_2\right\} > 0\right). \tag{B.7}
$$

By induction, it is straightforward to show that $P_C|_{\mathbf{H}_0,\mathbf{1}}$ can be written as,

$$
P_C\,|_{\mathbf{H},\mathbf{d=1}} = \prod_{n=0}^{\mathcal{K}-2} \Pr\left(\Re\left\{r_n r_{\acute{n}}\right\} > 0\right). \tag{B.8}
$$

Curriculum Vitae

Name:	Anas Saci

Post-Secondary Education and Degrees:

University of Western Ontario
London, ON, Canada
2020 Doctor of Philosophy in Electrical and Computer Engineering

Jordan University of Science and Technology
Irbid, Jordan
2013 Bachelor of Science in Electrical Engineering

Honours and Awards:

ECE Graduate Student Research Award,
Electrical and Computer Engineering Department
University of Western Ontario, London ON, Canada, 2019

ECE Graduate Travel Award, Western University
Electrical and Computer Engineering Department
University of Western Ontario, London ON, Canada, 2018

Graduate Teaching Assistant Award Nominee,
Electrical and Computer Engineering Department
University of Western Ontario, London ON, Canada, 2017

Outstanding presentation in the NSERC Create symposium
Electrical and Computer Engineering Department
University of Western Ontario, London ON, Canada, 2018

Related Work Experience:	Graduate Teaching and Research Assistant
	University of Western Ontario
	2015-2020

Publications:	[J1] A. Saci, A. Al-Dweik, A. Shami and Y. Iraqi, "One-shot blind channel estimation for OFDM systems over frequency-selective fading channels," *IEEE Trans. on Commun.,* vol. 65, no. 12, pp. 5445-5458, Dec. 2017.

[J2] A. Saci, A. Al-Dweik and A. Shami "Autocorrelation integrated gaussian based anomaly detection using vibration sensory data" submitted to *IEEE Sensors J.*

[J3] A. Saci, A. Al-Dweik and A. Shami "Direct data detection of OFDM signals over wireless channels" submitted to *IEEE Trans. on Veh. Technol.*

[C2] A. Saci, A. Shami and A. Al-Dweik, "Blind channel estimation using cooperative subcarriers for OFDM systems," *IEEE Int. Conf. Commun. (ICC)*, Kansas City, USA, May 2018.

[C1] A. Saci, A. Shami and A. Al-Dweik, "Cross-layer spectral efficiency of adaptive communications systems with QoS constraints," IEEE 86th *Veh. Technol. Conf. (VTC-Fall)*, Toronto, ON, 2017.