UNIVERSITÉ DU LUXEMBOURG

PhD-FSTM-2021-052
The Faculty of Sciences, Technology and Medicine

DISSERTATION

Defense held on 15/07/2021 in Luxembourg
to obtain the degree of

DOCTEUR DE L'UNIVERSITÉ DU LUXEMBOURG

EN INFORMATIQUE

by

Abderrahmane MAYOUCHE
Born on 14 December 1990 in Hammamet, (Algeria)

MACHINE LEARNING FOR MIMO DETECTION AND EAVESDROPPING WITH SYMBOL-LEVEL PRECODING COUNTERMEASURES

Dissertation defence committee

Dr Bjorn Ottersten, dissertation supervisor
Professor, Université du Luxembourg

Dr Symeon Chatzinotas
Professor, Université du Luxembourg

Dr Tegawendé Bissyande, Chairman
Associate professor, University of Luxembourg

Dr Ralf Müller
Professor, Friedrich-Alexander Universität Erlangen-Nürnberg

Dr Christos Masouros, Vice Chairman
Professor, University College London

"People Who Say It Cannot Be Done
Should Not Interrupt Those Who Are Doing It."

George Bernard Shaw

"To my parents, my wife, and my son Ibrahim."

Abstract

Multiple-input multiple-output (MIMO) technology is an integral part of many current wireless communication systems that can drastically improve the data rates and the spectral efficiency. One major performance limiting factor in MIMO communication is the inter-channel interference (ICI) that adversely affects the transmission's achievable rate, since the receiver has to deal with multiple interfering symbol streams that are transmitted concurrently through a channel subject to random noise and interference. In the case when the channel-state information (CSI) is known at the receiver, i.e., CSIR, it could be used by the latter to compensate for the undesired effects of ICI. Although the problem of symbol detection in MIMO systems — where the knowledge of CSIR is available — is a well studied problem with numerous classical detection methods, the complexity of optimal detection methods increase prohibitively in systems with large dimensions, making them impractical for real-time communication.

The problem of signal detection in precoded MIMO channels without explicit knowledge of the CSIR is challenging and still being considered in recent research. In particular, this problem is a common occurrence in systems where CSI at the receiver is not available, e.g., time-division duplex (TDD) systems. In this thesis, we investigate the problem of multi-antenna signal detection in the case of a highly distorted received signal due to the ICI effects. The core idea of this thesis is to use pilot data, without explicitly estimating the CSI, to improve the detection performance at the receiver. Motivated by low-complexity signal detection and given the accessibility to pilot data, which form an integral part of communications systems, in this thesis, we propose ML based techniques for MIMO detection in systems where the downlink transmission is precoded using imperfect CSI at the transmitter.

Firstly, in the context of a single-user MIMO system, we address the problem of MIMO detection when the received signals are highly distorted, i.e., the case where the signal distortion is caused by signals being precoded with a highly degraded CSI at the transmitter (CSIT). In this setting, we propose ML-based MIMO detectors robust to severe CSIT degradation. The second and third contributions relate to a downlink multi-user multiple-input single-output (MU-MISO) system, for which we propose ML-based detectors that are robust to inaccurate CSIT for uncoded and coded systems, respectively. Herein, the proposed ML detectors are presented as eavesdropping attacks, where, by using the proposed ML detectors, an eavesdropper (Eve) is able to learn the symbol detection function based on precoded pilots and to detect the transmitted symbols, intended for legitimate users, with high accuracy. To counteract these attacks, six symbol-level precoding (SLP)-based countermeasures are proposed with varying security, complexity, and power consumption trade-offs. Numerical results validate the effectiveness of the proposed ML-based detectors and the robustness to the harmful effects of ICI.

Acknowledgments

I would like to thank the following people: my supervisors Prof. Björn Ottersten and Prof. Symeon Chatzinotas for their close supervision, Prof. Ralf Müller for his valuable feedback on my PhD progress during the CET meetings and on the thesis as well, Dr. Wallace Martins, Dr. Christos Tsinos, and Dr. Danilo Spano who provided technical support, Dr. Adel Metref, my mentor, for his precious guidance, Dr. Karim Lounis, the security expert, for his suggestions, Rustem Galiullin and Farouk Damoun, the machine learning experts, for their tips and assistance.

I would like also to thank my family, my wife, and my friends who supported me in every way during this PhD journey.

Abderrahmane Mayouche
Luxembourg, July 2021

Contents

Preface **1**

Support of the Thesis . . . 1

Publications . . . 1

Journal Papers . . . 1

Conference Papers . . . 2

Publications not Included in the Thesis . . . 2

1 Introduction **3**

1.1 Background . . . 3

1.2 Motivation and Problem Definition . . . 7

1.3 Related Work . . . 10

1.4 Discussion on Leveraged Data and Machine Learning Tools . . . 11

1.4.1 Precoded Pilots as Training Data . . . 11

1.4.2 Learning from Data . . . 12

1.5 Thesis Outline and Contributions . . . 13

2 Overview of Signal Design and MIMO Detection **16**

2.1 Signal Design . . . 16

2.1.1 Preliminaries in Block-Level Precoding . . . 20

2.1.2 Symbol-Level Precoding Survey . . . 23

2.2 MIMO Detection . . . 28

2.2.1 Preliminaries in MIMO Detection . . . 28

3 Generic ML Framework for MIMO Detection **32**

3.1 Communication System Requirements . . . 32

3.2 Supervised Learning and Classification . . . 33
3.3 Training Phase . . . 34
3.3.1 Pilots Collection . . . 34
3.3.2 Pilots Pre-Processing . . . 35
3.3.3 Model Fitting . . . 36
3.4 Inference Phase . . . 37
3.4.1 Data Collection . . . 37
3.4.2 Data Pre-Processing . . . 37
3.4.3 ML-based Signal Detection . . . 37
3.5 Real-Time Considerations . . . 38
3.6 Summary . . . 39

4 Data-driven Precoded MIMO Detection Robust to Channel Estimation Errors 40
4.1 System Model . . . 41
4.1.1 CSIT Estimation through Uplink Training . . . 42
4.1.2 Downlink Transmission . . . 43
4.2 Learning-Based MIMO Detection Frameworks . . . 46
4.2.1 Learning-Based Framework for the Proposed MIMO Soft Detection Scheme . . . 46
4.2.2 Learning-Based Framework for the Proposed MIMO Hard Detection Scheme . . . 50
4.2.3 Scalability of the Proposed ML Detection Frameworks to Multi-User MIMO Systems . . . 52
4.3 Lightweight Implementation of the Proposed Detection Frameworks . . . 53
4.4 Numerical Results . . . 57
4.5 Summary . . . 63

5 Learning-Assisted Eavesdropping and Symbol-Level Precoding Countermeasures for Downlink MU-MISO Systems 64
5.1 System Model . . . 65
5.2 ML-Based Attack . . . 67
5.2.1 Training Phase . . . 67

5.2.2 Inference Phase . . . 69
5.2.3 ML Attack Formulation . . . 70
5.2.4 Attack Example on a Benchmark Scheme - CISPM . . . 70
5.3 Countermeasure - PLS Scheme . . . 72
5.3.1 Secure Design Principle . . . 72
5.3.2 PLS Schemes . . . 72
5.3.3 Attack Example on PLS Scheme . . . 79
5.4 Simulation Results . . . 79
5.5 Summary . . . 89

6 Multi-Antenna Data-Driven Eavesdropping Attacks and Symbol-Level Precoding Countermeasures for Coded MU-MISO Systems 90
6.1 System Model . . . 91
6.2 ML-Based Attacks . . . 92
6.2.1 Motivation . . . 92
6.2.2 Adversarial Model . . . 96
6.2.3 ML Framework for the Proposed Soft Decoding Scheme . . . 97
6.2.4 ML Framework for the Proposed Hard Decoding Scheme . . . 99
6.3 Countermeasure: PLS Schemes . . . 101
6.3.1 PLS Random Scheme . . . 102
6.3.2 PLS Eve-Min-Power Scheme . . . 103
6.4 Simulation Results . . . 105
6.4.1 Parameters, Metrics, and Benchmarks . . . 105
6.4.2 Selection of ML Algorithms for the Eve Attack . . . 106
6.4.3 Comparison and Insights . . . 107
6.5 Summary . . . 115

7 Conclusions and Future Works 117
7.1 Conclusions . . . 117
7.2 Limitations and Future Works . . . 118
7.2.1 Outlook for Online Learning . . . 118
7.2.2 Limitation of using Eve's CSI at the BS . . . 118
7.2.3 Outlook for Deep Learning . . . 118

7.2.4 Using Interleaving with Channel Coding 119
7.2.5 Using ML Algorithms that Directly Process Complex Numbers 119
7.2.6 Using More Benchmarks in Precoding and Detection 119

8 Appendix: Regression-Based ML Framework for Soft Decoding **120**

Bibliography **125**

List of Figures

2.1 ICI in an SU-MIMO system. 17
2.2 MUI in an MU-MISO system with K users. 17
2.3 Structure of the CISPM precoding scheme. 27

3.1 Overview of the proposed generic ML detection framework. 35

4.1 Coherence time structure. 42
4.2 CISPM precoding design for QPSK constellation. 44
4.3 Structure of the CISPM precoding scheme. 45
4.4 Overview of the proposed learning-based soft detector. 47
4.5 Overview of the proposed learning-based hard detector. 51
4.6 Histograms of the estimated likelihoods for the ML - Soft scheme using RZF precoding with $N_{\mathrm{t}} = 15$, $N_{\mathrm{r}} = 8$, $\eta \in \{0, 6\}$ dB, a frame size of 2000, and QPSK modulation. 56
4.7 CISPM precoding - Noiseless received signal an Rx's antenna. 58
4.8 RZF precoding - Noiseless received signal an Rx's antenna. 58
4.9 BER vs. η/γ [dB], with $\tau = 1$ (perfect CSIT). 60
4.10 BER vs. η/γ [dB], with $\tau = 0.9$ (degraded CSIT). 60
4.11 BER vs. η/γ [dB], with $\tau = 0.8$ (severe CSIT degradation). 61
4.12 Runtime per SP [ms] vs. N_{r} using RZF precoding and $\eta = 6$ dB. 62

5.1 Downlink MU-MISO system comprised of: a BS with N_{t} antennas, K single-antenna users, and one Eve with M antennas. 65
5.2 Summary of the ML-based attack. 67
5.3 Example of PLS scheme using QPSK modulation. 73
5.4 16-QAM constellation showing the 6 lines boundary regions. 78

5.5 16-QAM with $N_t = 15$, $K = 6$, 15 dB target SINR, and $M = 1$. 81
5.6 BER vs. number antennas at Eve, with $N_t = 15$, $K = 6$, and $\delta = 0.1$ 82
5.7 BER vs. target SINR, with $N_t = 15$, $K = 6$, $\delta = 0.1$, and $M = 3$. 83
5.8 BER at intended user vs. γ, with $N_t = 15$, $K = 6$, $\delta = 0.1$, and $M = 3$. . . . 84
5.9 Total transmit power vs. target SINR, with $N_t = 15$, $K = 6$, $\delta = 0.1$, and $M = 3$. 85
5.10 Rate/power efficiency vs. number of antennas at Eve, with $N_t = 15$, $K = 6$, and $\delta = 0.1$. 86
5.11 Total transmit power vs. N_t, with $K = 6$, $\delta = 0.1$, and $M = 3$. 86
5.12 BER vs. δ, with $N_t = 15$, $K = 6$, and $M = 3$. 87
5.13 Rate/power efficiency vs. σ_{ze}^2, with $N_t = 15$, $K = 6$, $\delta = 0.1$, and $M = 3$. . . 88

6.1 Symbols used in the two pilot signals intended for user k. 93
6.2 Noiseless received signals at user k and Eve when the BS uses ZF precoding with a mean power of 5 dB. 94
6.3 Noiseless received signals at user k and Eve when the BS uses CISPM precoding. 95
6.4 Received signal at Eve when BS sends pseu-random sequences to every user. . 95
6.5 Overview of the ML-based soft decoding scheme. 97
6.6 Overview of the ML-based hard decoding scheme. 100
6.7 BER at Eve vs. number of antennas at Eve, with $r = 1/3$, $N_t = 15$, $K = 6$, and $\eta = \gamma_k = 6$ dB. 108
6.8 LLRs distribution of the Soft - CC decoding scheme with $r = 1/3$ and $M \in \{1, 9\}$. 109
6.9 BER at Eve vs. η/γ_k [dB], with $r = 1/3$, $N_t = 15$, $K = 6$, and $M = 11$. 110
6.10 FER at Eve vs. η/γ_k [dB], with $r = 1/4$, $N_t = 15$, $K = 6$, and $M = 11$. . . . 111
6.11 Total transmit power [dBW] vs. target SINR [dB], with $N_t = 15$, $K = 6$, and $M \in \{1, 13\}$. 112
6.12 BER at Eve using Soft - CC vs. the pilot percentage, with $r = 1/3$, $N_t = 15$, $K = 6$, $M = 11$, $\eta = \gamma_k = 6$ dB, and a frame size of 900 symbols. 114
6.13 Runtime per SP [ms] vs. M of proposed and benchmark schemes with $N_t = 15$, $K = 6$, $\gamma_k = 6$ dB, and a frame size of 900 symbols. 115

8.1 Overview of the regression-based ML framework for soft decoding. 121

8.2 BER at Eve vs. number of antennas at Eve, with $r = 1/3$, $N_t = 15$, $K = 6$, and $\eta = \gamma_k = 6$ dB. 122
8.3 BER at Eve vs. η/γ_k [dB], with $r = 1/3$, $N_t = 15$, $K = 6$, and $M = 11$. 124
8.4 FER at Eve vs. η/γ_k [dB], with $r = 1/4$, $N_t = 15$, $K = 6$, and $M = 11$. . . . 125

List of Tables

4.1 Prediction accuracy of the proposed learning-based detectors with several state-of-the-art classifiers when using CIPSM and RZF precoding schemes. 56
4.2 Channel coding parameters used for the simulations 57

5.1 Performance of different classifiers for SLP-based dataset. 72
5.2 Performance of different classifiers for countermeasure SLP-based dataset. . . 79

6.1 Channel coding parameters used for the simulations 106
6.2 Prediction accuracy of our proposed ML-based decoding schemes with several classifiers when using ZF, CIPSM, PLS random, and PLS Eve-min-power precoding schemes. 107

Abbreviations

ADC	Analog-to-Digital Converter
AI	Artificial Intelligence
AN	Artificial Noise
AMP	Approximate Message-Passing
AWGN	Additive White Gaussian Noise
BER	Bit-Error Rate
BLP	Block-Level Precoding
BR	Binary Relevance
BS	Base Station
CC	Classifier Chain
CI	Constructive Interference
CISPM	Constructive Interference for Sum Power Minimization
CRC	Cyclic Redundancy Check
CSI	Channel-State Information
CSIR	Channel-State Information at the Receiver
CSIT	Channel-State Information at the Transmitter
DAC	Digital-to-Analog Converter
DetNet	Detection Network
DFE	Decision Feedback Equalization
DL	Deep Learning
EEST	Energy Efficiency for Secure Transmission
EP	Expectation Propagation
EPNet	EP Network
Eve	Eavesdropper

5G	Fifth Generation of Cellular Networks
FDD	Frequency-Division Duplex
FEC	Forward-Error Correction
FER	Frame-Error Rate
i.i.d.	independent and identically distributed
LLR	Log-Likelihood Ratio
LU	Legitimate User
MAP	Maximum A Posteriori
Mbps	Mega-bit per-second
MC-CDMA	Multi-Carrier Code Division Multiple Access
MCC	Multi-Class Classification
MCS	Modulation and Coding Scheme
MF	Matched Filter
MIMO	Multiple-Input Multiple-Output
ML	Machine Learning
MLC	Multi-Label Classification
MLD	Maximum Likelihood Detector
MMNet	Massive MIMO Network
MMSE	Minimum Mean Square Error
MRT	Maximum Ratio Transmission
MUI	Multi-User Interference
MU-MISO	Multi-User Multiple-Input Single-Output
NN	Neural Network
OAMNet	OAPM Network
OAMP	Orthogonal AMP
OFDM	Orthogonal Frequency-Division Multiplexing
PLS	Physical-Layer Security
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase Shift Keying
Rx	Receiver
RV	Random Variable
RZF	Regularized Zero-Forcing

SD	Sphere Decoding
SDR	Semi-Definite Relaxation
SER	Symbol-Error Rate
SINR	Signal-To-Interference-Plus-Noise Ratio
SLC	Single-Label Classification
SLP	Symbol-Level Precoding
SM	Spatial Multiplexing
SNR	Signal-To-Noise Ratio
SOC	Second-Order Cone
SP	Symbol Preriod
SVM	Support Vector Machine
SU	Single-User
TDD	Time-Division Duplex
Tx	Transmitter
ZF	Zero Forcing

Preface

This Ph.D. work has been carried out from July, 2018 to June, 2021, at the Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg, Luxembourg, under the supervision of Prof. Björn Ottersten and Prof. Symeon Chatzinotas. The annual evaluations of the Ph.D. progress were duly performed by the *comité d'encadrement de thèse* (CET) members Prof. Björn Ottersten, Prof. Symeon Chatzinotas, and Prof. Ralf Müller.

Support of the Thesis

This Ph.D. work has been fully supported by the Engineering and Physical Sciences Research Council (EPSRC) and the Luxembourg National Research Fund (FNR) project, entitled Exploiting Interference for Physical Layer Security in 5G Networks (CI-PHY) under grant number FNR/INTER/FNR-RCUK/17/11607830. The partial support from SIGCOM is also gratefully acknowledged.

Publications

As first author among senior co-authors, I have been the main contributor to the publications below. These publications are referred to in the text by J for Journal and C for Conference.

Journal Papers

[MMCO21, J1] Abderrahmane Mayouche, Wallace A. Martins, Symeon Chatzinotas, and Björn Ottersten, "Data-driven Precoded MIMO Detection Robust to Channel Estimation Errors," *IEEE Open J. Commun. Soc.*, vol. 2, pages 1144-1157, 2021.

[MST+20, J2] Abderrahmane Mayouche, Danilo Spano, Christos Tsinos, Symeon Chatzinotas, and Björn Ottersten, "Learning-Assisted Eavesdropping and Symbol-Level Precoding Countermeasures for Downlink MU-MISO Systems," *IEEE Open J. Commun. Soc.*, vol. 1, pages. 535-549, 2020.

[MMT+21a, J3] Abderrahmane Mayouche, Wallace A. Martins, Christos Tsinos, Symeon Chatzinotas, and Björn Ottersten, "Multi-Antenna Data-Driven Eavesdropping Attacks and Symbol-Level Precoding Countermeasures," *IEEE Open J. Vehicular Tech.*, pages. 1-1, 2021.

Conference Papers

[MST+19a, C1] Abderrahmane Mayouche, Danilo Spano, Christos Tsinos, Symeon Chatzinotas, and Björn Ottersten, "Machine Learning Assisted PHYSEC Attacks and SLP Countermeasures for Multi-Antenna Downlink Systems," IEEE Global Commun. Conf. (GLOBECOM), Waikoloa, HI, USA, pages 1-6, 2019.

[MMT+21b, C2] Abderrahmane Mayouche, Wallace A. Martins, Christos Tsinos, Symeon Chatzinotas, and Björn Ottersten, "A Novel Learning-based Hard Decoding Scheme and Symbol-Level Precoding Countermeasures," IEEE Wireless Commun. Netw. Conf. (WCNC), Nanjing, China, pages 1-6, 2021.

Publications not Included in the Thesis

[MST+19b, C3] Abederrahmane Mayouche, Danilo Spano, Christos Tsinos, Symeon Chatzinotas, and Björn Ottersten, "SER -Constrained Symbol-Level Precoding for Physical-Layer Security," IEEE Conf. Commun. Netw. Sec. (CNS), Washington, DC, USA, pages 1-5, 2019.

Chapter 1

Introduction

This chapter introduces the context and the problems addressed in this thesis. The motivation, research gaps, contributions and organization of the thesis are presented in the subsequent sections.

1.1 Background

Driven by the increasing demands of higher speeds and lower latency wireless communication services, e.g., Enhanced Mobile Broadband (eMBB), Internet of Things (IoT), autonomous driving, and Virtual Reality (VR), fifth generation (5G) cellular networks are currently being provisioned worldwide, providing the basic infrastructure for these services. According to the Cisco Annual Internet Report [Cis20], by 2023, about two-thirds of the global population will have Internet access with a 5.3 billion total users, nearly 29.3 billion devices will be connected, accounting for more than three times the world population, and 5G connection speeds averaging 575 Mbps, resulting in 13 times more than the average mobile network connection speed. While 5G strives to meet these expectations, its successor generation, named 6G wireless system, is being proposed to overcome several limitations in 5G [SBC19, YXXL19, GLT+20].

Multiple-input multiple-output (MIMO) technology, i.e., the technology of transmitting multiple data streams from multiple antennas and of performing detection of these streams at a multi-antenna receiver, is an integral part of 5G communication systems that can drastically improve the data rates and the spectral efficiency (measured in bits/s/Hz) [Win87, Fos96, Tel99]. The use of multiple transmit and receive antennas helps address the

problem of scarcity of bandwidth resources, which is a big limiting factor in wireless communication systems, by simultaneously sending several data streams over the same time-frequency resources through the MIMO wireless channel [GJJV03]. In fact, MIMO technology has been widely deployed in modern communication systems [GJJV03]; it is a necessary enabler to help meet the increasing data rates demands in 5G and also for future cellular networks. In this setting, we define downlink transmission from the transmitter to the receiver while uplink transmission from the receiver to the transmitter.

One major performance limiting factor in MIMO communication is the inter-channel interference (ICI) [GHH+10], also known as multi-user interference (MUI) in the context of multi-user systems, that adversely affect the transmission's achievable rate. Particularly, in the MIMO spatial multiplexing (SM) configuration where several information streams are sent in the same time-frequency resource, the receiver has to detect multiple interfering symbol streams, transmitted concurrently through a channel subject to random noise and interference [YH15].

One approach to mitigate the harmful effect of ICI/MUI is to compensate for interference at the receiver. However, this method requires the knowledge of channel-state-information (CSI) at the receiver, known as CSIR. CSI characterizes the wireless channel, i.e., the scattering and reflections patterns of the wireless signals from the transmitter to the receiver. The CSI knowledge is usually acquired at the receiver through pilot-assisted channel estimation and commonly represented in the complex domain. Herein, the transmitter sends pilot symbols, also called reference symbols, that are known to all parties in wireless communications standards, including the intended receiver that estimates the CSIR. Thus, the idea of MIMO detection in this approach is to use this information at the receiver to compensate for the undesired effects of ICI/MUI. The problem of symbol detection in MIMO systems where the knowledge of CSIR is available and accurate is a well studied problem with numerous classical detection methods [Lar09].

For instance, the maximum likelihood detector (MLD) is the optimal detector that yields minimum joint error probability of detecting all the transmitted symbols. However, it entails exhaustive search with exponential runtime complexity and perfect CSI, which makes it impractical in real time systems. Sphere decoding (SD) algorithms [AEVZ02, GN06] were proposed to overcome the computational cost of the MLD, although the original SD was optimal and exponential in complexity. SD algorithms employ lattice search and offer bet-

ter computational complexity with a rather a relatively lower accuracy compared to the full search. Several other detectors have been proposed with varying accuracy-complexity tradeoffs. Linear receivers [YH15, Lar09], i.e., matched filter (MF), zero forcing (ZF), and minimum mean squared error (MMSE) detectors, are amongst the most common suboptimal detectors which provide low computational complexity with good detection performance. Non-linear receivers, such as decision feedback equalization (DFE), approximate message passing (AMP) [JGMS15], and semidefinite relaxation (SDR) detectors [LMS+10, JO08], offer better accuracy at the expense of higher computational complexity. What most of these classical MIMO detectors share is the two-step-based detection, through which they first estimate CSIR, and only then employ this knowledge for detection.

For large MIMO systems, CSI estimation is a huge issue due to the required signaling overhead. Moreover, the complexity of these schemes increases prohibitively as the system size increases, i.e., as the number of transmit/receive antennas and simultaneous data streams grows. Thus, there is a need for *one-shot* detection, i.e., performing detection in one process without estimating the CSIR, and low-complexity MIMO detection schemes that can perform well in large system dimensions [KAHF20]. In [ZWMM03], a one-shot detection method was proposed for the uplink of multi-carrier code division multiple access (MC-CDMA) systems, that jointly performs detection and CSI estimation.

However, to address the low-complexity aspect in large MIMO systems, several machine-learning (ML) based detectors have been proposed recently for MIMO signal detection. Unlike classical MIMO detection schemes that are based on hypothesis testing [Ver98], ML solves statistical problems using examples of input-output pairs. ML is a core subset of artificial intelligence (AI), which is an ensemble of tools and algorithms intended for making predictions or decisions through learning patterns from data [JZR+17]. ML is typically used when the underlying distribution are not known but could be characterized from sample examples. Particularly, in the last decade, ML witnessed a big revolution thanks to the *deep learning* (DL), which usually involves neural networks with numerous operations and layers and can theoretically solve hard and large problems [YLH15]. DL have shown impressive results in numerous fields, e.g., speech processing [HDY+12] and computer vision [HZRS16]. For communication systems, there is a growing interest in using ML/DL for different purposes, including ML for channel estimation [YLJ18], error correcting codes [NBB16], and end-to-end detection over continuous signal [DCHB18].

For the MIMO detection problem where the ICI/MUI effect is mitigated at the receiver using CSIR, several ML-based techniques have been proposed that are based on the two-step approach [SDW19, HWJL18, KAHF20, ZHL+21, SPBL20]. Specifically, the authors in [SDW19] achieved excellent detection performance with a deep neural network architecture called DetNet, e.g., with a performance matching an SDR baseline for independent and identically distributed (i.i.d.) Gaussian channels while running 30× faster. The work in [HWJL18] introduced OAMPNet, a deep learning based scheme that mimics the orthogonal approximate message-passing (OAMP) algorithm [MP17], which outperforms the OAMP algorithm in both i.i.d and small-sized Kronecker model-based correlated channels. We note that DetNet and OAMNet are both trained offline: using a single detector trained over several channel matrices. On the other hand, MMNet in [KAHF20] is an adaptive neural-network based detection scheme tailored to realistic channels with spatial correlation and is suitable for online training, where the training is performed for each coherence time instead of offline training where the module could be trained in advance for a large number of coherence times by generating random channel coefficients. In the same context of online training and motivated by practical implementation, EPNet [ZHL+21] was proposed to perform signal detection by unfolding the expectation propagation (EP) algorithm and training the damping factors. To support coded systems, a neural-network MIMO detector with impairments was proposed in [SPBL20], where the MIMO detection algorithm design is based upon projected gradient descent iterations for orthogonal frequency-division multiplexing (OFDM) systems. Empirical results show the robustness of the proposed detection scheme against several common communication impairments, since the neural network (NN) does not assume any specific model. Another approach to mitigate the ICI/MUI is to pre-compensate for its undesired effect at the transmitter instead of at the receiver. This approach is commonly known as transmit precoding [BO01].

Similar to the aforementioned approach that requires CSIR, the precoding-based approach also requires the knowledge of CSI but at the transmitter side, i.e., CSIT. A common methodology to obtain CSIT is through pilot-assisted channel estimation, where prior to downlink transmission, the receiver sends pilots for the transmitter to estimate the CSIT. This setting is common in a *reciprocal* channel, i.e., uplink and downlink CSI are the same. For instance, this reciprocity applies to time-division duplex (TDD) communications systems [BM13, MH06, GSK05, Ott96]. In such systems, CSIT can be available without explicitly

estimating it at the receiver CSIR. Thus, it is interesting to consider systems that can take advantage of CSIT to mitigate the channel effect at the receiver when CSIR is not explicitly available.

Using the CSIT knowledge, the transmitter can perform ICI/MUI mitigation/exploitation processing to the intended data streams prior to transmission, thanks to the multiple antennas capability which provides sufficient degrees of freedom to handle the interference effect. In particular, the CSIT knowledge is used to precode the transmitted signal to attain several goals: a) maximize the signal-to-noise ratio (SNR), which indirectly mitigates the interference effect, b) minimize the ICI/MUI effect at the receiver, c) equalize/mitigate the effect of fading, and/or d) exploit the ICI/MUI for power gains, where in all case the performance of the precoder is directly affected by the quality of the CSIT. In this context, very recently, in [ZZZX20] a deep NN-based precoding technique was proposed for finite-alphabet inputs, where the authors employ deep NNs to learn the input-output relationship of a nearly optimal precoder for mutual information maximization.

Nonetheless, in systems with real-time constraints, it is extremely difficult to obtain good CSIT. Channel estimation errors can occur due to several sources, for instance, pilot length, Doppler effect, and hardware impairments [Sch08]. Thus, MIMO detection robust to CSIT imperfections should be investigated in order to support such systems with practical limitations. Overall, the problem of one-shot detection in a fading MIMO channel with imperfect CSIT was not explored. We note that, very recently, deep NN-based precoding techniques were proposed

1.2 Motivation and Problem Definition

TDD MIMO communication systems, where data is transmitted periodically, e.g., 1/2 or 1/3 of the time, play an important role in modern communication systems. TDD systems are particularly suitable for asymmetric transmission demands and also in cases where paired frequency is not available. Frequency resources are limited and prohibitively expensive, thus in cases where it not possible to secure paired spectrum, TDD systems are considered despite their higher deployment and operating costs when compared to frequency-division duplex (FDD) systems, as FDD requires fewer base stations for the same coverage. In addition, TDD systems are also considered when the range and mobility are limited, due to propagation

delays and channel variations. In TDD mode, data is transmitted in both directions between the transmitter and the receiver using the same frequency resource but in different time slots. In this setting, one can exploit channel reciprocity [Smi04] to exchange uplink and downlink data in the same coherence time. In this thesis, to maintain reciprocity, we assume that the channel does not have ferromagnetic materials and does not include the non-linear effects of RF-chains. Reciprocity is based on the premise of wireless signals traveling in uplink and downlink directions will undergo the same physical perturbations, i.e., reflection, refraction, diffraction, etc.

In addition to the aforementioned advantages of reciprocity in TDD systems, this mode of operations has some disadvantages as well. For instance, since the transmission in TDD occurs in the whole bandwidth half of the time, the power amplifier is active in this time only and the average radiated power is cut in half. As a result, if we transmit at maximum peak power, the link budget will suffer a 3 dB loss in SNR when compared to FDD. Another limitation is TDD systems is the necessity of guard periods to avoid the uplink-downlink interference, which further limits the total bandwidth. In this mode, the users do not start uplink transmission until cell edge users' downlink transmission is done, where the guard period, i.e., the waiting time, is defined by the base station. Another disadvantage of TDD is due to the inter-cell synchronization, to avoid interference between the uplink and the downlink among cells. Otherwise, the cell edge users might receive a downlink signal from its associated BS and an uplink signal from a neighboring user belonging to another cell. Despite these limitations, TDD is suitable for several technologies and use cases, amongst which we find massive MIMO, that constitutes a core technology in 5G wireless systems and works best in this mode.

Since in TDD the same frequency band is used in both directions, the impulse response of the channel observed between any transmitter's antenna and receiver's antenna should be the same regardless of the direction when the hardware if properly calibrated. Hence, CSIT is acquired in the uplink channel, usually with pilot-assisted channel estimation using uplink pilots, and used for downlink transmission (reverse channel). In this thesis, we consider multi-antenna systems subject to ICI/MUI effects, where between each transmit and receive antenna, the wireless channel has fading and noise effects. To mitigate the harmful effects of interference, we employ precoding, which requires the knowledge of CSIT. The quality of the estimated CSIT depends mainly on the pilot sequence length, the SNR, the particular

channel estimation technique that is being employed, and the hardware impairments [Sch08].

As mentioned in the previous section, the CSIT is employed by the transmitter to mitigate the harmful effects of ICI/MUI. In cases where the CSIT quality deteriorates, the effect of ICI/MUI will not be handled properly and thus the detection performance at the receiver will be poor or at best limited. In this thesis, *MIMO detection schemes robust to CSIT deterioration are proposed to support TDD systems with practical limitations*. Specifically, we develop one-shot ML-based MIMO detection schemes, that 1) perform MIMO detection without explicitly estimating the CSIR and 2) use ML to make the best use of prior information (pilot sequences) for detection purposes.

In addition, we also study a similar problem to the aforementioned one, where we investigate multi-antenna eavesdropping in multi-user multiple-input single-output (MU-MISO) systems. In such systems, a multi-antenna transmitter, commonly referred to as a base station (BS), sends simultaneously different data streams to several single-antennas users, known as *legitimate users* (LUs). In this context, the BS usually uses CSIT to precode the transmitted data in order to mitigate the ICI/MUI at the LUs. We note that the employed CSIT characterizes the propagation environment between the BS and the LUs. If the utilized CSIT at the BS is precise enough, the harmful MUI effects are managed well, and hence LUs can accurately detect the transmitted data. In parallel, given the broadcasting nature of the wireless channel, unintended receivers, e.g., an eavesdropper (Eve), may detect sensitive information [LCW17]. Herein, Eve receives a distorted signal as the transmitted signal was not intended for it. This is equivalent to the case where the BS used a completely inaccurate CSIT to serve a multi-antenna LU. Thus, for the eavesdropping attack to be successful, Eve needs to use MIMO detection schemes robust to highly deteriorated CSIT. Therefore, the second problem investigated in this thesis is essentially *MIMO detection schemes that are robust to completely inaccurate CSIT, intended for eavesdropping purposes in the physical-layer*. In this framework, to counteract these eavesdropping attacks, we propose numerous countermeasures that obstruct the detection process at Eve. In the subsequent section, we discuss the relevant works to these problems and the underlying research gap. To this end, we investigate systems with and without channel coding. Channel coding, also known as forward error correction (FEC), is the mechanism through which bit-errors could be detected and corrected in digital communication systems. We refer to "coded" systems for systems that employ channel coding and "uncoded" systems otherwise.

1.3 Related Work

This section encompasses the relevant works to the aforementioned problems and the underlying research gaps, with a focus on one-shot MIMO detection.

MIMO Detection Schemes Robust to CSIT Deterioration for TDD Systems with Practical Limitations

A few one-shot ML-based techniques have been proposed [LCK16, YYX+17, JHL17, NNT+20] for MIMO detection. MIMO detection schemes that do not use CSI, also known as blind MIMO detection, and are clustering based were proposed in [LCK16, YYX+17]. A major drawback to these approaches was identified by some ambiguity issues due to the underlying unsupervised learning detectors. However, the approaches in [JHL17, NNT+20] employ supervised learning instead, where downlink pilot sequences are sent to train the proposed learning frameworks by considering systems with low-resolution analog-to-digital converters (ADCs). The key idea in [JHL17] is to interpret the MIMO detection problem as a supervised classification problem, where one-shot based MIMO detection schemes were proposed, that exploits the knowledge of pilot sequences without explicitly estimating the CSIR. In [NNT+20], similar detectors were proposed that employ more efficient learning methods and leverage the knowledge of cyclic redundancy check (CRC) and the to-be-detected data to further assist the training process. However, none of these works consider CSIT, which is, as previously detailed, a common setting in TDD systems. Particularly, robust one-shot MIMO detection to CSIT deterioration has not been investigated in the literature.

MIMO Detection Schemes Robust to Completely Inaccurate CSIT

By 2023, there will likely be 5.7 billion total mobile users (71% of the world population) [Cis20]. In such a crowded environment, unintended receivers, e.g., an Eve, may decode sensitive information given the broadcasting nature of the wireless channel [LCW17]. As a result, security is of primary importance in next generation networks. In particular, PLS stands out as a powerful technology to complement encryption-based methods [WKX+18], including application-layer encryption.

The essence of PLS is to exploit the characteristics of the wireless channel, i.e., fading, noise, interference, and diversity, to attain an acceptable decoding performance at intended

users while obstructing the correct decoding at Eve. Alternatively, the aim of PLS is to increase the gap of correct decoding rates between intended users and Eve [HFA19]. PLS is foreseen to be used as a complementary layer of protection, in addition to the existing cryptography-based security methods. As the rise of quantum computing [NCL+20,MVZJ18] is threatening both symmetric and asymmetric cryptography, non-cryptographic-based methods such as PLS are needed [WML+20,WTZ+20,ZLZ+20]. In this setting, the *artificial noise* (AN) scheme [GN08] and its extensions [WYX12,WLXY13,WZX15,WWN15] have been proposed to improve PLS.

In the context of physical-layer eavesdropping LUs in MU-MISO system with CSIT-based precoding, a smart Eve approach was proposed in [XRS21] that exploits the statistical information learned from the precoded data. However, the proposed approach does not consider eavesdroppers with multiple antennas, which is increasingly being deployed in sophisticated devises, nor ML tools to learn from the precoded data, as ML algorithms are amongst the most efficient approaches to learn from data. Hence, more realistic eavesdropping attacks that take advantage of the recently available technologies in wireless devices, e.g., multi-antennas and ML tools, should be developed and evaluated to assess security vulnerabilities in modern communication systems. At a fundamental level, these attacks are basically MIMO detection schemes robust to completely inaccurate CSIT.

1.4 Discussion on Leveraged Data and Machine Learning Tools

In this section, we discuss the utilized data and tools in the proposed MIMO detection schemes.

1.4.1 Precoded Pilots as Training Data

In downlink transmission, pilot symbols, also known as reference signals, are known sequences to all parties and constitute an integral part of communication systems. They are often used to estimate parameters at the receiver to assist the detection process at the latter. Sometimes, the estimated parameters are fed back to the transmitter to be used to improve the communication performance. The downlink pilot sequences are commonly used for CSI and signal-to-interference-plus-noise ratio (SINR) estimation. Specifically, non-precoded or un-processed pilot symbols are typically used for CSIR estimation while precoded/processed

pilot symbols are intended for SINR estimation at the receiver [ETS14]. In this thesis, we are interested in the latter case, precoded pilot symbols, which use the same modulation and coding scheme (MCS) used for precoding the data.

In fact, pilot sequences are known entities to all parties. That is, the receiver knows exactly which symbols constitute these pilot sequences and their placement in the frame. In addition, the receiver also observes a signal during the transmission of these sequences, which we denote by received pilot signals. Consequently, the receiver can collect this data in input-output pairs, where during each symbol period (SP), the input represents the received pilot signal while the output is the actual transmitted pilot symbols. Considering a frame with N SPs dedicated for pilots, the receiver collects N examples of input-output pairs, that we denote as the training data. In this thesis, we leverage the training data knowledge to counteract the effect of CSIT imperfections in the detection process.

1.4.2 Learning from Data

As discussed earlier in Section 1.1, ML has been used to design low-complexity MIMO detectors that can perform well in large systems. Motivated by the former and given the nature of the training data we have access to, naturally, we employ ML to design MIMO detectors that solve the aforementioned problems described in Section 1.2 in the context of online learning. As opposed to offline learning where the training is performed in advance for a large number of coherence times, in online learning, the detector is optimized for each channel realization, i.e., the training is performed for each coherence time. ML algorithms are amongst the best and most efficient tools to solve statistical problems using examples of inputs and their desired outputs, with little or no prior information or assumptions on the functional relationship between input-output pairs. Besides, since the same CSIT and MCS used for pilots are also used for transmitting the data, consequently, the receiver can use ML to learn a function that maps an input to an output from the training set and use it to infer the transmitted data symbols. As the training data is labeled, we consider the supervised learning class of ML. However, since the training data is very limited in each coherence time, we resort to non NN based approaches [SB14]. Therefore, in this thesis, we propose ML-based frameworks that leverage precoded pilots as training data for 1) MIMO detection robust to CSIT imperfections and 2) multi-antenna eavesdropping attacks. To counteract these eavesdropping attacks, we propose symbol-level precoding (SLP) schemes that enhance physical-layer security (PLS),

by impeding the detection process at Eve.

1.5 Thesis Outline and Contributions

In the thesis, there are three main contributions, which are organized into five chapters. In short, in systems with CSIT and no explicit CSIR, we address the problem of MIMO detection robust to CSIT deterioration in TDD systems with practical limitations as well as the generalization to MIMO detection with fully inaccurate CSIT. We first start by highlighting some previous work on signal design and MIMO detection in Chapter 2. Next, we present a generic ML framework for MIMO detection which lays the foundation for the proposed three main contributions in Chapter 3. The first contribution is presented in Chapter 4, in which we address the first problem detailed in Section 1.2, where we study MIMO detection robust to CSIT imperfections for TDD systems with practical limitations. The second problem in Section 1.2 is studied in Chapters 6 and 5, for MU-MISO systems with and without channel coding, respectively. More details about each chapter and the publications related to each are presented below.

Chapter 2: Overview of Signal Design and MIMO Detection

In this chapter, we present a detailed literature review on the different signal design schemes, with an emphasis on SLP and its application to PLS and classical MIMO detection methods.

Chapter 3: Generic ML Framework for MIMO Detection

This chapter introduces a generic ML-based framework for signal detection that employs the available downlink precoded pilots as training data, laying down the basic structure for the MIMO detection schemes proposed in the following chapters.

Chapter 4: Data-driven Precoded MIMO Detection Robust to Channel Estimation Errors

In this chapter, we develop MIMO detection schemes robust to CSIT degradation, where we study the problem of symbol detection in downlink MIMO systems with precoding and channel coding but without explicit CSIR. To mitigate the effect of CSIT deterioration at the receiver, we propose ML-based techniques for hard and soft detection that use downlink

precoded pilot symbols — originally intended for SINR estimation — as training data. We validate the approach by proposing a lightweight implementation that is suitable for online training. Numerical results show that severe CSIT degradation impedes the correct detection when a conventional detector is used, however, the proposed ML detectors can achieve good detection performance even under severe CSIT deterioration. The findings of this chapter appeared in the following journal paper:

[MMCO21] Abderrahmane Mayouche, Wallace A. Martins, Symeon Chatzinotas, and Björn Ottersten, "Data-driven Precoded MIMO Detection Robust to Channel Estimation Errors," *IEEE Open J. Commun. Soc.*, vol. 2, pages 1144-1157, 2021.

Chapter 5: Learning-Assisted Eavesdropping and Symbol-Level Precoding Countermeasures for Downlink MU-MISO Systems

Herein, we leverage the proposed generic ML MIMO detection framework for eavesdropping purposes in the context of a downlink MU-MISO systems without channel coding, where an Eve is able to learn the symbol detection function based on precoded pilots and detects the transmitted symbols, intended for legitimate users, with high accuracy. To counteract this attack, we propose a novel SLP schemes that enhances PLS while guaranteeing a constructive interference effect at the intended users. In the numerical results, we validate both the eavesdropping attack as well as the countermeasures. The findings of this chapter appeared in the following papers:

[MST+20] Abderrahmane Mayouche, Danilo Spano, Christos Tsinos, Symeon Chatzinotas, and Björn Ottersten, "Learning-Assisted Eavesdropping and Symbol-Level Precoding Countermeasures for Downlink MU-MISO Systems," *IEEE Open J. Commun. Soc.*, vol. 1, pages 535-549, 2020.

[MST+19a] Abderrahmane Mayouche, Danilo Spano, Christos Tsinos, Symeon Chatzinotas, and Björn Ottersten, "Machine Learning Assisted PHYSEC Attacks and SLP Countermeasures for Multi-Antenna Downlink Systems," IEEE Global Commun. Conf. (GLOBECOM), Waikoloa, HI, USA, pages 1-6, 2019.

Chapter 6: Multi-Antenna Data-Driven Eavesdropping Attacks and Symbol-Level Precoding Countermeasures for Coded MU-MISO systems

In this chapter, we extend the previous chapter's work in two directions: 1) by extending the ML attack to support channel coding, and 2) by proposing more sophisticated PLS schemes in terms of security and low-complexity. In this setting, we exploit ML tools to design soft and hard decoding schemes by using precoded pilot symbols as training data. Simulation results validate both the ML-based eavesdropping attacks as well as the countermeasures, and show that the gains in security are achieved without affecting the decoding performance at the intended users. The findings of this chapter appeared in the following papers:

[MMT+21b] Abderrahmane Mayouche, Wallace A. Martins, Christos Tsinos, Symeon Chatzinotas, and Björn Ottersten, "A Novel Learning-based Hard Decoding Scheme and Symbol-Level Precoding Countermeasures," IEEE Wireless Commun. Netw. Conf. (WCNC), Nanjing, China, pages 1-6, 2021.

[**?**] Abderrahmane Mayouche, Wallace A. Martins, Christos Tsinos, Symeon Chatzinotas, and Björn Ottersten, "Multi-Antenna Data-Driven Eavesdropping Attacks and Symbol-Level Precoding Countermeasures," *IEEE Open J. Vehicular Tech.*, under revision.

Chapter 7: Concluding Remarks and Future Work

This chapter concludes the thesis and suggests some possible future directions to the current work.

Chapter 2

Overview of Signal Design and MIMO Detection

This chapter introduces the mathematical model and the underlying assumptions used in this thesis and provides an in-depth literature review and preliminaries on signal design and signal detection. For signal design section, we review block-level and symbol-level precoding, where we discuss the application of the latter on PLS. Regarding signal detection, we start by discussing classical signal detection schemes, followed by an overview of ML for communications and particularly a survey on ML-based detection schemes.

2.1 Signal Design

In this thesis, we consider two downlink communication systems, single-user (SU) MIMO and MU-MISO systems, to address important problems in two scenarios. The former system is considered in Chapter 4 while the latter is employed in Chapters 5 and 6. For the SU-MIMO system, as depicted in Figure 2.1, a multi-antenna transmitter (Tx) sends concurrently multiple data streams to a multi-antenna receiver (Rx), thus the interference in this case is of type ICI. In the MU-MISO system, as depicted in Figure 2.2, the BS sends multiple data streams simultaneously to several single-antenna users and the resulting interference is MUI. To mitigate the adverse effect of ICI/MUI, precoding at the Tx/BS is commonly used and is suitable for both systems. Precoding is a well-known technique to perform the downlink transmission in such systems, which exploits the transmitter's multi-antenna capability and spatially multiplexes the independent data stream intended for each receive antenna.

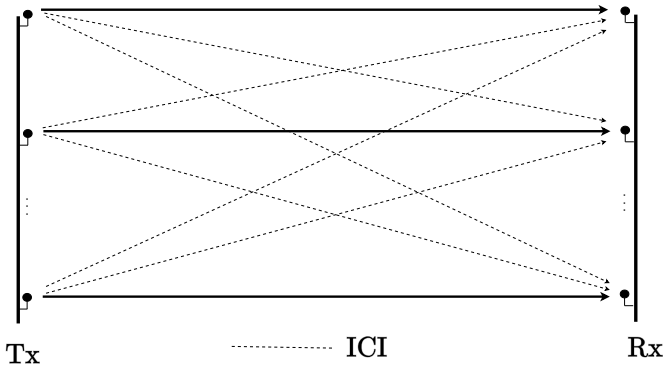


Figure 2.1: ICI in an SU-MIMO system.

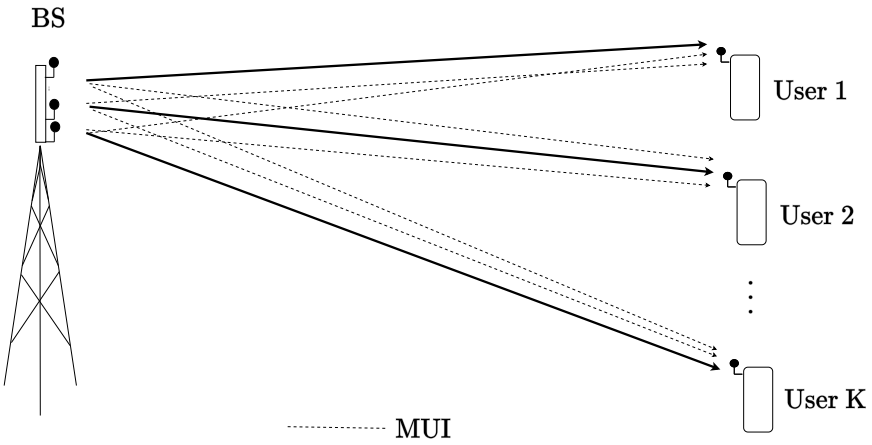


Figure 2.2: MUI in an MU-MISO system with K users.

To illustrate, let us consider an SU-MIMO with a Tx equipped with N_{t} antennas and communicates with an Rx with N_{r} antennas, supporting N_{r} independent data stream with $N_{\mathrm{r}} \leq N_{\mathrm{t}}$. If we assume $K = N_{\mathrm{r}}$, this MIMO configuration is equivalent to a downlink MU-MISO system with a BS of N_{t} transmit antennas and K single-antenna users supporting single-stream transmission each. The main goal of the precoder is to map the N_{r} data symbols onto N_{t} transmit antennas. Depending on the precoder design, this mapping achieves specific objectives under particular considerations in order to improve one or several aspect(s) of the system's performance, as we shall detail it later.

Let $d_i[n]$ denote the discrete-time data symbol intended for the ith receive antenna during the nth SP, where $i \in \{1, \ldots, N_{\mathrm{t}}\}$ and $E\{d_i[n]d_i[n]^*\} = 1$. We note that the precoder may have a linear, i.e., the precoder is linear function of the data symbols $\{d_i[n]\}_{i=1}^{N_{\mathrm{r}}}$, or non-linear structure, depending on the objectives and constraints of the underlying precoder design problem. Considering a linear structure, the precoder could be expressed using the precoding

$\mathbf{W} \in \mathbb{C}^{N_\mathrm{t} \times N_\mathrm{r}}$ matrix, which maps a linear combination of the data symbols $\{d_i[n]\}_{i=1}^{N_\mathrm{r}}$ onto N_t transmit antennas. This precoding matrix could be constructed as $\mathbf{W} \triangleq [\mathbf{w}_1 \mathbf{w}_2 \dots \mathbf{w}_{N_\mathrm{r}}]^\mathrm{T}$, with $\mathbf{w}_i \in \mathbb{C}^{N_\mathrm{t} \times 1}$ denoting the precoding vector for the ith data stream intended for the ith receive antenna. In particular, the precoding vector $\mathbf{w}_i$ represents the complex weights intended for the data symbol $d_i[n]$, which alters both the amplitude and the phase of $d_i[n]$ for each antenna element. The precoded signal vector for all receive antennas during the nth SP, $\mathbf{x_d}[n]$, can be expressed as

$$\mathbf{x_d}[n] = \mathbf{W}\mathbf{d}[n] = \sum_{i=1}^{N_\mathrm{r}} \mathbf{w}_i d_i[n], \tag{2.1}$$

where $\mathbf{d}[n] \triangleq [d_1[n] d_2[n] \dots d_{N_\mathrm{r}}[n]]^\mathrm{T}$ collects all receive antennas' data symbols. Assuming a frequency-flat fading channel between the Tx and the Rx, we define $\mathbf{h}_i \in \mathbb{C}^{N_t \times 1}$ as the vector containing the complex coefficients that characterize the propagation environment, i.e., channel gains and phases, between the N_t transmit antennas and the ith receive antenna. Thus, the received complex, baseband, and symbol-sampled signal at the ith receive antenna can be expressed as

$$y_i[n] = \mathbf{h}_i \mathbf{W}\mathbf{d}[n] + z_i[n] = \mathbf{h}_i \sum_{i=1}^{N_\mathrm{r}} \mathbf{w}_i d_i[n] + z_i[n], \quad i \in \{1, \dots, N_\mathrm{r}\}, \tag{2.2}$$

where $z_i[n] \sim \mathcal{CN}(0, \sigma_i^2)$ represents the additive circularly symmetric complex Gaussian noise at the ith receive antenna. Considering all receive antennas, the overall received signal during the nth SP can be written as

$$\mathbf{y}[n] = \mathbf{H}^\mathrm{T} \mathbf{W}\mathbf{d}[n] + \mathbf{z}[n], \tag{2.3}$$

where $\mathbf{z}[n] \triangleq [z_1[n] z_2[n] \dots z_{N_\mathrm{r}}[n]]^\mathrm{T}$, $\mathbf{y}[n] \triangleq [y_1[n] y_2[n] \dots y_{N_\mathrm{r}}[n]]^\mathrm{T}$, and $\mathbf{H} \triangleq [\mathbf{h}_1^\mathrm{T} \mathbf{h}_2^\mathrm{T} \dots \mathbf{h}_{N_\mathrm{r}}^\mathrm{T}]$ represents the channel matrix.

Given the aforementioned systems depicted in Figures 2.1 and 2.2 and $\mathbf{w}_i d_i[n]$ as the transmit signal intended for the ith receive antenna duing the nth SP, the received signal y_i could be decomposed into desired and interference components as

$$y_i[n] = \underbrace{\mathbf{h}_i \mathbf{w}_i d_i[n]}_{\text{desired}} + \underbrace{\mathbf{h}_i \sum_{j \neq i} \mathbf{w}_j d_j[n]}_{\text{ICI/MUI}} + z_i[n], \quad i \in \{1, \ldots, N_\mathrm{r}\}, \tag{2.4}$$

where the ICI/MUI is caused by simultaneous transmission to the remaining antennas other than i. Specifically, we can observe in (2.4) that the transmitted signal to the jth receive antenna interferes with the received signal at the ith antenna, resulting in the received signal $\mathbf{h}_i \mathbf{w}_j d_j[n]$ contributing to the desired signal, in general in an undesired way that may degrade the detection performance. Consequently, the average SINR at the ith antenna, SINR_i, is the ratio between the desired and undesired signal powers received from transmitting the i data stream averaged over the SP, i.e.,

$$\mathrm{SINR}_i \triangleq \frac{P_{\mathrm{D},i}}{P_{\mathrm{I},i} + P_{\mathrm{N},i}}, \quad i \in \{1, \ldots, N_\mathrm{r}\}, \tag{2.5}$$

where $P_{\mathrm{N},i} = \sigma_i^2$ represents the noise power at the ith antenna, and $P_{\mathrm{S},i}$ and $P_{\mathrm{I},i}$ denote the average desired received signal power and the interference power for the ith antenna, respectively, and are defined as

$$\begin{aligned} P_{\mathrm{D},i} &\triangleq E\left\{\mathbf{h}_i \mathbf{w}_i d_i[n] d_i^*[n] \mathbf{w}_i^\mathrm{H} \mathbf{h}_i^\mathrm{H}\right\} \\ &= \mathbf{h}_i \mathbf{w}_i \, E\left\{d_i[n] d_i^*[n]\right\} \, \mathbf{w}_i^\mathrm{H} \mathbf{h}_i^\mathrm{H} \\ &= \mathbf{h}_i \mathbf{w}_i \mathbf{w}_i^\mathrm{H} \mathbf{h}_i^\mathrm{H}, \end{aligned} \tag{2.6}$$

and

$$\begin{aligned} P_{\mathrm{I},i} &\triangleq E\Big\{ \Big\| \mathbf{h}_i \Big(\sum\nolimits_{j \neq i} \mathbf{w}_j d_j[n] \Big) \Big\|^2 \Big\} \\ &= \mathbf{h}_i \Big(\sum\nolimits_{j \neq i} \mathbf{w}_j E\left\{d_j[n] d_j^*[n]\right\} \mathbf{w}_j^\mathrm{H} \Big) \mathbf{h}_i^\mathrm{H} \\ &= \sum_{j \neq i} \mathbf{h}_i \mathbf{w}_j \mathbf{w}_j^\mathrm{H} \mathbf{h}_i^\mathrm{H}, \end{aligned} \tag{2.7}$$

such that the symbols $\{d_i[n]\}_{i=1}^{N_\mathrm{r}}$ are mutually uncorrelated, i.e., $E\{s_i[n] s_j^*[n]\} = 0$ for all $i, j \in \{1, 2, \ldots, N_\mathrm{r}\}, i \neq j$. As a consequence, (2.5) becomes

$$\mathrm{SINR}_i = \frac{\mathbf{h}_i \mathbf{w}_i \mathbf{w}_i^\mathrm{H} \mathbf{h}_i^\mathrm{H}}{\sum_{j \neq i} \mathbf{h}_i \mathbf{w}_j \mathbf{w}_j^\mathrm{H} \mathbf{h}_i^\mathrm{H} + \sigma_i^2}, \quad i \in \{1, \ldots, N_\mathrm{r}\}. \tag{2.8}$$

From (2.8), we can observe that the precoder design directly influences the strength of the ICI/MUI. But in all cases, the ICI/MUI may degrade the system's performance if

not handled properly. Therefore, precoding techniques are of paramount importance for interference management. As mentioned earlier, precoding techniques are employed at the Tx/BS to properly manage the ICI/MUI, e.g., mitigate, eliminate, or convert the interference into a usual signal contribution. As a prerequisite to closed-form based precoding techniques, we start by reviewing cost-function based precoding.

2.1.1 Preliminaries in Block-Level Precoding

We start by reviewing some linear block-level precoding (BLP) schemes. We note that precoder design is the same for both SU-MIMO as well as MU-MISO systems. Assuming a separate design of the Tx and the Rx, i.e., where the precoding at the Tx and the detection at the Rx are not designed jointly, these systems are equivalent from a transmitter perspective if we consider $N_{\mathrm{r}} = K$. In the following, we set N_{r} to denote the number of antennas at the Rx for both systems for simplicity. Thus, we define the instantaneous channel matrix $\mathbf{H} \in \mathbb{C}^{N_{\mathrm{t}} \times N_{\mathrm{t}}}$ between the N_{t} transmit antennas and N_{r} receive antennas. A common assumption in the following precoders' derivations is: $\mathbf{H}$ is assumed to be known at the Tx/BS.

MF Precoder

The MF precoder [JUN05], also known as conjugate beamforming and maximum ratio transmission (MRT) [Lo99], aims to maximine the power of the received signal without any consideration of the ICI/MUI. The corresponding optimization problem can be expressed as

$$\underset{\mathbf{W}}{\text{maximize}} \quad \frac{\left|E\left\{\mathbf{d}^{\mathrm{H}}\mathbf{y}\right\}\right|^2}{E\left\{\|\mathbf{z}\|^2\right\}} \quad \text{subject to} \quad \left(\mathbf{W}E\left\{\mathbf{d}\mathbf{d}^{\mathrm{H}}\right\}\mathbf{W}^{\mathrm{H}}\right) = p, \tag{2.9}$$

where p is a fixed average transmit power. The solution to this optimization problem can simply be obtained as

$$\mathbf{W}_{\mathrm{MF}} = \mathbf{H}^{\mathrm{H}} \tag{2.10}$$

ZF Precoder

The ZF precoder, as the name implies, attempts to fully cancel the power of the interference such that $\mathbf{H}\mathbf{W}_{\mathrm{ZF}} = \mathbf{I}_{N_{\mathrm{t}}}$ [WES08]. It is considered as an evolution of the MF precoder to limit

the adverse effects of ICI/MUI. The interference cancellation property of this precoder comes at the price of a slightly more complex precoder than the MF while lowering the received power. The corresponding precoding matrix is derived by solving the following optimization problem:

$$\underset{\mathbf{W}}{\text{minimize}} \quad E\left\{\|\mathbf{W}\mathbf{d}\|^2\right\} \quad \text{subject to} \quad \mathbf{H}\mathbf{W} = \mathbf{I}. \tag{2.11}$$

The solution to (2.11) is given by

$$\mathbf{W}_{\text{ZF}} = \left(\mathbf{H}^{\text{H}}\mathbf{H}\right)^{-1}\mathbf{H}^{\text{H}} \tag{2.12}$$

We note that $\mathbf{W}_{\text{ZF}}$ is essentially the left pseudo-inverse of the matrix $\mathbf{H}^{\text{H}}$. One common drawback of this precoding approach is its low power efficiency, as it suppresses the interference below the noise floor at the Rx, which is unnecessary. One way of handing this is inefficiency is through regularization, as in the next scheme.

Regularized ZF Precoder

The regularized ZF (RZF) precoding main goal is to maximize the sum of the SINR. Although this approach does not produce an optimal solution, it minimizes the interference while optimizing the received power. It is considered as a precoding approach between the MF and ZF precoders, which can be expressed as a linear combination of MF and ZF precoders [RM16]. Applying the regularized inversion method in [ZO95], the RZF precoding matrix can be expressed as

$$\mathbf{W}_{\text{RZF}} = \left(\mathbf{H}^{\text{H}}\mathbf{H} + \alpha\mathbf{I}_{N_{\text{t}}}\right)^{-1}\mathbf{H}^{\text{H}} \tag{2.13}$$

where α is a regularization parameter whose value is fixed during the transmission, often and $\mathbf{I}_{N_{\text{t}}}$ denotes the $N_{\text{t}} \times N_{\text{t}}$ identity matrix. We note that, for all of the aforementioned linear precoding schemes, in order to respect the power constraint $E[\|\mathbf{W}\|^2] = 1$, we normalize $\mathbf{W}$ as follows: $\widehat{\mathbf{W}} = \frac{\mathbf{W}}{\|\mathbf{W}\|_{\text{F}}}$.

Cost-Function Based Precoding

We revise that in wireless communications, a typical goal is to increase signal power at each receive antenna while reducing the ICI/MUI that is caused by the energy leakages from the transmit antennas. While it is feasible to design the precoder to minimize signal power at a receive antenna, it is challenging to jointly minimize this power and simultaneously minimize the interference leakage. In general, this optimization problem is a non-deterministic polynomial-time (NP)-hard problem [LDL11]. Nevertheless, a simple structure of optimal linear precoding was introduced in [SJU08,BBO14]. In the following, we present two common structure for cost-function based precoding design.

Power Minimization with SINR Constraints: One common approach in designing the precoder is by reducing the transmit power if the available resources allow, while simultaneously constraining the power minimization problem by some SINR requirements, targeted for each receive antenna. In general, the power minimization approach usually a straightforward problem with a simple optimal solution structure [BO99, BBO14]. This optimization problem could be formulated as follows

$$\underset{\mathbf{w}_1,\ldots,\mathbf{w}_{N_\mathrm{r}}}{\text{minimize}} \quad \sum_{i=1}^{N_\mathrm{r}} \|\mathbf{w}_i\|^2 \quad \text{subject to} \quad \mathrm{SINR}_i \geq \gamma_i,\ i \in \{1,\ldots,N_\mathrm{r}\}, \tag{2.14}$$

where γ_i represents the target SINR for the ith receive antenna described in (2.8). We note that the SINR constraints in (2.14) are not convex, however, they could be formulated as convex constraints as follows. Using the SINR expression presented in (2.8), we can express the SINR constraint in (2.14) as

$$\frac{\mathbf{h}_i \mathbf{w}_i \mathbf{w}_i^{\mathrm{H}} \mathbf{h}_i^{\mathrm{H}}}{\sum_{j \neq i} \mathbf{h}_i \mathbf{w}_j \mathbf{w}_j^{\mathrm{H}} \mathbf{h}_i^{\mathrm{H}} + \sigma_i^2} \geq \gamma_i. \tag{2.15}$$

Performing some straightforward algebraic steps on (2.15) lead to the following expression

$$\mathbf{h}_i \left(\sum_{j \neq i} \mathbf{w}_j \mathbf{w}_j^{\mathrm{H}} - \frac{1}{\gamma_i} \mathbf{w}_i \mathbf{w}_i^{\mathrm{H}} \right) \mathbf{h}_i^{\mathrm{H}} + \sigma_i^2 \leq 0, \tag{2.16}$$

which is a convex second-order cone (SOC) constraint. The advantage of having a convex constraint is, when the objective function is also convex, consequently the corresponding optimization problem will be convex as well and could be solved using convex optimization

solvers [BV04].

General Transmit Precoding Optimization: In the case when the transmit power is restricted in the system, the aforementioned power minimization approach could not be used. To address this limitation, another precoding design having the power restrictions as constraint while maximizing some performance metric. Usually, the corresponding cost function is a function of the receive antennas' target SINRs. When the total transmit power is upper bounded by p, the corresponding optimization problem could be formulated as

$$\underset{\mathbf{w}_1,\ldots,\mathbf{w}_{N_\mathrm{r}}}{\text{maximize}} \quad f\big(\mathrm{SINR}_1, \mathrm{SINR}_2, \ldots, \mathrm{SINR}_{N_\mathrm{r}}\big) \quad \text{subject to} \quad \sum_{i=1}^{N_\mathrm{r}} \|\mathbf{w}_i\|^2 \leq p, \tag{2.17}$$

where $f(\cdot)$ is strictly increasing in the target SINR_i for all $i \in \{1, 2, \ldots, N_\mathrm{r}\}$.

Contrary to power minimization problems, the power-constrained precoding design problems of the form of (2.17) are usually difficult to solve, or even NP-hard for some common choices of the objective function $f(\cdot)$, e.g., the sum-rate function given as

$$f\big(\mathrm{SINR}_1, \mathrm{SINR}_2, \ldots, \mathrm{SINR}_{N_\mathrm{r}}\big) = \sum_{i=1}^{N_\mathrm{r}} \log_2(1 + \mathrm{SINR}_i).$$

In general, precoding techniques could be categorized depending on the switching rate, i.e., how often the precoding coefficients are updated, into two categories: BLP and SLP.

BLP considers the ICI/MUI as harmful and should be mitigated [TG06, YL07, DGAA13, SSH04]. In this situation, the precoding is limited to alleviate the interference along the whole frame as it uses only the knowledge of CSI. This results in reducing the average amount of interference in the frame.

2.1.2 Symbol-Level Precoding Survey

For BLP, in both closed-form linear precoding methods [JUN05] and optimization based schemes [BO01], only the the CSI is used by the precoder and the interference is always treated as detrimental. However, the non-linear precoding methods Tomlinson-Harashima precoding (THP) [Tom71, HM72, WFVH04] and vector perturbation (VP) precoding [PHS05, HPS05] exploit both the CSI and the data symbols in the symbol-by-symbol precoding design. Nevertheless, the problem of these schemes is that they are difficult to implement in practical communications systems due to the heavy encoding and decoding process. In [FW03], an

enhanced THP-based precoding method was proposed for MIMO channels using lattice-reduction-aided equalization strategies. To alleviate the computational shortcomings of THP, in [MGM08], convex precoding was proposed using non-discrete alphabets. In [dMM08], convex precoding was introduced for MIMO channels in high dimensions. In the thesis, we focus on the "constructive interference" (CI) [LSK+20] variation of SLP, however, for simplicity, in the following we use SLP instead of CI-based SLP.

In SLP, the interference can be controlled on a symbol-by-symbol basis. This way permits to rotate each interfering signal to be in the correct detection region, thus eliminating the inter-user interference at each symbol slot. Therefore, SLP techniques [ASK+18] manage the ICI/MUI at the price of a higher switching rate at the precoder [MA09, Mas11, ACO17, KYK14, AM17, ACO16, TDCO20]. That is, SLP implies that the interference can be controlled on a symbol basis, i.e., interference is controlled from an instantaneous point of view. Namely, SLP promotes a new way of handing interference where not all interference is harmful. Instead of cancelling or mitigating interference, SLP aims to turn it into an additional source of power that results in power gains at the Rx [MRS+13, ZKM+14]. Thus, constructive interference (CI) is defined as the interference that pushes the received constellation points deeper into their detection regions, leading to an increased received signal power. Therefore, SLP is by default a CI-based precoding.

Early work of SLP focused on CI exploitation through adapting linear precoding methods. In [MA07, MA09], CI and its counterpart destructive interference (DI) were introduced as types of MUI for selective SLP precoding, where CI is preserved while DI is cancelled out through ZF precoding. In [Mas11], instead of eliminating the DI, the SLP precoder aims to align it with the intended data symbols in order to transform it into CI. In the literature, there has been other work that aim to exploit CI using non-linear precoding methods. In [MSR12], a complex scaling-based approach for a single user was introduced to improve interfering signals alignment with their corresponding intended symbols, which was further optimized to minimize the transmit power. In [GM14], the complex scaling was extended for several users.

Recently, the increase of computational power allowed for optimization to be used for the CI-based precoding to improve specific aspects of the system's performance [MZ15b, MZ15a, ACO15b, ACO16]. In [MZ15b, MZ15a], regions where interference is constructive were formally defined for phase-shift keying (PSK) constellations. This definition has relaxed the

previous signal design approach where the interfering signals have to be strictly aligned to the intended symbols, instead designing the signals to lie in the constructive regions, which further improved the performance. This advanced CI approach was called *non-strict phase rotation* in [LM18b], and was thereafter adopted in the SLP precoding designs in [LM16,LM18a,LML+20, LMV+21,LM18c,LMLV19,LM17a]. We note that most of the aforementioned work [MSR12, GM14], [MA07, MA09, Mas11], [MZ15b, MZ15a, ACO15b, ACO16], [ACO14, ACO15c], the precoding schemes were derived only for PSK modulations. However, in [LML+20,LM17a], [ACO15a,ACO17,LM17b] the symbol-scaling concept was introduced where CI exploitation precoding was extended to multi-level modulations, such as quadrature amplitude modulation (QAM). Besides, it was shown in [ASK+18, MA07, MA09, Mas11, MZ15b, MZ15a, ACO15b, ACO16,LM18b,LM16,LM18a,LML+20,LMV+21] that SLP could used to transmit more data streams simultaneously than linear precoding. Generally, in linear precoding, the number of supported data streams is constrained by the total number of transmit antennas N_{t}. Contrarily, SLP enables more data streams than N_{t}. In [DTCO21], SLP was developed for OFDM MIMO systems one-bit digital-to-analog converters (DACs) and analog-to-digital converters (ADCs) at the Tx and the Rx, respectively, in order to reduce the transmit power. For the same system, SLP was designed for constant envelope mode of transmission, where the transmitted signals have constant amplitude regardless of the channel realization or the information symbols [TDCO20]. In [KMCO21], propose an optimized SLP precoder for downlink MU-MIMO systems in the finite block length regime based on discrete constellation rotations. The authors in [HKO20] designed an SLP precoder for MU-MISO systems robust to CSIT imperfections. SLP design for low complexity transmitter architecture for large-scale antenna array systems was investigated in [DGCO19].

Overall, SLP have clear advantages over conventional precoding as it achieves significant performance gains in terms of bit/symbol-error rate (BER/SER) while maintaining a lower transmit power. By exploiting both CSI and data symbols for CI precoding, interference can be transformed into power gains at the Rx. Besides, in SLP the constraints are enforced for each SP, as opposed to BLP where the constraints are met on average for a block of symbols. In addition, SLP reduces the complexity at the Rx. In conventional BLP, the Rx have to estimate the channel and compensate for the phase rotation induced by the channel effect [Mas18]. SLP, on the other hand, the received symbols are placed in the constructive areas of the intended detection regions, thus requiring no phase-compensation at the Rx. As

a result, it eliminates the need for CSI estimation at the latter. This advantage is particularly relevant for mobile users whose computational power is limited.

However, as any other technique, has disadvantages as well. The major downside of SLP is computational complexity, which is higher than BLP. That is, in BLP, the precoder is updated on a block basis, i.e., an ensemble of symbols that is defined by how often the channel changes, in SLP, on the other hand, the precoder is updated for each SP, thus the increase of the computational complexity. Further, as most of the literature on SLP where the precoder design requires solving an optimization problem at a symbol rate, this leads to an extra computational power when compared with BLP that usually supports closed-form solutions and optimized the precoder on a block-level rate. In addition, SLP schemes are designed mostly for un-coded communication systems; the optimality of the performance could not be guaranteed unless SLP precoding is jointly designed with channel coding.

A typical SLP scheme is the constructive interference for sum power minimization (CISPM) SLP scheme, which is designed to exploit the ICI/MUI for power gains. This scheme propels the Rx's received signals deeper into the correct detection region of the desired symbol for each receive antenna. The CISPM precoded signal for the nth SP can be computed as

$$\begin{aligned}\mathbf{x}_{\mathrm{d}}(\mathbf{d},\mathbf{H},\boldsymbol{\gamma},\sigma_z)=&\arg\min_{\mathbf{x}}\ ||\mathbf{x}||^2 \\ &\text{subject to} \\ &\mathrm{Re}\{\mathbf{h}_j^{\mathrm{H}}\mathbf{x}\}\trianglelefteq\sigma_z\sqrt{\gamma_j}\mathrm{Re}\{d_j\},\ \forall j \\ &\mathrm{Im}\{\mathbf{h}_j^{\mathrm{H}}\mathbf{x}\}\trianglelefteq\sigma_z\sqrt{\gamma_j}\mathrm{Im}\{d_j\},\ \forall j,\end{aligned} \tag{2.18}$$

where $\gamma_j \geq 0$ is the target SINR for the jth receive antenna with $\boldsymbol{\gamma} = [\gamma_1 \dots \gamma_{N_{\mathrm{r}}}]^{\mathrm{T}} \in \mathbb{R}^{N_{\mathrm{r}}\times 1}$ representing the target[1] SINR for all Rx's antennas, and the operator $\trianglelefteq$ guarantees that the real/imaginary parts of received signal lie in the same detection region as the data symbols d_k. In the case of quadrature phase-shift-keying (QPSK) constellation, for instance, when $d_k = 1+1j$, the operator $\trianglelefteq$ simplifies to $\geq$ for both constraints. In the case of QPSK constellation, for instance, when $d_j = 1+1j$, the operator $\trianglelefteq$ simplifies to $\geq$ for both constraints.

As depicted in Figure 2.3, the optimization problem in eq. (2.18) takes inputs: the CSIT $\mathbf{H}$, the symbols to transmit $\mathbf{d}$, the target SINR at the N_{r} antennas $\boldsymbol{\gamma}$, and the noise standard deviation σ_z. The objective function's aim is to minimize the transmit power subject to

[1]We note that herein the target SINR is a parameter used in the CI constraints, not the SINR expression.

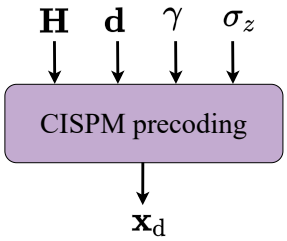


Figure 2.3: Structure of the CISPM precoding scheme.

some constructive interference constraints that are applied to each receive antenna. The constraints' aim is to place the real/imaginary parts of the noiseless received signal at the jth antenna, $\mathbf{h}_j^{\mathrm{H}}\mathbf{x}$, in the detection region corresponding to the real/imaginary parts of the jth intended symbols to transmit. Specifically, with a minimum value of $\sigma_z\sqrt{\gamma_j}$ to guarantee a specific target SINR for each receive antenna. In other words, this scheme propels the Rx's received signals deeper into the correct detection region of the desired symbol.

Thus, the CISPM scheme minimizes the transmit power while guaranteeing a certain target SINR at the Rx through constructive interference constraints. Contrary to the RZF scheme where the precoding matrix $\mathbf{W}_{\mathrm{RZF}}$ is used for the entire coherence time, in the SLP approach, for each SP, the precoding module directly designs the transmitted signal vector $\mathbf{x}_{\mathrm{d}}$ based on both the CSIT $\mathbf{H}$ and the input data symbols $\mathbf{d}$.

Application of SLP to Physical-Layer Security

In the context of PLS, SLP has been introduced as a new way for attaining PLS [LSK+20, ASK+18, MA09]. Although not originally conceived as a PLS method, SLP is more secure than BLP as the precoder is redesigned for each SP. In [LLLS20, FLLL21], secure SLP precoding schemes were proposed in the context of a MISO wiretap channel while considering only a single-antenna Eve. The authors in [KSM+16] proposed to use SLP to enhance the security by increasing the symbol-error rate (SER) at a multi-antenna Eve. In this thesis, we propose several SLP-based PLS schemes, which are presented in Chapters 5 and 6, with varying security, complexity, and power consumption requirements, that increase the BER/SER at Eve.

2.2 MIMO Detection

In the following, we start by revising preliminaries in MIMO detection, we then review the literature for ML for communications as an introduction to ML-based techniques for MIMO detection, which is one of the main topics of this thesis.

2.2.1 Preliminaries in MIMO Detection

MIMO technology has been widely deployed in modern communication systems, however, most receivers rely heavily on accurate CSI in order to detect the data symbols sent by each transmit antenna [TEG04]. Particularly, in the MIMO spatial multiplexing (SM) configuration where several information streams are sent in the same time frequency resource, the Rx has to detect multiple interfering symbol streams transmitted concurrently through a channel subject to random noise and interference [YH15]. In a communication system where the Tx/BS sends different data streams from N_{t} transmit antennas to a multi-antenna Rx, or equivalently to several single antenna users, the Rx observes a superposition of several transmitted data symbols. From the Rx point of view, the problem of detection is to separate the different transmitted symbols. The main challenge at the Rx is to counteract the effect of ICI/MUI, by accurately separating the different data symbols. The goal of *signal detection* is to infer the information data in signal vector $\mathbf{x}$ from the received signal vector $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z}$, where $\mathbf{H}$ is the channel matrix between the transmitter and the receiver and $\mathbf{z}$ is Gaussian noise. ICI and MUI are caused by the harmful effect of the channel matrix $\mathbf{H}$.

We define $\mathbf{y}[n]$ as the received signals at the N_{r} Rx's antennas during the nth SP and $\mathbf{H}$ as the channel matrix between the N_{t} transmit antennas and the Rx with N_{r} antennas, where N_{r} different data streams are sent simultaneously, i.e., one data stream for each receive antenna[2]. We note that the users herein refers to the case of an SU-MIMO system, which is equivalent to N_{r} single-antenna users in a MU-MISO system. For illustration purposes, we consider the SU-MIMO system notation: a Tx equipped with N_{t} transmit antennas and a Rx with N_{r} antennas. Thus, the received signal at all Rx's antennas can be written as

$$\mathbf{y}[n] = \mathbf{H}^{\mathrm{H}}\mathbf{x}_{\mathrm{d}}[n] + \mathbf{z}[n], \tag{2.19}$$

[2]We assume that the number of antennas at the Tx is greater than or equal to the number of antennas at the Rx, $N_{\mathrm{t}} \geq N_{\mathrm{r}}$.

where $\mathbf{x_d}[n] \in \mathbb{X}^{N_t \times 1}$ is the precoded transmitted signal with $\mathbb{X}$ denoting the finite set of constellation points and $\mathbf{z}[n] \in \mathbb{C}^{N_r \times 1}$ collects the independent additive white Gaussian noise (AWGN) components with variance σ_z^2 each at all Rx's antennas. In the following, for simplicity, we drop the index n.

Most MIMO detection techniques have been developed for coherent detection, which require the assumption of having estimated or perfect CSI at the Rx (CSIR). In this setting, the goal of the Rx is to compute the maximum likelihood estimate $\hat{\mathbf{x}}_d$ of $\mathbf{x}_d$, which could be expressed as

$$\hat{\mathbf{x}}_d = \underset{\mathbf{x} \in \mathbb{X}^{N_t}}{\arg\min} \quad \|\mathbf{y} - \mathbf{H}^H \mathbf{x}\|_2. \tag{2.20}$$

The optimization problem in (2.20) is NP-hard due to the final alphabet constraint $\mathbf{x} \in \mathbb{X}^{N_t}$ [PDM17]; the MLD can provide an optimal solution and involves an exhaustive search, which makes it impractical in real time systems. To overcome the computational cost of the MLD, several detectors with varying complexity-performance trade-offs have been proposed in the last decades.

Linear MIMO detectors are based on a linear transformation of the received signal $\mathbf{y}$. There generally known for their low complexity, but incur a considerable loss in performance when compare to the MLD. Their decision statistic could be expressed as

$$\hat{\mathbf{x}}_d = \mathbf{T}\mathbf{y} \tag{2.21}$$

where $\mathbf{T}$ is the linear transformation matrix that could be designed using various criteria.

Neglecting the noise term in (2.19) leads to a system of linear equations. The solution of this linear is known as the ZF detector [YH15] and is expressed as

$$\hat{\mathbf{x}}_{ZF} = (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H . \mathbf{y} \tag{2.22}$$

The linear transformation or filtering matrix $\mathbf{T}$ in (2.21) could also be designed using the MMSE criterion [YH15], by minimizing the mean-square error between the transmitted signal $\mathbf{x}$ and the received signal after using the transformation matrix $\mathbf{T}$. This detector is commonly being referred to as the MMSE detector where $\mathbf{T}$ could be obtained by solving the

following optimization problem

$$\mathbf{T}_{\text{MMSE}} = \arg\min_{\mathbf{T}} \quad E\big(\|\mathbf{x} - \mathbf{T}\mathbf{y}\|_2^2. \tag{2.23}$$

Using the orthogonality principle [Kay93], the MMSE detector's decision statistic may be derived as

$$\hat{\mathbf{x}}_{\text{MMSE}} = \mathbf{T}_{\text{MMSE}}\mathbf{y} \tag{2.24}$$

$$= (\mathbf{H}^{\text{H}}\mathbf{H} + \sigma_z^2\mathbf{I}_{N_{\text{t}}})^{-1}\mathbf{H}^{\text{H}}\mathbf{y} \tag{2.25}$$

We note that herein we assumed perfect CSIR, i.e., the Rx knows perfectly the channel matrix $\mathbf{H}$. The main advantage of linear detection is its low complexity and ease of implementation, which make them attractive for practical implementations. However, they perform substantially worse than the MLD. In the high accuracy regime, we find sphere decoding (SD) algorithms [AEVZ02, GN06], that conduct a search over solutions $\mathbf{x} \in \mathbb{X}^{N_{\text{t}}}$ such that

$$\|\mathbf{y} - \mathbf{H}^{\text{H}}\mathbf{x}\|_2 \leq r. \tag{2.26}$$

where the larger the radius parameter r, the larger set of possible solutions, and consequently the higher the algorithm's complexity. We refer the reader to [YH15, Lar09] for a comprehensive review of MIMO detection schemes that assume the knowledge of CSIR.

On the other hand, few works have been proposed for the case when CSIR is not available, which is commonly being referred to as *blind* MIMO communication [LT02, ACH08, CSH10, LCCK13, SF13]. In [LCCK13, SF13], data detection techniques for blink MIMO communications were proposed, where [LCCK13] was developed specifically for space-shift-keying modulation while [SF13] was developed particularly for phase-shift-keying modulation. Several learning-based blind and semi-blind have been proposed, which will be discussed in the ML-based MIMO detection subsection.

ML-based MIMO Detection

As mentioned in Chapter 1, several ML based approaches have been reported for MIMO detection. In this thesis, we investigate the problem of one-shot symbol detection in a fading

MIMO channel with CSIT, where we leverage the downlink precoded pilot symbols as training data and proposed ML based frameworks for MIMO detection robust to CSIT degradation. In the next chapter, we introduce a generic ML framework for the proposed ML MIMO detection frameworks.

Chapter 3

Generic ML Framework for MIMO Detection

In this chapter, we present an overview of the proposed ML framework for MIMO detection, that is generic and can function with any supervised learning algorithm.

3.1 Communication System Requirements

The idea in this chapter is to model the MIMO detection problem as a supervised learning problem, where the ML model is trained using the received pilot signals and their corresponding pilot symbols. Once the trained ML model is obtained, it can be used to infer the actual data symbols using the received data signals. In this chapter, we propose a MIMO detection framework that is generic and can function with any supervised ML algorithm and/or channel coding scheme.

For this detection framework to function, the underlying communication system must have: 1) precoded pilots in the downlink, to be used as training data, which are commonly available in communication systems for SINR estimation purposes [ETS14, ADM+07], 2) multi-antennas at the Rx, to have enough degrees of freedom to separate the N_{r} data streams, 3) enough computational and memory resources at the Rx to run ML algorithms. In the following, we discuss how this framework could be model as a supervised learning classification problem, and 4) the coherence time length should be long enough in order to have sufficient pilots for the training to converge.

3.2 Supervised Learning and Classification

The main task of supervised learning, which is a subset of ML, is to learn a function that maps an input to an output based on input-output pairs examples. In the context of MIMO detection, the inputs are the received pilot signals. In the ML nomenclature, these inputs are commonly being referred to as *features*. The outputs, on the other hand, are the actual pilot symbols, which as mentioned earlier, are known entities in communication systems' standards. These outputs are referred to as *labels*. Thus, the input-output pairs in our context are the received pilots signals with their corresponding pilot symbols.

Since we consider online training in this thesis, where the training is performed for each coherence time, the training set contains N examples of (received pilot signals)/(actual pilot symbols) pairs while the evaluation set consists of $T - N$ examples of received data signals. We note that, for each example in the the training/evaluation set, the number of features depend directly on the number of antennas at the Rx N_{r}. We also stress that, the nature of the labels define the type of the supervised learning problem, either a *classification* or a *regression* problem. That is, if the labels are of discrete value, the supervised learning problem is a classification problem whereas the regression applies to cases where the labels are continuous values.

Since in our case, the pilot symbols are discrete, i.e., drawn from a modulation constellation, the ML problem is a classification problem. The size of the label-set is determined by the modulation alphabet size, for instance, for QPSK modulation, there are 4 different symbols, thus the label-set is of size 4. Using ML terminology, the label-set size equals the number of *classes* in the classification problem. When the number of classes is 2, we refer to *binary* classification, however, when the number of classes is greater than 2, the classification problem becomes *multi-class*. It is important not to confuse the number of labels with the label-set size. For each example in the training set, the number of labels is equal to the number of simultaneous data streams sent by the Tx, which equals to the number of receive antennas in the MIMO system considered in this thesis. Thus for each antenna at the Rx receives a different pilot symbol. When the label size $\mathbb{L} > 1$, the classification problem becomes a *multi-label* classification (MLC) problem.

Therefore, the nature of the classification problem depends on the label-set size, which defines whether it is a single or multi class, and the number of labels associated with each

example, i.e., one label leads to single-label classification else we refer to multi-label classification. In our setting, the label size depends on the constellation size while the number of labels for each example depends on the number of different pilot/data streams sent by the Tx, which is assumed to be the same as the number of antennas at the Rx in this thesis. Thus, the training set has N examples while the evaluation set contains $T-N$ examples. The number of features is the same in both training and evaluation sets and is directly dependent on the number of antennas at the Rx. Herein, we emphesis that the labels are only part of the training set and the number of labels equals the number of pilot/data streams sent, which also equals the number of antennas at the Rx in this thesis.

Essentially, the training set provides data to the ML algorithm to learn the mapping function between the input features and output labels pairs. This phase is referred to as training phase, which will result in a trained ML model. The evaluation set, on the other hand, is used as an input to the trained ML model to predict the labels, which represent the transmitted data symbols. This prediction occurs during the inference phase. Next, we present in details the training and inference phases of the proposed ML MIMO detection framework. For a thorough explanation of the proposed ML-based detection framework, in the following, we use RZF precoding as an example. However, the proposed decoding frameworks are valid for any precoding technique used at the BS.

3.3 Training Phase

The training phase comprises of three steps: 1) pilots collection, 2) pilots pre-processing, and 3) model fitting. In this phase, for the nth SP, the Tx sends $\mathbf{p}[n] \in \mathbb{C}^{N_\mathrm{r}\times 1}$ pilot symbols with a transmit power of ρ_p. Using RZF precoding, as per eq. (2.13), the transmitted pilot signals during the nth SP, $\mathbf{x}_\mathrm{p}[n] \in \mathbb{C}^{N_\mathrm{t}\times 1}$, can be expressed as follows

$$\mathbf{x}_\mathrm{p}[n] = \sqrt{\rho_\mathrm{p}}\mathbf{W}_\mathrm{RZF}\mathbf{p}[n]. \tag{3.1}$$

We note that the pilot symbols $\mathbf{p}[n]$ are pseudo-random sequences for each receive antenna.

3.3.1 Pilots Collection

In this step, the Rx constructs the training set using the received precoded pilots. We note that precoded pilots are the downlink pilots that are transmitted after precoding that uses

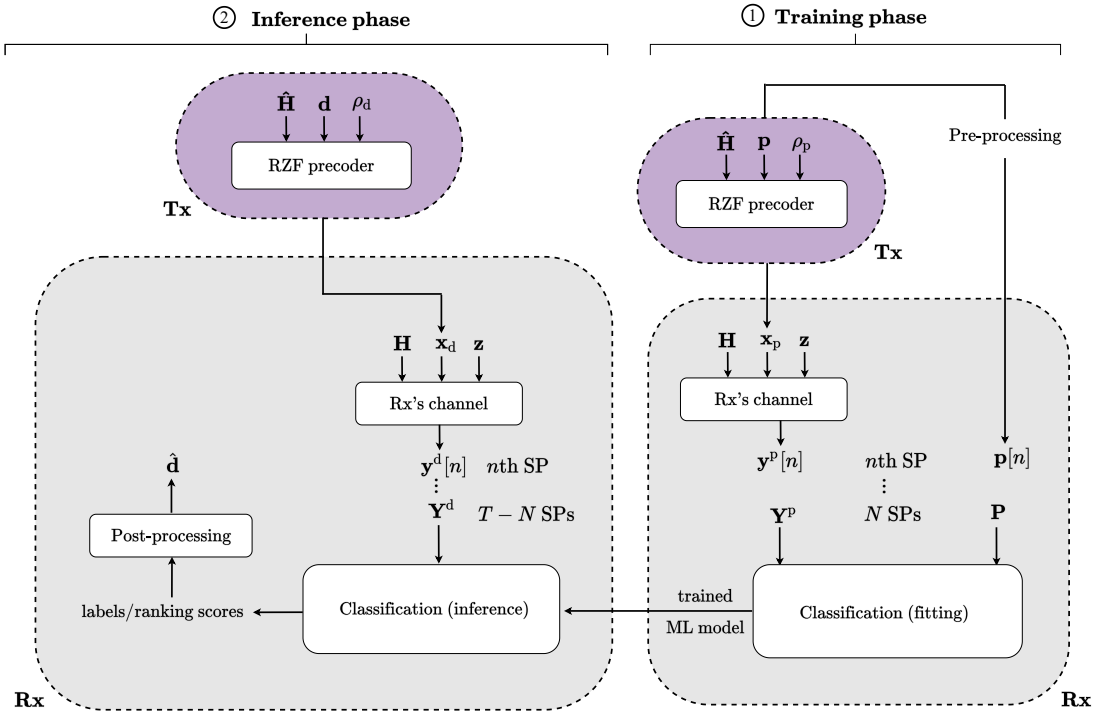


Figure 3.1: Overview of the proposed generic ML detection framework.

the CSIT. The CSIT is obtained through uplink pilots and used for downlink transmission, exploiting the reciprocity in TDD systems. Thus, For the nth SP, the overall received pilot signal at all Rx's antennas, $\mathbf{y}^{\mathrm{p}}[n] \in \mathbb{C}^{N_{\mathrm{r}} \times 1}$, can be written as

$$\mathbf{y}^{\mathrm{p}}[n] = \mathbf{H}\mathbf{x}_{\mathrm{p}}[n] + \mathbf{z}, \tag{3.2}$$

where $\mathbf{x}_{\mathrm{p}} \in \mathbb{C}^{N_{\mathrm{t}} \times 1}$ is the transmitted precoded pilot signal from the N_{t} Tx's antennas during the nth SP. Thus, for N SPs, the received pilot signals are collected in $\mathbf{Y}^{\mathrm{p}} \in \mathbb{C}^{N \times N_{\mathrm{r}}}$ as follows

$$\mathbf{Y}^{\mathrm{p}} = \mathbf{y}^{\mathrm{p}}[n], n \in \{1, \ldots, N\}, \tag{3.3}$$

As depicted in Figure 3.1, for the N SPs of each coherence time, the Rx creates a single set that collects together the received pilot signal $\mathbf{Y}^{\mathrm{p}}$ with the corresponding pilot symbols $\mathbf{P}$, which constitutes the training set.

3.3.2 Pilots Pre-Processing

In this step, pre-processing on the training set is performed for two main purposes: a) for the training set to be supported by the ML algorithms and b) to adapt the formatting of

the training set to the machine learning task, e.g., soft or hard detection. Concerning the suitability aspect of ML, the features in $\mathbf{Y}^{\mathrm{p}}$ are complex-valued and are not suitable for common ML algorithms. This is generally addressed by considering real and imaginary parts separately. Namely, each example, i.e., row, in the features set $\mathbf{Y}^{\mathrm{p}}$ will have N_{r} complex values of the form $a + bi$. Consequently, the new features set $\mathbf{Y}^{\mathrm{p}}$ will contain only the a and b terms of each complex value, i.e., only real numbers. Thus, with this transformation, the features set becomes $\mathbf{Y}^{\mathrm{p}} \in \mathbb{R}^{N \times 2N_{\mathrm{r}}}$.

Regarding the adaptation of the training set format to the ML task, for soft or hard detection, we define two main representations, the bit and the decimal representations. The bit representation, where the pilot symbols in $\mathbf{p}$ are represented in bits, e.g., $p_j \in \{\text{“00”}, \text{“01”}, \text{“11”}, \text{“10”}\}$ in the case of QPSK modulation, is used for soft detection because the soft outputs, i.e., log-likelihood ratios (LLRs), are computed on a per-bit basis. The decimal representation, however, is employed for hard detection, such that $p_j \in \{0, 1, 2, 3\}$ in the case of QPSK modulation.

We note that more sophisticated pre-processing could be applied. For instance: converting the complex value features to polar coordinates, performing computing the mean and standard deviation of all features and add them as features, or adding other relevant entities as features. This sub-field of data analysis is commonly being referred to as *features engineering*.

3.3.3 Model Fitting

As depicted in Figure 3.1, the training set is fed to the fitting module to obtain the trained ML model. Depending on the Rx's intention, the number of labels and the label set size in each example in the training set, $\mathbb{P}$, determine whether the classification problem is single/multi label and binary/multi-class, respectively. For any combination of these, we will have a training set that we denote as $\mathcal{D}_{\mathrm{p}} = \{\mathbf{Y}^{\mathrm{p}}, \mathbf{P}\}$.

The goal of fitting is to generate a well-fitted model to accurately predict new features, that are of similar nature to the ones used in the training phase by minimizing the *bias*. A high bias leads to underfitting, i.e., the model is unable to predict well the labels in the training phase, whereas overfitting manifests when the model predicts very well the training data but poorly the data outside of the training set [Bur19]. The output of this phase is a trained ML model that will be used subsequently in the inference phase.

3.4 Inference Phase

The goal of this phase, as the name implies, it to infer the transmitted symbols from the received data signals during the data transmission phase of the coherence time, that contains $T - N$ SPs. This phase comprises of: a) data collection phase where the evaluation set is constructed using the data received signals, b) data pre-processing where the data is converted to a form that is suitable for ML algorithms and the intended task, and c) signal detection where the transmitted symbols, $\mathbf{d}$, are inferred.

3.4.1 Data Collection

As depicted in Figure 3.1, the overall received data signals at all Rx's antennas during the nth SP, $\mathbf{y}^{\mathrm{d}}[n] \in \mathbb{C}^{N_{\mathrm{r}} \times 1}$, can be expressed as

$$\mathbf{y}^{\mathrm{d}}[n] = \mathbf{H}^{\mathrm{H}}\mathbf{x}_{\mathrm{d}}[n] + \mathbf{z}[n]. \tag{3.4}$$

where $\mathbf{x}_{\mathrm{d}} \in \mathbb{C}^{N_{\mathrm{t}} \times 1}$ is the transmitted precoded data signal from the N_{t} Tx's antennas during the nth SP. As mentioned previously, the Rx collects all the $T - N$ data signals into one, $\mathbf{Y}^{\mathrm{d}}$, that we refer to as the evaluation set.

3.4.2 Data Pre-Processing

In this step, we note that the same pre-processing performed in the training phase must be performed too on the evaluation set, i.e., the obtained mapping function learned how to map the pre-procesed features to the labels, thus the same pre-processing has to be applied to the evaluation set. For instance, the features have to be transformed from complex-values into real-value to be supported by ML algorithms.

3.4.3 ML-based Signal Detection

The last step in the inference phase is to infer the transmitted symbols $\mathbf{d}$, where $\hat{\mathbf{d}}$ denotes an estimate of $\mathbf{d}$. For that, the Rx first feeds the evaluation set $\mathbf{Y}^{\mathrm{d}}$ along with the trained ML model to the classification inference module, which in return will outputs labels as well as *ranking scores*. The labels are of the same nature as the labels in the training set, whereas ranking scores are uncalibrated values that do not constitute probability densities, they signify the confidence level of the inference. For example, if the labels are in the form of bits, that is,

if $s(b_0)$ and $s(b_1)$ represent the scores of the predicted bit's possibilities b_0 and b_1, respectively, and if $s(b_0) \leq s(b_1)$, then there are more chances that the inferred bit is b_0, i.e., the higher the score, the higher the likelihood of the possibility associated with that score. As will be demonstrated in Chapters 4 and 6, these scores could be used to obtain log-likelihood ratios that are often needed for soft decoders when channel coding is employed in the system.

Once the labels and ranking scores are obtained, the final step is to perform post-processing to obtain the estimated transmitted data symbols $\hat{\mathbf{d}}$. This processing includes, but not limited to, demapping or conversion of the labels to symbols (and eventually bits) and hard/soft decoding in case the transmitted bits were coded. The goal of this ML detection framework is to maximize the inference accuracy, i.e., to have the inferred data symbols $\hat{\mathbf{d}}$ as close as possible to the actual sent data symbols $\mathbf{d}$. The BER is often employed in this thesis as a performance metric to assess the prediction accuracy of the proposed ML detection framework.

3.5 Real-Time Considerations

As mentioned earlier, considering a coherence time of T SPs, a total of N SPs are dedicated to pilots while $T - N$ SPs are allocated for informational data. In each SP, the Tx sends N_r different streams simultaneously, thus the Rx has to detect $N_\mathrm{r} \times (T - N)$ data symbols within each coherence time. In the proposed ML MIMO detection framework, in each coherence time, before detecting the informational data, the ML has to be first trained using the $N_\mathrm{r} \times N$ received pilot signals and their corresponding $N_\mathrm{r} \times N$ actual pilot symbols. We note that the SPs dedicated to pilots are not necessarily placed at the beginning of the frame, they are usually interleaved with data SPs. Regardless of their placement in the coherence time, depending on the communication standard used, the Rx knows the pilots and data positioning in each coherence time. Thus, the first step at the Rx is to collect the pilot signals and construct the training set. This part is straightforward and does not require any specific computation, thus could be performed in real time. In the same way, the Rx collects the evaluation set using the received data signals.

Next, the Rx performs some pre-processing to the training and evaluation sets to adapt the data type/format to the ML algorithm and the desired output as well. This step usually involves simple data manipulations and hence does not increase much the complexity of

the proposed ML detector. However, the training and inference phases, detailed in the previous section, incur some computational complexity. Still, as will shall demonstrate in the next chapter, the proposed ML detectors rely on lightweight implementations that make the underlying detectors' computational complexity only marginally higher than closed-form based detectors such as the MMSE detector, making them suitable for real-time systems with online learning.

3.6 Summary

In this chapter, we presented the proposed generic ML detection framework that leverages the availability of downlink precoded pilots as training data. Given the nature of the MIMO detection problem, the ML framework is modeled using as a supervised learning classification problem. This chapter lays the foundation of the MIMO detection framework, which constitutes the core of the ML detectors proposed in Chapter 4 and the eavesdropping attacks in Chapters 5 and 6.

Chapter 4

Data-driven Precoded MIMO Detection Robust to Channel Estimation Errors

In this chapter, we study the problem of one-shot MIMO detection in downlink coded MIMO systems robust to CSIT deterioration. In this context, we investigate the impact of imperfect CSIT on the detection performance at a multi-antenna receiver. We first model the CSIT degradation based on channel estimation errors to investigate its impact on the detection performance at the receiver. To mitigate the effect of CSIT deterioration at the latter, we propose learning-based techniques for hard and soft detection that use downlink precoded pilot symbols as training data. We note that these pilots are originally intended for SINR estimation. We validate the approach by proposing a lightweight implementation that is suitable for online training using several state-of-the-art classifiers. We compare the BER and the runtime complexity of the proposed approaches where we achieve superior detection performance in harsh channel conditions while maintaining low computational requirements. Specifically, numerical results show that severe CSIT degradation impedes the correct detection when a conventional detector is used. However, the proposed learning-based detectors can achieve good detection performance even under severe CSIT deterioration, and can yield 4-8 dB power gain for BER values lower than 10^{-4} when compared to the classic linear MMSE detector.

The rest of the chapter is organized as follows: Section 4.1 introduces the system model,

where we discuss CSIR estimation and downlink transmission. In Section 4.2, we propose our novel learning-based detection frameworks whereas in Section 4.3 we present a lightweight implementation of the proposed frameworks that is suitable for online training. Simulation results are discussed in Section 4.4, followed by the conclusion in Section 4.5.

4.1 System Model

We consider a downlink MIMO system with N_t antennas at the BS and N_r at the user[1]. Herein, the BS sends simultaneously N_r different data streams to the multi-antenna user. The transmission from the BS to the user (downlink) and the transmission from the user to the BS (uplink) share the same frequency resource and operate in TDD. Each coherence interval is divided into three phases: uplink training, downlink payload data transmission, and uplink data transmission, however, in this work, we focus on the downlink data transmission. Hence, for our case, we consider a coherence time with two phases: uplink training and downlink data transmission. In the uplink training phase, the users send orthogonal pilot sequences to the BS, one for every antenna so that the multiple channels can be estimated simultaneously at the BS. The obtained channel estimates at the latter are used to precode the transmit signals in the downlink. For simplicity, in the following, we refer to the BS by "Tx" and the user by "Rx".

The notation and assumptions adopted in this work are as follows:

- We let h_{ij} denote the channel coefficient between the ith Tx's antenna and the jth Rx's antenna. We assume that $h_{ij}, i \in \{1, \ldots, N_\mathrm{t}\}, j \in \{1, \ldots, N_\mathrm{r}\}$, are i.i.d. random variables (RVs). This assumption models the case wherein the Tx and the Rx are distributed over a wide area, and hence, the set of scatterers is likely to be different for each transmit/receive antenna.

- The channel matrix $\mathbf{H} \in \mathbb{C}^{N_\mathrm{t} \times N_\mathrm{r}}$, with entries $h_{ij}, i \in \{1, \ldots, N_\mathrm{t}\}, j \in \{1, \ldots, N_\mathrm{r}\}$, embodies small-scale fading, which is assumed to be static during each coherence time and changes independently from one coherence time to another. We consider Rayleigh channel to reflect a rich scattering environment in an urban setting, where there is no line-of-sight component between the Tx's and the Rx's antennas.

[1] We assume that the number of antennas at the BS is greater than or equal to the number of antennas at the user, $N_\mathrm{t} \geq N_\mathrm{r}$.

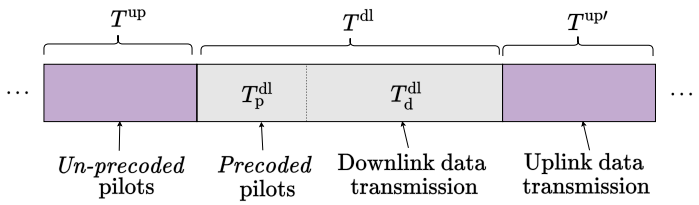


Figure 4.1: Coherence time structure.

- We assume channel reciprocity, i.e., in each coherence time, the channel coefficients in the uplink are the conjugate of those in the downlink. This assumption requires TDD operation and appropriate calibration of the hardware chains.

- We consider a block-fading channel where the channel remains constant for T symbol periods (SPs), i.e., the coherence time length. Since we operate in TDD, as depicted in Figure 4.1, the coherence time is shared between uplink pilots for CSIT estimation, downlink transmission, and uplink transmission, where T^{up}, T^{dl}, and $T^{\mathrm{up}\prime}$ represent their allocated time in SPs, respectively.

- We let d_j denote the data symbol associated with the jth data stream intended for the jth Rx's antenna, which satisfies $E[|d_j|^2] = 1$. In each SP, the Tx sends $\mathbf{d} = [d_1 \dots d_{N_{\mathrm{r}}}]^{\mathrm{T}}$ from the N_{t} antennas, where the symbols $d_j, j \in \{1, \dots, N_{\mathrm{r}}\}$, are mutually independent.

4.1.1 CSIT Estimation through Uplink Training

The Tx computes the channel estimate $\hat{\mathbf{H}} \in \mathbb{C}^{N_{\mathrm{t}} \times N_{\mathrm{r}}}$ using uplink pilots, as depicted in Figure 4.1, that will be used subsequently to define the precoder for the downlink transmission phase that encompasses T^{dl} SPs. The quality of the channel estimate depends mainly on the SNR level, the pilot length, the particular channel estimation technique that is being employed, and the hardware impairments, e.g., uncalibrated hardware chains in this context [Sch08]. In our work, we use the Gauss-Markov formulation to model $\hat{\mathbf{H}}$ in order to reflect these imperfections [RRZ19], ranging from perfect estimation to a completely inaccurate estimation. In this setting, the channel estimate $\hat{\mathbf{H}}$ is obtained using the *actual* channel $\mathbf{H}$ as

$$\hat{\mathbf{H}} = \tau \mathbf{H} + \sqrt{1 - \tau^2}\mathbf{E} \tag{4.1}$$

where the scalar $\tau \in [0,1]$ specify the quality/accuracy of the instantaneous CSI — $\tau = 1$ corresponds to perfect CSIT while $\tau = 0$ indicates that $\hat{\mathbf{H}}$ is completely incorrect and uncorrelated with the actual CSI —, and $\mathbf{E} \in \mathbb{C}^{N_\mathrm{t} \times N_\mathrm{r}}$ represents the random error [TEG04, DB09] where each term follows a circularly symmetric normal distribution $\mathcal{CN}(0,1)$.

4.1.2 Downlink Transmission

Once the channel estimate $\hat{\mathbf{H}}$ is obtained, the Tx uses it to precode pilot and data symbols to transmit to the Rx during T_p^dl and T_d^dl SPs, respectively, as depicted in Figure 4.1.

In each SP, the Tx sends N_r symbols to the Rx, which could be pilot or data symbols. Hence, for the nth SP, the received data signal y_j at the jth antenna of the Rx can be expressed as

$$y_j[n] = \mathbf{h}_j^\mathrm{H} \mathbf{x_d}[n] + z_j[n] \tag{4.2}$$

$\mathbf{h}_j = [h_{1j} \dots h_{N_\mathrm{r}j}]^\mathrm{T} \in \mathbb{C}^{N_\mathrm{t} \times 1}$ is the channel from the N_t transmit antennas to the jth receive antenna, $\mathbf{x_d}[n] \in \mathbb{C}^{N_\mathrm{t} \times 1}$ is the precoded transmitted signal with power $\rho_\mathrm{d} \geq 0$, and $z_j[n]$ is the AWGN at the jth receive antenna with variance σ_z^2.

The above model can be rewritten in a more compact form by collecting the received signal at all Rx's antennas:

$$\mathbf{y}[n] = \mathbf{H}^\mathrm{H} \mathbf{x}_\mathrm{d}[n] + \mathbf{z}[n], \tag{4.3}$$

where $\mathbf{z}[n] \in \mathbb{C}^{N_\mathrm{r} \times 1}$ collects the independent AWGN components of all antennas. Consequently we define the transmit SNR as $\eta = \frac{\rho_\mathrm{d} N_\mathrm{t}}{\sigma_z^2}$.

In our work, we consider the case when the Tx uses the RZF [PHS05], which is a linear BLP, as well as the case of a conventional SLP, i.e., the constructive interference for sum power minimization (CISPM) approach in [ACO15c].

The RZF precoding main goal is to maximize the sum of the SINR. It minimizes the interference signal while optimizing the received power. It is considered as a precoding approach between the MF and ZF precoders. The RZF precoding matrix can be expressed as

$$\mathbf{W} = \hat{\mathbf{H}} \left(\hat{\mathbf{H}}^\mathrm{H} \hat{\mathbf{H}} + \alpha \mathbf{I}_{N_\mathrm{t}} \right)^{-1} \tag{4.4}$$

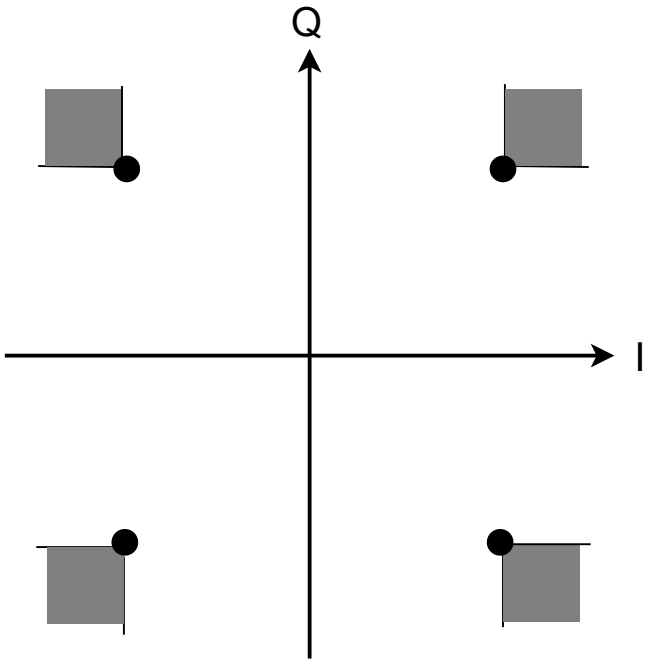


Figure 4.2: CISPM precoding design for QPSK constellation.

where α is a regularization parameter whose value is fixed during the transmission and $\mathbf{I}_{N_\mathrm{t}}$ denotes the $N_\mathrm{t} \times N_\mathrm{t}$ identity matrix.

We note that, in order to respect the power constraint $E[\|\mathbf{W}\mathbf{d}\|^2] = 1$, we normalize $\mathbf{W}$ as follows: $\widehat{\mathbf{W}} = \frac{\mathbf{W}}{\|\mathbf{W}\|_\mathrm{F}}$. Hence, using this notation, the RZF precoded signal for the nth SP can be expressed as

$$\mathbf{x}_\mathrm{d}[n] = \sqrt{\rho_\mathrm{d}}\widehat{\mathbf{W}}\mathbf{d}[n]. \tag{4.5}$$

As for the CISPM SLP scheme, it is designed to exploit the MUI for power gains. This scheme propels the noiseless Rx's received signals deeper into the correct detection region of the desired symbol for each receive antenna. To illustrate, as depicted in Figure 4.2, for a QPSK constellation, the gray shaded regions represent the areas where the noiseless received signal lays after precoding with the CISPM scheme and the black circles are the constellation point placed at the target SINR value. The CISPM precoded signal for the nth SP can be computed as

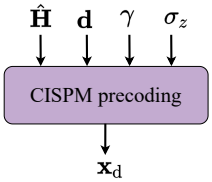


Figure 4.3: Structure of the CISPM precoding scheme.

$$\begin{aligned}\mathbf{x}_\mathrm{d}(\mathbf{d},\hat{\mathbf{H}},\boldsymbol{\gamma},\sigma_z)=&\arg\min_{\mathbf{x}}||\mathbf{x}||^2 \\ &\text{subject to} \\ &\mathrm{Re}\{\hat{\mathbf{h}}_j^\mathrm{H}\mathbf{x}\}\trianglelefteq\sigma_z\sqrt{\gamma_j}\mathrm{Re}\{d_j\},\ \forall j \\ &\mathrm{Im}\{\hat{\mathbf{h}}_j^\mathrm{H}\mathbf{x}\}\trianglelefteq\sigma_z\sqrt{\gamma_j}\mathrm{Im}\{d_j\},\ \forall j,\end{aligned} \tag{4.6}$$

where $\hat{\mathbf{h}}_j\in\mathbb{C}^{N_\mathrm{t}\times 1}$ is the channel estimate from the N_t transmit antennas to the jth receive antenna, $\gamma_j\geq 0$ is the target SINR for the jth receive antenna with $\boldsymbol{\gamma}=[\gamma_1\dots\gamma_{N_\mathrm{r}}]^\mathrm{T}\in\mathbb{R}^{N_\mathrm{r}\times 1}$ representing the target SINR for all Rx's antennas, and the operator $\trianglelefteq$ denotes[2] the correct detection region [ACO17].

As depicted in Figure 4.3, the optimization problem in eq. (4.6) takes inputs: the estimated CSIT $\hat{\mathbf{H}}$, the symbols to transmit $\mathbf{d}$, the target SINR at the N_r antennas $\boldsymbol{\gamma}$, and the noise standard deviation σ_z. The objective function's aim is to minimize the transmit power subject to some constructive interference constraints that are applied to each receive antenna. The constraints' aim is to place the real/imaginary parts of the noiseless received signal at the jth antenna, $\hat{\mathbf{h}}_j^\mathrm{H}\mathbf{x}$, in the detection region corresponding to the real/imaginary parts of the jth intended symbols to transmit. Specifically, with a minimum value of $\sigma_z\sqrt{\gamma_j}$ to guarantee a specific target SINR for each receive antenna. In other words, this scheme propels the Rx's received signals deeper into the correct detection region of the desired symbol.

Thus, the CISPM scheme minimizes the transmit power while guaranteeing a certain target SINR at the Rx through constructive interference constraints. Contrary to the RZF scheme where the precoding matrix $\widehat{\mathbf{W}}$ is used for the entire coherence time, in the SLP approach, for each SP, the precoding module directly designs the transmitted signal vector $\mathbf{x}_\mathrm{d}$ based on both the CSIT $\hat{\mathbf{H}}$ and the input data symbols $\mathbf{d}$. Even though SLP schemes are computed for every SP, an efficient implementation in hardware has been proposed in [KMA+19].

[2]For further detailed information, the reader should refer to [LSK+20, ASK+18].

4.2 Learning-Based MIMO Detection Frameworks

In this section, we propose two ML detection frameworks, where the Rx uses the transmitted precoded pilot symbols as training data to accurately hard/soft decode the transmitted symbols. We note that these precoded pilots are already used in some communication standards [ADM+07] for the purpose of SINR estimation. Thus, we propose to leverage this existing knowledge at the Rx to train our learning-based detectors. Both of these frameworks are comprised of training and inference phases. In the following, we use RZF precoding as an example for a complete explanation. However, the proposed detection frameworks are valid for any precoding scheme. We stress that in our work we propose to use machine learning only at the Rx, whereas we assume conventional precoding where RZF is used as an example for the precoding design. The proposed learning-based frameworks are generic and function with any ML or channel coding algorithm, thus mathematical derivations of these algorithms are omitted in this section.

4.2.1 Learning-Based Framework for the Proposed MIMO Soft Detection Scheme

As pointed out earlier, the detection occurs at the Rx during the downlink part of the coherence time T^{dl}, where the training phase takes place during the $T_{\mathrm{p}}^{\mathrm{dl}}$ SPs followed by the inference phase in the remaining $T_{\mathrm{d}}^{\mathrm{dl}}$ slots. We note that these pilots might be interleaved with data and do not have to be sequential.

As depicted in Figure 4.4, our learning-based soft detector encompasses two steps: 1) training phase, where the ML model is trained using the $T_{\mathrm{p}}^{\mathrm{dl}}$ precoded pilot symbols, 2) inference phase, where probability densities are estimated and employed to calculate the LLRs which are then fed to a soft decoder. Below we detail each phase.

Training Phase

In each SP of $T_{\mathrm{p}}^{\mathrm{dl}}$, the Tx sends N_{r} pilot symbols, $\mathbf{p} = [p_1 \dots p_{N_{\mathrm{r}}}]^{\mathrm{T}}$ satisfying $E[|p_j|^2] = 1$, which are pseudo-random sequences with each symbol corresponding to one Rx antenna. The corresponding received pilot signal at the jth antenna during the nth SP, $y_j^{\mathrm{p}}[n] \in \mathbb{C}$, can be written as

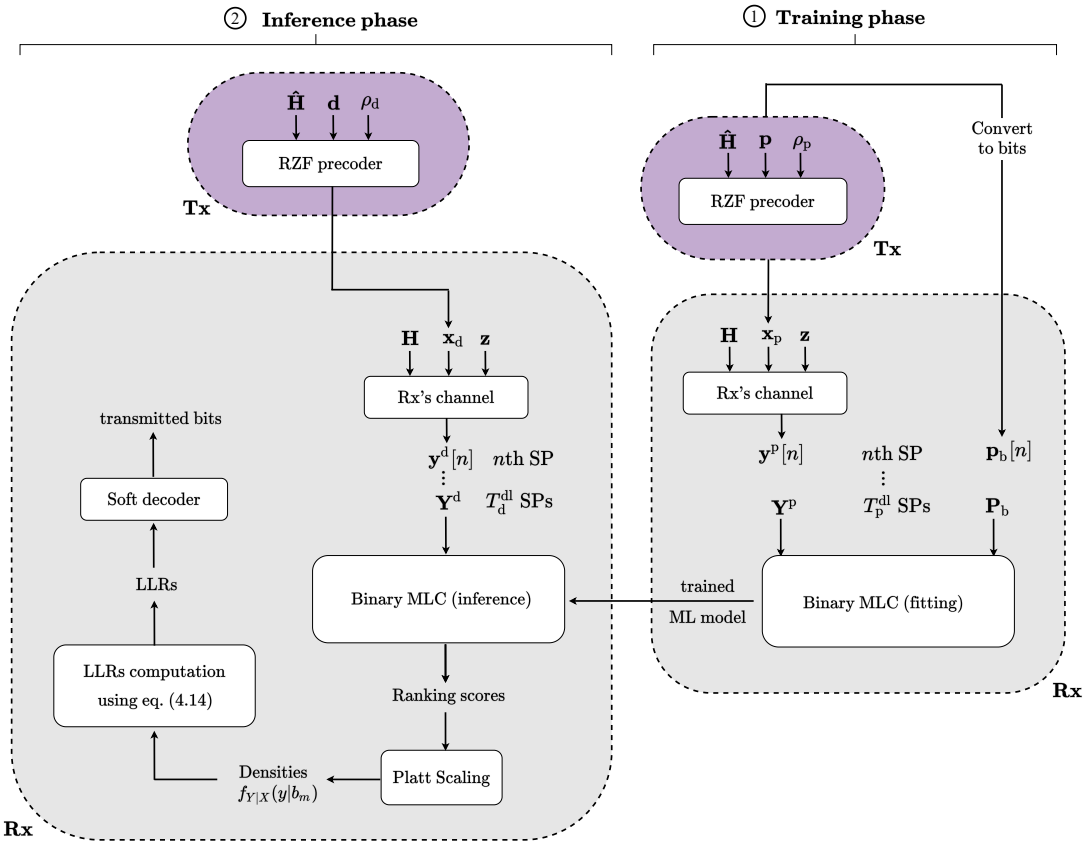


Figure 4.4: Overview of the proposed learning-based soft detector.

$$y_j^{\mathrm{p}}[n] = \mathbf{h}_j^{\mathrm{H}} \mathbf{x}_{\mathbf{p}}[n] + z_j[n], \tag{4.7}$$

where $\mathbf{x}_{\mathbf{p}} \in \mathbb{C}^{N_{\mathrm{t}} \times 1}$ is the RZF precoded transmitted pilot signal with power $\rho_{\mathrm{p}} \geq 0$ such that $\mathbf{x}_{\mathbf{p}} = \sqrt{\rho_{\mathrm{p}}} \widehat{\mathbf{W}} \mathbf{p}$.

Thus, collecting all of the received pilot signals at all antennas, the equivalent received pilot signals during the nth SP, $\mathbf{y}^{\mathrm{p}}[n] \in \mathbb{C}^{N_{\mathrm{r}} \times 1}$, can be expressed as

$$\mathbf{y}^{\mathrm{p}}[n] = \mathbf{H}^{\mathrm{H}} \mathbf{x}_{\mathrm{p}}[n] + \mathbf{z}[n]. \tag{4.8}$$

As depicted in Figure 4.4, for the nth SP, the Rx obtains $\mathbf{y}^{\mathrm{p}}[n]$ and $\mathbf{p}$. The Rx creates a single set that collects together the received pilot signal $\mathbf{y}^{\mathrm{p}}[n]$ and the pilot symbols $\mathbf{p}_{\mathrm{b}}$ corresponding to the bit representation of $\mathbf{p}$, for the LLRs to be computed on a per-bit basis. Specifically, the subscript b in $\mathbf{p}_{\mathrm{b}}$ stands for "bit", where the pilot symbols are represented in bits, e.g., $p_j \in \{\text{“00”}, \text{“01”}, \text{“11”}, \text{“10”}\}$ in the case of QPSK modulation.

As illustrated in Figure 4.4, to perform the training, the Rx first collects all the $T_{\mathrm{p}}^{\mathrm{dl}}$ received pilot signals $\mathbf{y}^{\mathrm{p}}$ with their pilot symbols $\mathbf{p}_{\mathrm{b}}$ in one set, which is referred to as the *training set*, which can be written as

$$\mathcal{D}_{\mathrm{p}}^{\mathrm{s}} = \{\mathbf{Y}^{\mathrm{p}}, \mathbf{P}_{\mathrm{b}}\}, \tag{4.9}$$

where $\mathbf{Y}^{\mathrm{p}} \in \mathbb{C}^{T_{\mathrm{p}}^{\mathrm{dl}} \times N_{\mathrm{r}}}$ are the received pilot symbols at the Rx during $T_{\mathrm{p}}^{\mathrm{dl}}$ SPs and $\mathbf{P}_{\mathrm{b}} \in \mathbb{C}^{T_{\mathrm{p}}^{\mathrm{dl}} \times N_{\mathrm{r}}}$ are the transmitted pilot symbols for $T_{\mathrm{p}}^{\mathrm{dl}}$ SPs. In the ML context, $\mathbf{Y}^{\mathrm{p}}$ and $\mathbf{P}_{\mathrm{b}}$ are referred to as features[3] and labels, respectively. This class of ML problems is named supervised ML. Specifically, this ML problem is considered as a binary MLC problem [TK07] as more than one label is used for each example and the format of the label is binary; each row in the training set $\mathcal{D}_{\mathrm{p}}^{\mathrm{s}}$ represents an example. In our case, an example is any received pilot signal $\mathbf{y}^{\mathrm{p}}$ in a given SP.

Hence, as depicted in Figure 4.4, the training dataset is inputed to the MLC fitting module that will output the so-called "trained ML model". Our goal is to generate a well-fitted model to accurately predict new features, that are of similar nature to the ones used in the training, by minimizing the *bias*. A high bias leads to underfitting, i.e., the model is

[3]We note that the features in $\mathbf{Y}^{\mathrm{p}}$ are complex-valued and are not suitable for common ML algorithms. This is generally addressed by considering real and imaginary parts separately.

unable to predict well the labels in the training, whereas overfitting manifests when the model predicts very well the training data but poorly the data outside of the training set [Bur19].

Inference Phase

This phase, on the other hand, takes place during the $T_{\mathrm{d}}^{\mathrm{dl}}$ SPs of the downlink part of the coherence time. In particular, as depicted in Figure 4.4, in each SP of $T_{\mathrm{d}}^{\mathrm{dl}}$, the Tx sends N_{r} data symbols. The overall received data signals at all antennas of the Rx during the nth SP, $\mathbf{y}^{\mathrm{d}}[n] \in \mathbb{C}^{N_{\mathrm{r}} \times 1}$, can be expressed as

$$\mathbf{y}^{\mathrm{d}}[n] = \mathbf{H}^{\mathrm{H}} \mathbf{x}_{\mathrm{d}}[n] + \mathbf{z}[n]. \tag{4.10}$$

As depicted in Figure 4.4, the Rx first collects the $T_{\mathrm{d}}^{\mathrm{dl}}$ received data signals in one set $\mathcal{D}_{\mathrm{d}} = \mathbf{Y}^{\mathrm{d}} \in \mathbb{C}^{T_{\mathrm{d}}^{\mathrm{dl}} \times N_{\mathrm{r}}}$. In ML terminology, the set $\mathcal{D}_{\mathrm{d}}$ is the test/evaluation dataset.

Generally, the goal of classification is to predict labels. In this context, however, we are not interested in the predicted labels (bits) but rather in the corresponding predicted densities (soft outputs). These soft outputs are used subsequently to compute the (LLRs), which indicate the reliability of the predicted bits.

Before tackling the computation of these LLRs, we first provide an overview of LLRs computation by recalling some fundamental definitions in binary detection [Gal06]. Let X be a binary RV, acting as the correct hypothesis, with possible values $\{b_0, b_1\}$ and a priori probabilities p_0 and p_1. Herein, X models one bit in a transmitted symbol. Let Y be an RV with conditional probability density $f_{Y|X}(y|b_m)$ that is finite and non-zero for all $y \in \mathbb{R}$ and $m \in \{0, 1\}$. In our context, Y models the received signal at the Rx's antenna for a given SP. We note that the conditional densities $f_{Y|X}(y|b_m), m \in \{0, 1\}$, are called *likelihoods*. The marginal density of Y is given by $f_Y(y) = p_0 f_{Y|X}(y|b_0) + p_1 f_{Y|X}(y|b_1)$. Hence, the *a posteriori* probability of X can be expressed as

$$f_{X|Y}(b_m|y) = \frac{p_m f_{Y|X}(y|b_m)}{f_Y(y)}, \tag{4.11}$$

where $m \in \{0, 1\}$. To maximize the probability of correct detection, the maximum a posteriori (MAP) rule can be written as

$$\frac{p_0 f_{Y|X}(y|b_0)}{f_Y(y)} \underset{\tilde{X}=b_1}{\overset{\tilde{X}=b_0}{\gtrless}} \frac{p_1 f_{Y|X}(y|b_1)}{f_Y(y)}, \tag{4.12}$$

where $\tilde{X}$ denotes the decision on the RV X. Rearranging (4.12) and canceling $f_Y(y)$, we obtain the *likelihood ratio*

$$\Lambda(y) = \frac{f_{Y|X}(y|b_0)}{f_{Y|X}(y|b_1)} \underset{\tilde{X}=b_1}{\overset{\tilde{X}=b_0}{\gtrless}} \frac{p_1}{p_0}, \tag{4.13}$$

where the quantity $\frac{p_1}{p_0}$ is called the *threshold* and depends only on the a priori densities. Hence, the log-likelihood ratio $\mathrm{LLR}(y)$ can be expressed as follows:

$$\mathrm{LLR}(y) = \ln\left[\frac{f_{Y|X}(y|b_0)}{f_{Y|X}(y|b_1)}\right]. \tag{4.14}$$

Therefore, in order to compute the LLRs, we need to first calculate the densities in eq. (4.14).

As depicted in Figure 4.4, to obtain the densities in eq. (4.14), we feed the test dataset to the binary MLC inference module along with the trained ML model. However, the binary MLC inference module does not output densities, but rather predicted labels with their ranking scores. These scores are uncalibrated values that do not constitute probability densities; these scores signify the confidence level of the inference. That is, if $s(b_0)$ and $s(b_1)$ represent the scores of the predicted bit's possibilities b_0 and b_1, respectively, and if $s(b_0) \leq s(b_1)$, then $f_{Y|X}(y|b_0) \leq f_{Y|X}(y|b_0)$. Fortunately, there are existing methods to convert these ranking scores into densities [ZE02]. In Section 4.3, we discuss the details of an efficient implementation that estimates densities from ranking scores, which will be used subsequently in the experiments.

Once the likelihoods $f_{Y|X}(y|b_m)$ are obtained, the LLRs can be computed using eq. (4.14), after which the Rx can simply feed the computed LLRs to the soft decoder (e.g., a Viterbi decoder [Vit06]) to obtain the transmitted data.

4.2.2 Learning-Based Framework for the Proposed MIMO Hard Detection Scheme

The proposed learning-based hard MIMO detector is also comprised of two phases: 1) training phase, where the ML model is trained using the precoded pilot symbols as training data; 2)

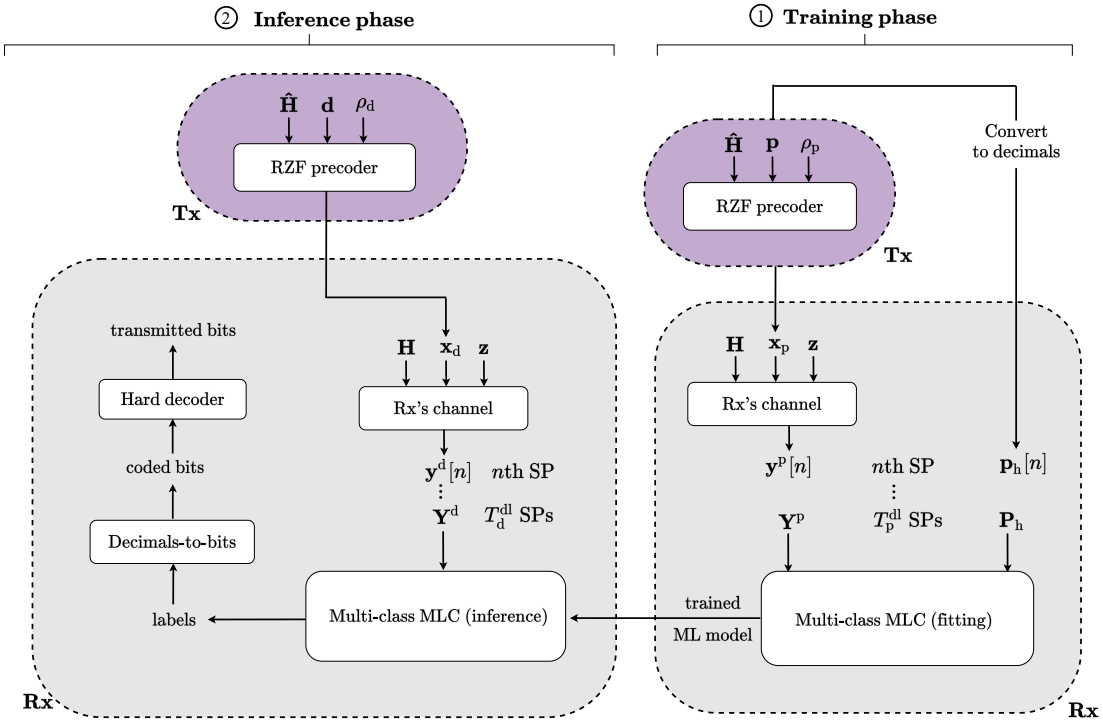


Figure 4.5: Overview of the proposed learning-based hard detector.

inference phase, where the module directly predicts the coded bits (in the form of decimals), which are then mapped into bits to finally be fed to a conventional hard decoder to recover the transmitted bits.

Training Phase

As illustrated in Figure 4.5, during the nth SP within T_p^dl, the Tx sends $\mathbf{p}[n] \in \mathbb{C}^{N_\mathrm{r}\times 1}$ which becomes $\mathbf{x}_\mathrm{p}[n] \in \mathbb{C}^{N_\mathrm{t}\times 1}$ after precoding. Subsequently, the Rx receives at all its antennas, $\mathbf{y}^\mathrm{p}[n] \in \mathbb{C}^{N_\mathrm{r}\times 1}$, as detailed in eq. (4.8).

In addition to the $\mathbf{y}^\mathrm{p}[n]$, the Rx also has the knowledge of the pilot symbols $\mathbf{p}[n]$. We denote the decimal representation of $\mathbf{p}[n]$ by $\mathbf{p}_\mathrm{h}[n]$, where the subscript h stands for "hard"; e.g., $p^j \in \{0, 1, 2, 3\}$ in the case of QPSK modulation.

As illustrated in Figure 4.5, the Rx creates a single set that maps the received signal $\mathbf{y}^\mathrm{p}[n]$ to the corresponding pilot symbols $\mathbf{p}_\mathrm{h}[n]$. Thus, the training set $\mathcal{D}_\mathrm{p}^\mathrm{h}$ is

$$\mathcal{D}_\mathrm{p}^\mathrm{h} = \{\mathbf{Y}^\mathrm{p}, \mathbf{P}_\mathrm{h}\}, \tag{4.15}$$

where $\mathbf{P}_\mathrm{h} \in \mathbb{C}^{T_\mathrm{p}^\mathrm{dl}\times N_\mathrm{r}}$ are the the pilot symbols during T_p^dl SPs. Since the labels are represented

in decimals, considering a modulation order of M, for each label there are M classes. In the case of non-binary modulations, this problem is a multi-class MLC problem [TK07]. Thus, as depicted in Figure 4.5, the training dataset $\mathcal{D}_{\mathrm{p}}^{\mathrm{h}}$ is fed to the multi-class MLC fitting module that in sequence outputs a trained ML model, which will be used thereafter in the inference phase.

Inference Phase

As depicted in Figure 4.5, for the nth SP, the Tx sends the data symbols $\mathbf{d}[n]$ to the Rx in the form of the precoded signal $\mathbf{x}_{\mathrm{d}}[n]$. The corresponding received signals at the Rx's antennas is $\mathbf{y}^{\mathrm{d}}[n]$, as detailed in eq. (4.10).

Therefore, for $T_{\mathrm{d}}^{\mathrm{dl}}$ SPs, the overall received data signal is $\mathbf{Y}^{\mathrm{d}} = \{\mathbf{y}^{\mathrm{d}}[n]\}, n \in \{1, \ldots, T_{\mathrm{d}}^{\mathrm{dl}}\}$, which constitutes the evaluation set. Thus, for the inference, we feed the set $\mathbf{Y}^{\mathrm{d}}$ to the multi-class MLC inference module. Contrary to the proposed soft detection scheme, herein we are interested in predicting the labels, i.e., hard outputs. As depicted in Figure 4.5, to obtain the labels, we feed the evaluation set as well as the previously trained model to the multi-class MLC inference module. We note that the predicted labels are in the form of decimals, i.e., the same nature of the labels used in the training phase. Once the labels are predicted, they will be first mapped into bits to obtain the coded bits, which will then be fed to a hard decoder to finally obtain the transmitted bits.

4.2.3 Scalability of the Proposed ML Detection Frameworks to Multi-User MIMO Systems

Even though this work investigates the case of a single-user (SU) MIMO system, the proposed detection frameworks can be extended to multi-user (MU) MIMO systems [TKCO18]. To illustrate, let us consider an MU-MIMO system with N_{t} antennas at the Tx and K Rxs with N_i antennas each with $N_i K$ total downlink streams such that $N_{\mathrm{t}} > N_i K$. For the nth SP, the received signal at the ith Rx, $\mathbf{y}_i[n] \in \mathbb{C}^{N_i \times 1}$, can be expressed as follows

$$\mathbf{y}_i[n] = \mathbf{H}_i^{\mathrm{H}} \mathbf{x}_{\mathrm{d}}[n] + \mathbf{z}_i[n], \tag{4.16}$$

$\mathbf{H}_i \in \mathbb{C}^{N_{\mathrm{t}} \times N_i}$ is the channel matrix from the Tx to the ith Rx and $\mathbf{z}_i[n] \in \mathbb{C}^{N_i \times 1}$ collects the independent AWGN components of all of the ith Rx's antennas. For each SP, the ith

Rx receives N_i pilot symbols. Therefore, the ith Rx maps each received signal $\mathbf{y}_i[n]$ with the corresponding pilot symbols vector $\mathbf{p}_i \in \mathbb{C}^{N_i \times 1}$. Similarly to the SU scenario, the Rx first collects the received signals with their corresponding pilots for the $T_{\mathrm{p}}^{\mathrm{dl}}$ SPs then construct the training and evaluation datasets using the exact same approach in Sections 4.2.1 and 4.2.2.

All in all, the only difference with respect to the single-user case is the second dimension of the matrices $\mathbf{Y}^{\mathrm{p}}, \mathbf{P}_{\mathrm{b}}, \mathbf{P}_{\mathrm{h}}, \mathbf{Y}^{\mathrm{d}}$ constructing the training and evaluation datasets, having N_i for the ith Rx in the MU use case instead of N_{r} for the single-user scenario.

4.3 Lightweight Implementation of the Proposed Detection Frameworks

In this section, we first discuss our efficient implementation to solve the MLC problem for both of the proposed soft and hard detectors. For the soft detector in particular, we propose a fast algorithm to estimate the probability densities from the ranking scores, which will be used for LLRs computation, and demonstrate its efficacy by plotting the distribution of the predicted densities. Next, we discuss the considered state-of-the-art classifiers used to solve the classification problem for each label. To clarify, the proposed efficient implementation is hierarchical; we propose an implementation for the MLC problem as well as the specific classifiers used for each label. This implementation is further detailed below. We note that the chosen algorithms and approaches for the proposed learning-based detectors have been carefully investigated in terms of suitability for online learning, where Rx optimizes its detector for every coherence time.

We should mention that we did not use deep learning [Den14] in this context despite its high performance in several areas mainly because it requires considerable amount of training data, which is not available in our case. Indeed, in each coherence time, only a limited portion of downlink transmission is dedicated to pilots. As we consider online learning, i.e., the detector is optimized for each coherence time, the pilots of one coherence time could not be combined with the pilots of another coherence time, thus the limitation of the training data. Besides, deep learning typically requires millions of parameters to train, which makes it prohibitively expensive in our application scenario.

In the literature, the proposed methods for solving MLC problems can be grouped into two main categories: a) problem transformation methods and b) algorithm adaptation meth-

ods [TK07]. Transformation methods are those that transform the MLC problem and decompose it into multiple single-label classification (SLC) problem instances, whereas adaptation methods are designed to solve MLC problems directly. Transformation methods are simple and efficient, however, adaptation methods are designed for maximal efficiency, which usually makes them more complex and more computationally intensive compared to transformation methods.

Motivated by the online training approach, we adopt the transformation approach to solve the MLC problem because of its sufficient efficiency and low complexity. In this setting, we consider two transformation methods, binary relevance (BR) [ZLLG18] and classifier chain (CC) [YWF+15]. BR is the most simple and efficient method to solve MLC problems, where multiple SLC are trained independently and their individual outputs are combined to form the multi-label output. Even though this method was designed for binary (two class) labels, as the name implies, it is also implemented for multi-class SLC problems. Despite its popularity and simplicity, its only drawback is that it does not consider label correlations. CC, on the other hand, takes into account the correlation between labels by using the outputs of the previously trained classifiers as features for the subsequent ones in the chain, except for the first classifier. In CC, a chain of SLC is constructed where each classifier, in addition to the related input label, also uses the inferences of other classifiers, thus considering the correlation between labels. We refer to these transformation implementations by "BR" and "CC" accordingly.

The time complexity of BR and CC algorithms is $\mathcal{O}\big(N_{\mathrm{r}} f_c(|\mathbf{Y}^{\mathrm{p}}|, |\mathcal{D}_{\mathrm{p}}^{\mathrm{s|h}}|)\big)$ and $\mathcal{O}\big(N_{\mathrm{r}} f_c(|\mathbf{Y}^{\mathrm{p}}| + N_{\mathrm{r}}, |\mathcal{D}_{\mathrm{p}}^{\mathrm{s|h}}|)\big)$, respectively, where $f_c(\cdot)$ is the complexity of the underlying classifier [RPHF09]. $f_c(\cdot)$ is heavily dependent on the classifier used and the solver used for its implementation, which makes it challenging to obtain a closed-form big $\mathcal{O}$ representation of it. For a quantitative analysis of the time complexity, we measure the total time it takes the algorithm to finish (in milliseconds), which is commonly being refereed to as *runtime* [SDW19]. Consequently, we present the runtime complexity analysis in the following section.

Concerning the proposed soft detector, we adopt an efficient and fast method to accurately estimate the densities $f_{Y|X}(y|b_m)$ from the outputted ranking scores of the MLC module, namely, the Platt scalling approach in [Pla99]. This method is used to transform the uncalibrated scores generated by the classification module into densities. Platt scaling works by fitting a logistic regression model to the classifier's scores. The densities $f_{Y|X}(y|b_m)$

according to the Platt scaling algorithm can be computed as

$$f_{Y|X}(y|b_m) = \frac{1}{1+\exp(Af_y(b_m)+B)}, \tag{4.17}$$

where $f_y(b_m)$ is the classifier ranking score and scalars A and B are the sigmoid parameters [Pla99] learned by the algorithm, which are calculated using a cross-entropy loss function and an internal threefold cross-validation to prevent overfitting.

Regarding the implementation of the SLCs, we have experimented with several state-of-the-art classifiers.[4] For the details about the classifiers' hyper-parameters used in this experiment, we have used the default parameters of the Python modules, where module "scikit-learn" version "0.23.1" was used to implement the "CC" method as well as the "MaxEnt", "SVM", "R_Forest", "KNN", "Decision_Tree", and "Extra_Trees" classifiers, module "scikit-multilearn" version "0.2.0" was used for the "BR" method, module "xgboost" version "1.1.1" was employed for the "XGB" classifier, and module "lightgbm" version "2.3.1" was used for the "LightGBM" classifier implementation. We note that the same implementation is used in the following section.

To evaluate the performance of the employed ML algorithms, we use the *prediction accuracy* metric, which is the ratio of number of correct predictions to the total number of predictions. In Table 4.1, we compare the prediction accuracy of the proposed soft and hard detectors using the proposed implementations, considering both RZF and CISPM precoding schemes with QPSK modulation. For a fair comparison between these precoding schemes, we set the transmit SNR η and the target SINR at the jth receive antenna γ_j to the same value such that all the examined schemes have the same transmit power. For simplicity, we set $\gamma_j = \gamma, j \in \{1,\ldots,N_\mathrm{r}\}$. The parameters used for this experiment are: $\tau = 0.8$ (severe CSIT degradation), $N_\mathrm{t} = N_\mathrm{r} = 8$, $\eta = \gamma_j = 6$ dB such that signal powers of pilot and data signals is the same, $\rho_\mathrm{p} = \rho_\mathrm{d}$, and a frame size of 300 symbols with $T_\mathrm{p}^\mathrm{up} = T_\mathrm{d}^\mathrm{up}$. We note that these results represent the averaged results over 100 different channel realizations. We also note that this accuracy applies before channel decoding, i.e., by comparing the ML predicted labels to the actual coded transmitted symbols.

As observed in Table 4.1, the prediction accuracy values are high despite the severe CSIT degradation ($\tau = 0.8$) and the relatively small SNR $\eta = 6$ dB employed. We also observe,

[4] We note that we did not experiment with neural-network based classifiers because of their relative high complexity and the scarcity of the training data.

Table 4.1: Prediction accuracy of the proposed learning-based detectors with several state-of-the-art classifiers when using CIPSM and RZF precoding schemes.

Classifiers	**CISPM**				**RZF**			
	CC		**BR**		**CC**		**BR**	
	Soft	**Hard**	**Soft**	**Hard**	**Soft**	**Hard**	**Soft**	**Hard**
MaxEnt	0.9439	0.9322	0.9398	0.9341	0.9096	0.8967	0.9058	0.8995
SVM	0.9015	0.8800	0.9011	0.8806	0.8792	0.8609	0.8793	0.8642
R_Forest	0.8419	0.8305	0.8440	0.8313	0.8337	0.8194	0.8376	0.8234
KNN	0.7874	0.7674	0.7888	0.7719	0.7712	0.7428	0.7754	0.7557
Decision_Tree	0.7579	0.7334	0.7582	0.7364	0.7410	0.7169	0.7427	0.7192
Extra_Trees	0.8207	0.8212	0.8577	0.8393	0.8116	0.8108	0.8467	0.8270
LightGBM	0.8667	0.8452	0.8673	0.8464	0.8500	0.8309	0.8495	0.8326
XGB	0.8611	0.8424	0.8611	0.8434	0.8450	0.8279	0.8456	0.8285

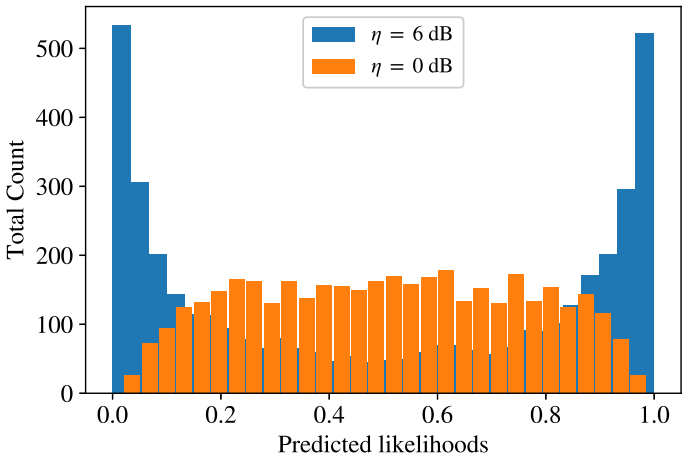


Figure 4.6: Histograms of the estimated likelihoods for the ML - Soft scheme using RZF precoding with $N_{\mathrm{t}} = 15$, $N_{\mathrm{r}} = 8$, $\eta \in \{0, 6\}$ dB, a frame size of 2000, and QPSK modulation.

for both of the precoding schemes, the CC approach achieves better results when using the Soft scheme, however, the BR approach slightly outperforms the CC approach when using the Hard scheme. Furthermore, MaxEnt classifier achieves the highest prediction accuracy amongst all classifiers regardless of the precoding scheme employed and the detector used. Therefore, in the numerical results, to achieve the highest detection performance, we adopt the CC approach for the soft detector and the BR method for the hard one, where both of them employ the MaxEnt classifier to solve the SLC problem for each label. In the numerical results, we refer to these implementations by "ML - Soft" for the soft detection scheme and "ML - Hard" for the hard detection one.

With regards to the employed likelihoods estimation method for the "ML - Soft" scheme, Figure 4.6 depicts the distribution of the estimated likelihoods $f_{Y|X}(y|b_m)$ in the presence of

Table 4.2: Channel coding parameters used for the simulations

Parameters	**Values**
Code rate	½
Decoder decision technique	Hard, Soft
Number of frames	100
Trace-back length	96
Constraint length	9

severe CSIT degradation ($\tau = 0.8$). The parameters used for this simulation are: $N_{\mathrm{t}} = 15$, $N_{\mathrm{r}} = 8$, $\eta \in \{0, 6\}$ dB, a frame size of 2000 with QPSK modulation and RZF precoding. We stress that a predicted probability of 0.5 indicates that the predictor is fully unsure of the predicted bit, on the other hand, a value close to 1 means the predictor is very confident that the predicted bit is a 1 whereas a probability close to 0 implies predictor is confident it is a 0. For a transmit SNR of 0 dB, we observe that the predicted likelihoods are spread around 0.5, which indicates a poor prediction performance. However, for $\eta = 6$ dB, the prediction accuracy is high as evidenced by the likelihoods values distributed mostly around values 0 and 1, thus demonstrating the efficacy of the adopted likelihood estimation method.

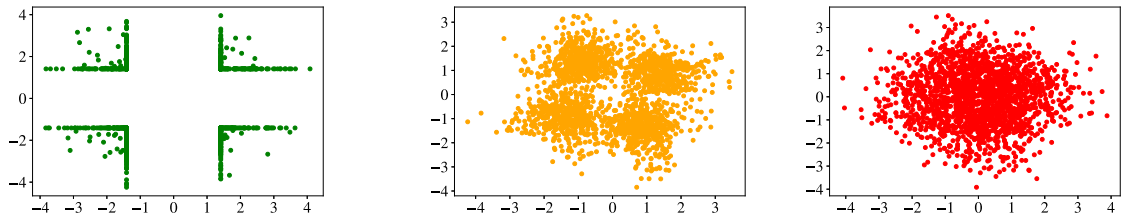
4.4 Numerical Results

Herein, we demonstrate the performance of the proposed detection frameworks using Monte Carlo simulations. We consider a Rayleigh flat-fading MIMO system with $N_{\mathrm{t}} = 15$ and $N_{\mathrm{r}} = 8$, with QPSK constellation with Gray mapping and channel coding, in which we use convolutional coding [JZ15] and Viterbi decoding [Vit06] with the parameters in Table 4.2 and the Python module "scikit-commpy" version "0.5.0". Unless otherwise specified, we use a frame size of 300 symbols, $T_{\mathrm{p}}^{\mathrm{up}} = T_{\mathrm{d}}^{\mathrm{up}}$, $\rho_{\mathrm{p}} = \rho_{\mathrm{d}}$, and $\sigma_z^2 = \alpha = 1$. Further, we consider CISPM and RZF precoding.

To our knowledge, we do not have direct competitors that addressed the same problem that we are investigating here. However, we compare our results with a *conventional* Rx that directly detects the received signals without any processing, since in the considered system, the Tx uses precoding to mitigate the channel effect. Subsequently, we refer to "Conv - Soft" to indicate soft decoding and "Conv - Hard" for the hard decoding. We also compare with another type of Rx that uses the downlink precoded pilot signals to estimate the effective

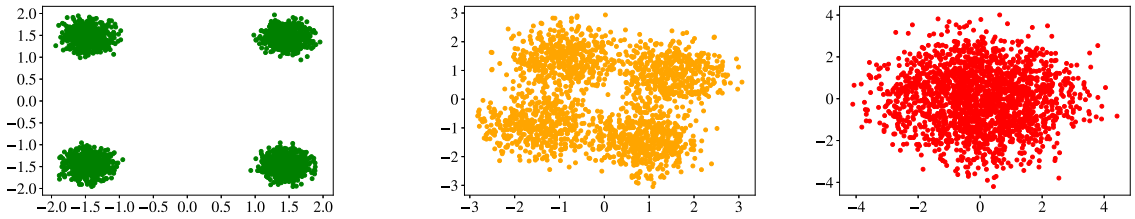
channel and equalize the receive data signals accordingly. Specifically, for this benchmark Rx, we adopt the least-squares (LS) method for the effective channel estimation and the linear MMSE detector that applies the SNR-regularized channel's pseudo-inverse and rounds the output to the nearest constellation point. In the following, for this Rx, we refer to "MMSE - Soft" to indicate soft decoding and "MMSE - Hard" for the hard decoding implementation.

In this work we consider several severity levels of CSIT degradation by varying the parameter τ in eq. (4.1). We note that a value of $\tau = 0.99$ identifies an optimistic channel imperfection, where even with such a small error in the CSI estimation, the degradation is notable [RRZ19]. In this work, we consider three scenarios of CSIT imperfections: $\tau = 1$ for perfect CSIT, $\tau = 0.9$ to indicate a moderately degraded CSIT estimate, and $\tau = 0.8$ to reflect severe CSIT degradation. We first analyze the noiseless received signal in each of these scenarios considering both RZF and CISPM precoding schemes.



(a) CISPM ($\tau = 1$) (b) CISPM ($\tau = 0.9$) (c) CISPM ($\tau = 0.8$)

Figure 4.7: CISPM precoding - Noiseless received signal an Rx's antenna.



(a) RZF ($\tau = 1$) (b) RZF ($\tau = 0.9$) (c) RZF ($\tau = 0.8$)

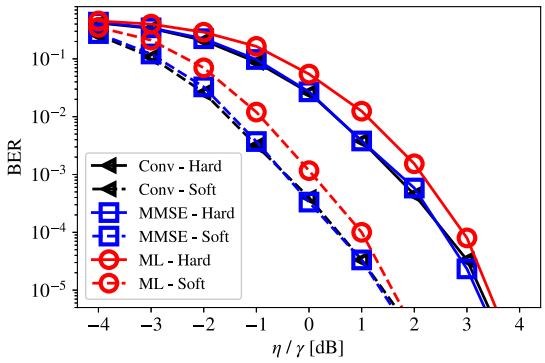
Figure 4.8: RZF precoding - Noiseless received signal an Rx's antenna.

Figures 4.7 and 4.8 depicts the noiseless received signal at an Rx's antenna when using CISPM and RZF precoding, respectively, for different values of τ. The parameters used for this simulation are: $N_{\mathrm{t}} = 15$, $N_{\mathrm{r}} = 8$, and $\eta = \gamma = 6$ dB. We start by evaluating the case when the Tx uses CISPM precoding. Figure 4.7a plots the scatterplot of the noiseless received

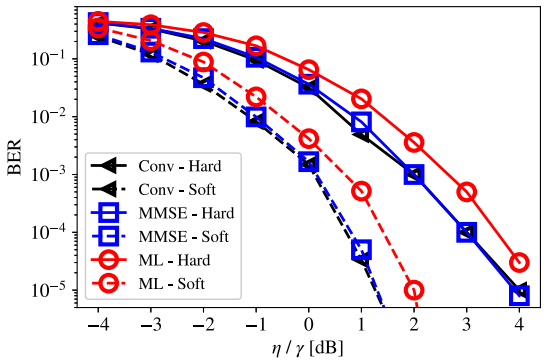
signal in the case of perfect CSIT ($\tau = 1$). We can clearly see that the received constellation points are positioned in the corresponding detection regions; this scheme is guaranteeing a minimum target SINR value of $\gamma = 6$ dB for some transmitted symbol while propelling the rest deeper into the detection region. Therefore, we can conclude that the Tx was fully able to mitigate the ICI effect when using perfect CSIT. In the case of a degraded CSIT ($\tau = 0.9$), as depicted in Figure 4.7b, we can observe that the constellation points that were mostly lying at the edge of the guaranteed SINR in Figure 4.7a got deviated in all directions, which is evidenced by the elliptic/circular shape cloud, due to the CSIT imperfections. However, in the case of severe CSIT imperfections ($\tau = 0.8$), as depicted in Figure 4.7c, we can no longer straightforwardly map the received constellation points to the QPSK detection regions. This is due to the ICI effect.

Similarly, when the Tx uses RZF precoding, we observe a similar phenomenon. Particularly, in the case of perfect CSIT ($\tau = 1$), as depicted in Figure 4.8a, the noiseless received constellation also exhibits a clear separation between the 4 QPSK symbols. As opposed to ZF precoding, RZF does not fully cancel out the ICI effect, we see the constellation points forming a circular cloud in each detection region, where the cloud-effect is caused by the ICI effect, thus, this is how RZF offers a trade-off between ZF and MF precoding. Figure 4.8b shows the degraded CSIT case, where we can observe that the CSIT imperfections led to deviations of the constellation points, causing the 4 "clouds" to get bigger and where their edge is near the detection regions. This means after noise adds up, some of the constellation points will cross into the opposite regions leading to detection errors. However, in the case of severe CSIT degradation ($\tau = 0.8$) in Figure 4.8c, the "clouds" are fully merged and centered at coordinates $(0, 0)$. The Tx in this case is not able to mitigate the ICI effect.

Figure 4.9 depicts the BER as a function of η/γ [dB] in low SNR regime in the case of perfect CSIT ($\tau = 1$). Figure 4.9a plots the BER as a function of η/γ. As expected, the higher the η/γ, the lower the BER, thus the better the detection performance. Conventional detectors outperform the proposed learning-based detectors, where both converge as η/γ increases. This is due to the fact that learning-based schemes loose some performance because of the training and the inference phases. The MMSE detector, however, achieves the same detection performance as the conventional detector, as the Tx have already mitigated all the ICI effect possible. As expected, for all detectors, soft decoding always outperforms hard decoding. The BER performance when using RZF precoding is depicted in Figure 4.9b, where



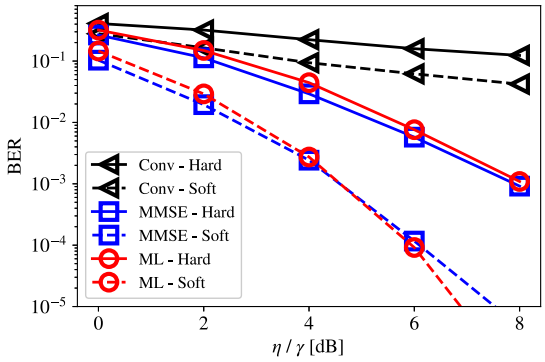
(a) CISPM precoding ($\tau = 1$)



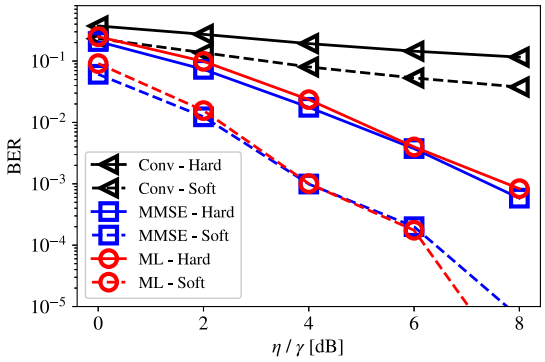
(b) RZF precoding ($\tau = 1$)

Figure 4.9: BER vs. η/γ [dB], with $\tau = 1$ (perfect CSIT).

the same observations apply to the RZF case as well. Overall, we note that all detectors can achieve very low BER with very low η/γ, which indicates the effectiveness of all detectors in an unrealistic perfect CSI scenario, conventional, MMSE, and the proposed ones.



(a) CISPM precoding ($\tau = 0.9$)

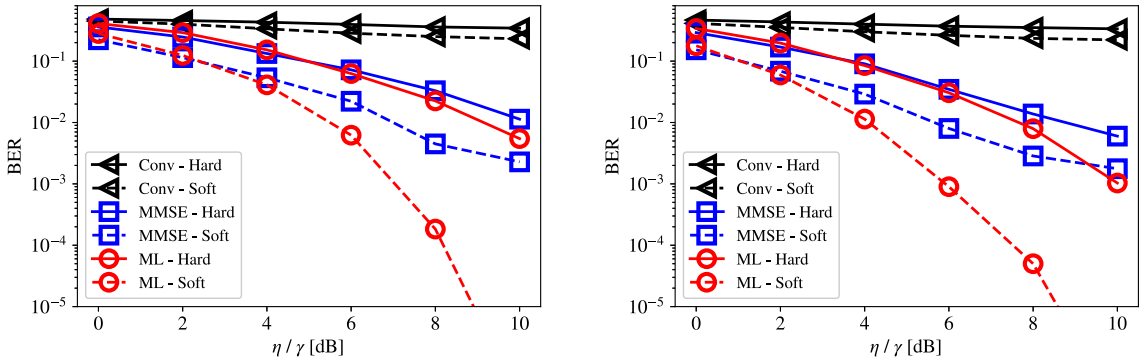


(b) RZF precoding ($\tau = 0.9$)

Figure 4.10: BER vs. η/γ [dB], with $\tau = 0.9$ (degraded CSIT).

Figure 4.10 plots the BER as a function of η/γ [dB] in the case of a moderately degraded CSIT ($\tau = 0.9$). In particular, Figure 4.10a plots the BER as a function of η/γ in the case of CISPM precoding. Similarly, for all detectors, the higher the η/γ, the lower the BER. In particular, the conventional detector's performance flattens in high SNR, which is due to the ICI effect as depicted in Figures 4.7 and 4.8. Nonetheless, MMSE and the proposed detectors' BER decreases linearly with η/γ with the proposed ML detectors outperforming MMSE ones in high SNR. And as expected, for all detectors, soft decoding always outperforms hard decoding. When the Tx uses RZF precoding, as depicted in Figure 4.10b, we observe a

similar behavior. In particular, the proposed ML detectors outperform the MMSE detector as η/γ increase. This implies that the ML approach can leverage better the knowledge of the pilot symbols η/γ, contrary to the MMSE detector where the detection performance that improves linearly with the increase of η/γ.



(a) CISPM precoding ($\tau = 0.8$) (b) RZF precoding ($\tau = 0.8$)

Figure 4.11: BER vs. η/γ [dB], with $\tau = 0.8$ (severe CSIT degradation).

Figure 4.11 plots the BER as a function of η/γ [dB] in the case of severe CSIT degradation ($\tau = 0.8$). In particular, Figure 4.11a plots the BER as a function of η/γ in the case of CISPM precoding. As expected, the BER when using the conventional detector is high, even in high SNR regime. However, the BER gets slightly lower as η/γ increases, but flattens out in high SNR. The MMSE detector's BER, on the other hand, improves linearly with η/γ but attaining a limited detection performance in high SNR. Nonetheless, when using the proposed learning-based detectors, the obtained BER drastically decreases as η/γ increases. This is due to the learning aspect of the proposed ML-based detectors; even in high ICI effect (induced by severe CSIT degradation), the ML detectors learns from the sent precoded pilots the input-output relationships and use it effectively in the inference. When the Tx uses RZF precoding, similar to Figure 4.11a, for all detection schemes, the higher the η/γ, the lower the BER, with the proposed learning-based detectors immensely outperforming the conventional and MMSE detectors and soft decoding achieving better detection performance than hard decoding. In addition, severe CSIT degradation impede the correct detection when using conventional detectors even in high SNR regime. Nonetheless, the proposed learning-based detectors can achieve under 10^{-4} BER values with η/γ as low as 8 dB. Overall, we observe that RZF precoding leads to better detection performance than CISPM precoding.

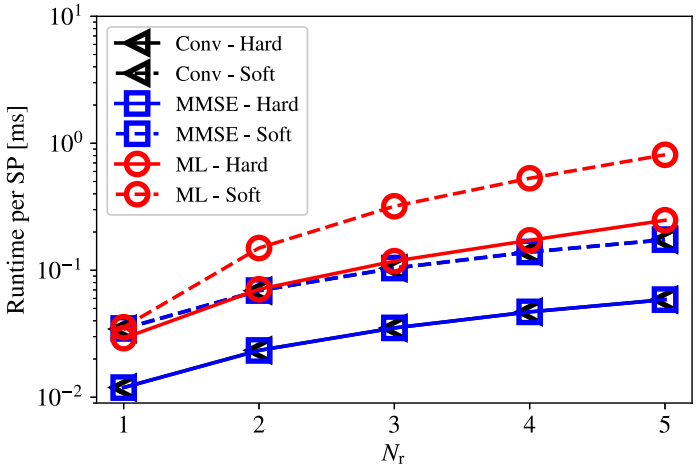


Figure 4.12: Runtime per SP [ms] vs. N_r using RZF precoding and $\eta = 6$ dB.

To quantitatively evaluate the time complexity of the different detectors, in Figure 4.12 we plot the runtime per SP as a function of N_r with $\eta = 6$ dB and RZF precoding. As expected, the higher the number of antennas at the Rx, N_r, the higher the runtime, with soft detection consuming more computation time than hard detection for all schemes. We also observe that "Conv" and "MMSE" detectors exhibit a comparable runtime, as both are based on closed-form implementations. The runtime of the "ML - Hard" scheme is higher than the "Conv/MMSE" hard detection ones but shows a similar performance to "Conv/MMSE" soft detection methods. For the "ML - Soft" scheme, when $N_r = 1$, it achieves the same runtime as the soft implementation of 'Conv/MMSE" schemes, whereas for $N_r > 1$, the runtime of the "ML - Soft" is higher than the one for the soft implementation of "Conv/MMSE" schemes. This difference is due to the "CC" implementation of the "ML - Soft" scheme, where the training and inference phases of the SLC modules are performed sequentially for each received antenna/stream. Overall, the proposed ML detectors consumes marginally higher runtime than the closed-form based implementations.

We conclude this section by summarizing the insights from the numerical results as follows:

- Soft decoding scheme always outperform hard decoding, i.e., soft values provide extra information to the decoder that allow for better restitution of the original data.
- In the case of perfect CSIT, RZF and CISPM precoding schemes are almost equivalent in performance.

- Conventional and MMSE detectors outperform the proposed learning based detectors in perfect CSIT, however, the former detectors are vulnerable to CSIT imperfections.
- The proposed learning-based detectors are much more robust to CSIT imperfections, thanks to their learning aspect that exploits the availability of the precoded SINR pilots in the downlink.
- Leveraging these precoded pilot symbols by using a classic detector like MMSE leads to limited BER improvement with SNR increase, as opposed to drastic BER improvement when using the proposed learning-based detectors.
- Overall, the proposed learning-based detectors achieve remarkable detection performance in severe channel conditions while having low computational complexity.

4.5 Summary

In this paper, we studied the problem of one-shot ML-based MIMO detection robust to CSIT deterioration. We investigated the impact of CSIT imperfections on coded MIMO detection in systems with precoding and without explicit CSIR knowledge. We modeled the CSIT imperfections using the Markov-Gauss formulation [RRZ19] to reflect the degradation due to channel estimation errors. In this setting, we proposed soft and hard learning-based detection frameworks that leverage the availability of downlink precoded pilots, originally intended for SINR estimation, as training data. Moreover, we proposed a lightweight implementation by using fast and efficient ML algorithms and methods to support online training. Numerical results showed that CSIT imperfections inhibits correct detection when using a conventional MIMO detector. We showed that even when a conventional Rx exploits the downlink precoded pilots to estimate the effective channel and uses the MMSE detector to compensate for ICI, the resulting performance is not good under severe CSIT degradation. However, the proposed learning-based detectors are substantially more robust to CSIT degradation, where the proposed ML scheme can achieve 4-8 dB power gain for a BER value under 10^{-4} when compared to the MMSE receiver under severe CSIT degradation, while retaining low computational complexity.

Chapter 5

Learning-Assisted Eavesdropping and Symbol-Level Precoding Countermeasures for Downlink MU-MISO Systems

In this chapter, we address the problem of one-shot MIMO detection robust to inaccurate CSIT at a multi-antenna Eve in an MU-MISO system without channel coding. Particularly, we introduce an ML based detection attack, where an Eve is able to learn the symbol detection function based on precoded pilots. With this ability, an Eve can correctly detect symbols with a high probability. To counteract this attack, we propose a novel SLP scheme that enhances PLS while guaranteeing a constructive interference effect at the intended users. Contrary to conventional SLP schemes, the proposed scheme is robust to the ML-based attack. In particular, the proposed scheme enhances security by designing Eve's received signal to lie at the boundaries of the detection regions. This distinct design causes Eve's detection decisions to be based almost purely on noise. The proposed countermeasure is then extended to account for multi-antennas at the Eve and also for multi-level modulation schemes. In the numerical results, we validate both the detection attack and the countermeasures and show that this gain in security can be achieved at the expense of only a small additional power consumption at the transmitter, and more importantly, these benefits are obtained without affecting the performance at the intended user.

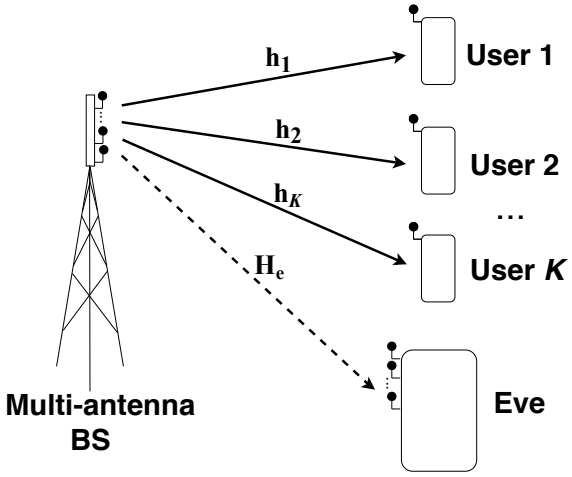


Figure 5.1: Downlink MU-MISO system comprised of: a BS with N_{t} antennas, K single-antenna users, and one Eve with M antennas.

The rest of this chapter is organized as follows: Section 5.1 introduces the system model. In Section 5.2, we present the proposed ML-based attack, whereas in Section 5.3 we propose novel PLS schemes as a countermeasures. Simulation results are presented in Section 5.4 and Section 5.5 concludes this chapter.

5.1 System Model

As depicted in Figure 5.1, we consider a single cell MU-MISO downlink system, where the BS is equipped with N_{t} transmit antennas serving K single-antenna users, with $K \leq N_{\mathrm{t}}$, and one multi-antenna eavesdropper with M antennas. We assume a block fading channel $\mathbf{h}_j \in \mathbb{C}^{1\times N_{\mathrm{t}}}$ between the transmit BS antennas and the j-th user. The received signal at the j-th user can be expressed as

$$y_j[n] = \mathbf{h}_j \mathbf{x}[n] + z_j[n] \tag{5.1}$$

where $y_j[n] \in \mathbb{C}$ is the received signal at the j-th user in the symbol slot n, $\mathbf{x}[n] \in \mathbb{C}^{N_{\mathrm{t}}\times 1}$ is the transmitted vector from the N_{t} transmit antennas, and $z_j[n] \in \mathbb{C}$ is the AWGN at the j-th user with variance σ_z^2.

The above model can be rewritten in a matrix form by collecting the received signal at all users in vector $\mathbf{y}[n] \in \mathbb{C}^{K\times 1}$ as

$$\mathbf{y}[n] = \mathbf{H}^{\mathrm{T}}\mathbf{x}[n] + \mathbf{z}[n] \tag{5.2}$$

where $\mathbf{H} = [\mathbf{h}_1^{\mathrm{T}} \dots \mathbf{h}_K^{\mathrm{T}}] \in \mathbb{C}^{N_{\mathrm{t}} \times K}$ represents the system channel matrix and $\mathbf{z}[n] \in \mathbb{C}^{K \times 1}$ gathers the independent AWGN components of all users, with a variance of σ_z^2 each. We note that $\mathbf{H}$ is assumed to be known at the BS through pilot-assisted channel estimation [SACO18].

Similarly, the received signal at an Eve with M antennas, $\mathbf{y}_{\mathrm{e}}[n] \in \mathbb{C}^{M \times 1}$, can be expressed as follows

$$\mathbf{y}_{\mathrm{e}}[n] = \mathbf{H}_{\mathrm{e}}^{\mathrm{T}}\mathbf{x}[n] + \mathbf{z}_{\mathrm{e}}[n] \tag{5.3}$$

where $\mathbf{H}_{\mathrm{e}} = [\mathbf{h}_{\mathrm{e},1}^{\mathrm{T}} \dots \mathbf{h}_{\mathrm{e},M}^{\mathrm{T}}] \in \mathbb{C}^{N_{\mathrm{t}} \times M}$ represents the system channel matrix between the BS and the multi-antennas Eve with $\mathbf{h}_{\mathrm{e},i} \in \mathbb{C}^{1 \times N_{\mathrm{t}}}$ being the channel coefficients between the BS and the i-th antenna at the Eve, and $\mathbf{z}_{\mathrm{e}}[n] \in \mathbb{C}^{M \times 1}$ gathers the independent AWGN components at all M antennas, with a variance of σ_e^2 each.

In conventional block-level precoding, the transmitted vector $\mathbf{x}[n]$ is modeled as $\mathbf{W}\mathbf{d}_{\mathrm{a}}[n]$, with $\mathbf{W}$ being the precoding matrix and $\mathbf{d}_{\mathrm{a}}[n] \in \mathbb{C}^{K \times 1}$ the data information intended for the K legitimate users. Specifically, the precoding matrix $\mathbf{W}$ is designed depending only on the CSI. For this reason, this type of precoding is commonly being referred to as channel-level or block-level precoding [SACO18, ASK+18]. Consequently, the precoder $\mathbf{W}$ changes only when the CSI changes and remains constant for several symbol slots, making the relation between $\mathbf{x}[n]$ and $\mathbf{d}_{\mathrm{a}}[n]$ linear.

In symbol-level precoding approach, however, the precoding module directly designs the transmitted signal vector $\mathbf{x}[n]$ based on both the CSI $\mathbf{H}$ and the input data symbols $\mathbf{d}_{\mathrm{a}}[n]$, hence the symbol-level nomenclature, i.e., the precoded signal $\mathbf{x}[n]$ changes at every symbol slot [ACO15c]. Therefore, this scheme optimizes the transmit vector $\mathbf{x}[n]$ without any intermediate steps (such as designing $\mathbf{W}$) while constructively exploiting the inter-user interference. As a result, the relation between the transmit vector $\mathbf{x}[n]$ and the input symbol vector $\mathbf{d}_{\mathrm{a}}[n]$ is no longer linear, as in the case of block-level precoding, and is inherently embedded into the precoding module. We note that the data symbols, $\mathbf{d}_{\mathrm{a}}[n]$, are assumed to be uncorrelated and drawn from a generic multi-level constellation having unit average power. We also assume that the channel to the Eve is known at the BS and that Eve is part of the system. The latter assumption provides Eve the advantage to know the modulation

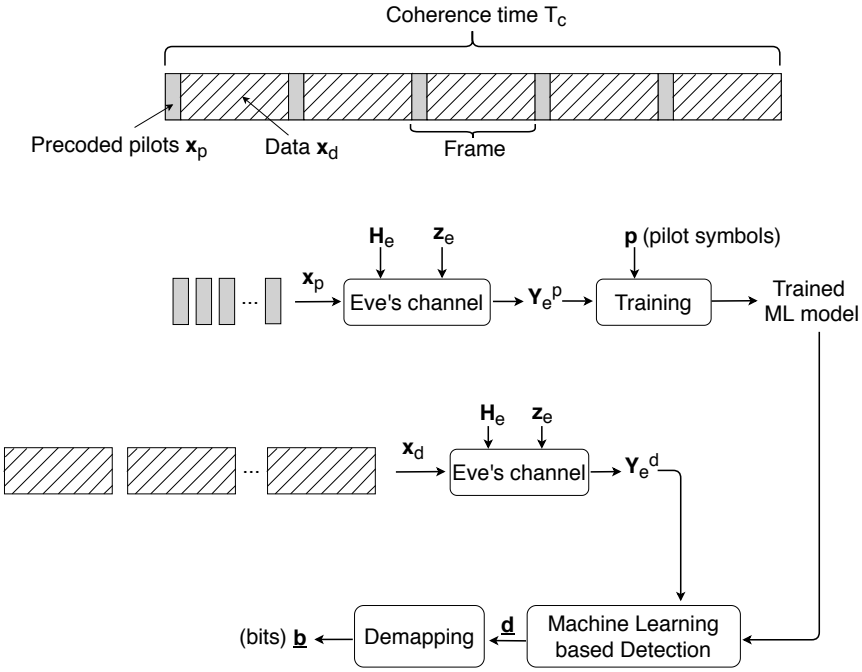


Figure 5.2: Summary of the ML-based attack.

and coding parameters used. For ease of notation, we drop the time index n in the remainder of the chapter.

5.2 ML-Based Attack

In this section, we introduce the ML-based attack, where an Eve can detect another user's symbols with a sufficiently low BER. Specifically, Eve would know the precoded pilot symbols used and their placement in the frame. This side information, which is usually publicly available in standards, can be exploited by the Eve and used to eavesdrop another user via the use of ML tools. This attack encompasses two phases 1) training phase and 2) inference phase. An overview of the attack is provided in Figure 5.2. Alongside, we present the two phases of the attack followed by a formulation of it. We then conclude this section by presenting a practical example of the ML-based attack on a SLP benchmark scheme.

5.2.1 Training Phase

As shown in Figure 5.2, the BS sends multiple frames within one coherence time T_c, where at the beginning of each frame, we find precoded pilots, $\mathbf{x}_p \in \mathbb{C}^{N_t \times 1}$, for channel and SNR estimation. The received pilot signals at the i-th antenna of the Eve, $y_{e,i}^p \in \mathbb{C}$, can be written

as follows

$$y_{\mathrm{e},i}^{\mathrm{p}} = \mathbf{h}_{\mathrm{e},i}\mathbf{x}_{\mathrm{p}} + z_{e,i} \tag{5.4}$$

where $z_{e,i} \in \mathbb{C}$ is the AWGN at the Eve with variance $\sigma_{e,i}^2$. The overall received pilot signal at Eve at all antennas, $\mathbf{y}_{\mathrm{e}}^{\mathrm{p}} \in \mathbb{C}^{1\times M}$, can be written as

$$\mathbf{y}_{\mathrm{e}}^{\mathrm{p}} = \mathbf{H}_{\mathrm{e}}\mathbf{x}_{\mathrm{p}} + \mathbf{z}_{\mathrm{e}}. \tag{5.5}$$

As $y_{\mathrm{e},i}^{\mathrm{p}}$ is Eve's received pilot signal, it knows beforehand the corresponding pilot symbol p that was transmitted. We note that the transmitted signal $\mathbf{x}_{\mathrm{p}}$ is a function of all user symbols, however the introduced attack targets a specific user, for which we know the transmitted pilot symbols in advance. In other words, the Eve is not trying to decode the data of all the users, it instead attempts to decode the data of a single user. As such, the Eve would create a mapping between the received signal $\mathbf{y}_{\mathrm{e}}^{\mathrm{p}}$ and the corresponding labels p. Thus, the Eve can exploit the knowledge of the precoded pilots, that are sent regularly according to communication standards for SNR estimation, in order to improve its detection performance.
We note that, as the number of antennas at Eve, M, increases, the number of input features increase accordingly, which often leads to better prediction accuracy. In essence, each antenna at Eve receives a different distorted copy of the same transmitted signal $\mathbf{x}_{\mathrm{p}}$, the more different copies of $\mathbf{x}_{\mathrm{p}}$ received by Eve, the better the machine-learning model performance will be, thus resulting in an improved symbol detection accuracy.

In the case of QPSK, there are only 4 possible pilot symbols, hence 4 classes. In the ML world, these classes are commonly being referred to as labels. As such, the corresponding machine learning problem is a supervised ML problem [SB14]. Meanwhile, since the labels are discrete, i.e., constellation points, the problem is categorised as a classification problem. Thus, the training set D contains N training points, i.e., $\{\mathbf{y}_{\mathrm{e},n}^{\mathrm{p}}, p_n\}, n \in \{1 \dots N\}$, where $y_{\mathrm{e},n}^{\mathrm{p}}$ represents the n-th received pilot signal at Eve, while p_n is the corresponding constellation point (label) associated with the observations $\mathbf{y}_{\mathrm{e},n}^{\mathrm{p}}$. We further define the training set D as

$$\{\mathbf{y}_{\mathrm{e},n}^{\mathrm{p}}, p_n\} \sim f(\mathbf{y}, p), n \in \{1 \dots N\} \tag{5.6}$$

where the operator $\sim$ signifies that the pairs $\{\mathbf{y}_{\mathrm{e},n}^{\mathrm{p}}, p_n\}, n \in \{1 \dots N\}$ are i.i.d with probability

distribution $f(\mathbf{y}, p)$. The training set D can be written in a more compact form as below

$$D = \{\mathbf{y}_{\mathrm{e}}^{\mathrm{p}}, \mathbf{p}\} \tag{5.7}$$

where $\mathbf{y}_{\mathrm{e}}^{\mathrm{p}} \in \mathbb{C}^{N \times M}$ is the received pilot symbols at Eve and $\mathbf{p} \in \mathbb{C}^{N \times 1}$ are the corresponding transmitted pilot symbols.

For simplicity, we denote the real and imaginary parts of $\mathbf{Y}_{\mathrm{e}}^{\mathrm{p}}$ as two real numbers, called input features, while we represent each pilot symbol in $\mathbf{p}$ using four[1] decimal $c = \{0, 1, 2, 3\}$, with each class corresponding to one QPSK symbol. Hence, the training set becomes $\{\mathbf{y}_{\mathrm{e}}^{\mathrm{p,r}}, \mathbf{y}_{\mathrm{e}}^{\mathrm{p,im}}, \mathbf{c}_{\mathrm{p}}\}$ with $\mathbf{y}_{\mathrm{e}}^{\mathrm{p,r}} \in \mathbb{R}^{N \times M}$ being the real part of Eve's received pilot signals, $\mathbf{y}_{\mathrm{e}}^{\mathrm{p,im}} \in \mathbb{R}^{N \times M}$ is its imaginary part, and $\mathbf{c}_{\mathrm{p}} \in \mathbb{N}^{N \times 1}$ are their corresponding classes. Based on the training set D, we derive a predictor (the trained ML model) that predicts a class l based on the observation y_{e}.

5.2.2 Inference Phase

As depicted in Figure 5.2, the BS sends precoded data, $\mathbf{x}_{\mathrm{d}} \in \mathbb{C}^{N_{\mathrm{t}} \times 1}$, to the users. The received symbol at Eve in each symbol period, $\mathbf{y}_{\mathrm{e}}^{\mathrm{d}} \in \mathbb{C}^{M \times 1}$, can be written as follows

$$\mathbf{y}_{\mathrm{e}}^{\mathrm{d}} = \mathbf{H}_{\mathrm{e}}^{\mathrm{T}} \mathbf{x}_{\mathrm{d}} + \mathbf{z}_{\mathrm{e}}. \tag{5.8}$$

where $\mathbf{x}_{\mathrm{d}} \in \mathbb{C}^{N_{\mathrm{t}} \times 1}$ represents the transmitted precoded signal from the N_{t} transmit antennas intended for all the users during one symbol period. If we assume that there are L data symbols in one coherence time, $\mathbf{y}_{\mathrm{e}}^{\mathrm{d}} \in \mathbb{C}^{L \times M}$ represents the collection of all received symbols at Eve during one coherence time T_{c}.

The goal of classification is to predict a label d for a new, usually unobserved, received precoded signals, $\mathbf{y}_{\mathrm{e}}^{\mathrm{d}}$, that is outside the pilot signals. Therefore, a machine learning-based detection can be performed over $\mathbf{y}_{\mathrm{e}}^{\mathrm{d}}$ using the trained ML model, as shown in Figure 2. The output of the ML-based detector block is the symbols vector $\underline{\mathbf{d}} \in \mathbb{C}^{L \times 1}$, which is an estimate of $\mathbf{d} \in \mathbb{C}^{L \times 1}$ (symbols intended for a specific user). Then, we perform demapping over $\underline{\mathbf{d}}$ to obtain the corresponding bit-vector $\underline{\mathbf{b}}$. Finally, we compare $\underline{\mathbf{b}}$ to $\mathbf{b}$ (the actual bits sent to a specific user) to obtain the BER at Eve.

[1]The number of classes depends on the modulation order used. In the case of QPSK, the number of classes equals four.

5.2.3 ML Attack Formulation

As stated above, the idea of supervised learning classification is to find a robust mapping h between the input features $\mathbf{y}_{\mathrm{e}}^{\mathrm{p}}$ and the classes $\mathbf{p}$ using the training dataset D. To further illustrate, we give the example of the support vector machine (SVM) classifier. The goal of SVM classifier is to separate the four[2] classes using lines that are usually hyperplanes.
The hyperplanes can be described using the below equation

$$\mathbf{w}\,\mathbf{y}_{\mathrm{e}} + b = 0 \tag{5.9}$$

where $\mathbf{w}$ is the normal to the hyperplane and $\frac{b}{\|\mathbf{w}\|}$ is the perpendicular distance from the hyperplane to the origin. Support vectors, as their names imply, are the separating hyperplanes and the goal of the SVM algorithm is to orientate the hyperplanes in such a way to be as far as possible from the closest members of the different classes. Hence, implementing the SVM classifier boils down to selecting the parameters $\mathbf{w}$ and b that best achieve the aforementioned goal through the use of the training data. Once these parameters are estimated, the trained ML model can be used to directly predict the transmitted symbols from observing any received signal at the Eve during the same coherence time T_{c}.

5.2.4 Attack Example on a Benchmark Scheme - CISPM

As a Benchmark, we use the approach in [ACO15c], the CISPM scheme. This particular scheme is designed to exploit inter-user interference for power gains at the intended users, in other words, this scheme propels the intended users' received signals deeper into the correct detection region of the desired symbol for each user. Although this scheme applies no processing towards Eve's received signal, it still provides security gains. Since the transmitted signals are designed to have CI only with the intended users channels, Eve's received signal would in all likelihood fall in a different region than the correct one, as his channel is different than the intended user's one. Hence, the benchmark scheme is inherently secure against a conventional eavesdropper. The corresponding optimization problem is defined as

[2]The number of classes depends on the modulation order. For QPSK, the number of classes equals four.

$$\mathbf{x}(\mathbf{d},\mathbf{H},\gamma)=\arg\min_{x}||\mathbf{x}||^2 \tag{5.10}$$
$$\text{subject to}$$
$$\text{Re}\{\mathbf{h}_j\mathbf{x}\} \trianglelefteq \sigma_z\sqrt{\gamma_j}\text{Re}\{d_j\}, \quad j\in\{1\dots K\}$$
$$\text{Im}\{\mathbf{h}_j\mathbf{x}\} \trianglelefteq \sigma_z\sqrt{\gamma_j}\text{Im}\{d_j\}, \quad j\in\{1\dots K\}$$

where γ_j is the target SINR for the j-th user, $\gamma=[\gamma_1\dots\gamma_K]\in\mathbb{R}^{K\times 1}$ represents the target SINR for all users. This problem is convex as both objective function and constraints are convex and can be solved efficiently using second order cone programming [BV04].

Although the CISPM scheme is secure against conventional eavesdropper, it can not stand against a sophisticated Eve that employs machine learning for symbol detection. As we shall demonstrate subsequently, the CISPM is vulnerable to the ML-based attack as it uses no specific pre-processing for Eve's received signal.

The setup of the experiment is as follows. We consider a BS using the CISPM scheme to precode the transmit signal $\mathbf{x}$ intended to the K users. Particularly, both transmit signals $\mathbf{x}_\text{p}$ and $\mathbf{x}_\text{d}$ are designed using the CISPM scheme. In Table 5.1, we show the symbol detection accuracy[3] of the different classifiers when the CISPM scheme is used. We note that, for this simulation we used QPSK modulation, $N_\text{t}=10$, $K=6$, and a single antenna Eve. In this particular experiment, we used 100 symbols as pilots and 1000 symbols as data. In this setting, we used Matlab for data generation and Python for classification and performance analysis. We should mention that we did not use deep learning [Den14] in this context despite its high performance mainly because it requires considerable amount of training data that is not available in our case. In fact, in each coherence time, only one portion of the frame is dedicated to pilot symbols, which in turn serve as the training data, and since the channel and data change in each frame, training could be done only within the frame itself, hence the limitation of training data. We notice that the symbol detection accuracy is relatively high when using such a scheme. This is due to the fact that at the BS, there was no particular constraint or processing towards the received signal at Eve. Therefore, Eve is using the power of ML tools to be able to still discriminate between the intended symbols for a particular user, even though the precoded signal was specifically designed for channel vectors that are considerably different than Eve's one.

[3]This accuracy refers to the accuracy of the trained ML model, i.e., the learning block of the diagram in Figure 5.2

Table 5.1: Performance of different classifiers for SLP-based dataset.

Classifier	**Symbol detection accuracy**
Support Vector Machines	0.7
Gradient Boosting Machine	0.63
Logistic Regression	0.71
K-Nearest Neighbors	0.66
XGBoost	0.68
Light GBM	0.67

5.3 Countermeasure - PLS Scheme

In this section, we first present a novel secure principle for designing secure precoding schemes that counteract the ML-based Eve. Then, based on this principle, we propose a precoder that yields high achievable BER at the Eve. Since the formulation of the latter scheme is non-convex, we propose an equivalent convex formulation. As a tradeoff between security and transmit power, we propose three other convex secure precoding schemes that are more energy efficient.

5.3.1 Secure Design Principle

As presented in the earlier section, an Eve with multiple antennas can achieve a decent detection performance that allows it to detect most of the received symbols by using the power of machine leaning. In order to dramatically worsen its detection performance, we propose a novel design principle of the precoded signal. We deliberately force the received signal at the Eve to lie at the boundaries of the detection regions.

This specific design has two advantages 1) to increase PLS as the detection decisions at the Eve will be mostly made depending on noise and 2) for energy efficiency purposes since it involves only a small deviations of the received constellation point.

5.3.2 PLS Schemes

In this section, we introduce the SLP-based countermeasure, that we call subsequently PLS scheme. Similar to [MST+19b], the idea is to design the transmitted signal $\mathbf{x}$ so as to have constructive interference at the intended users, and at the same time, to confuse the Eve by maximizing its detection uncertainty using the above secure design principle.

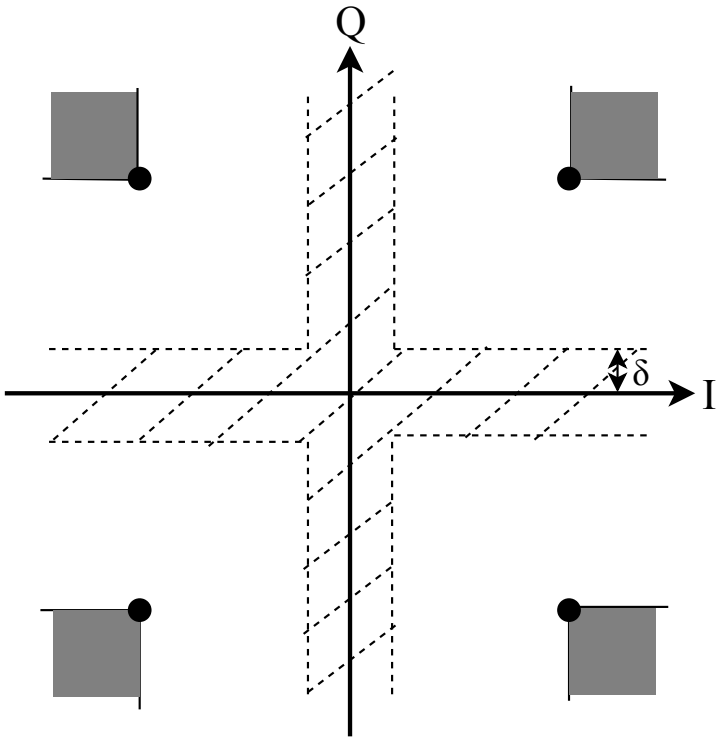


Figure 5.3: Example of PLS scheme using QPSK modulation.

The PLS scheme is demonstrated in Figure 5.3. Herein, we adopt the example of QPSK modulation for illustration purposes, where the dark circles represent the constellation points. We design the transmitted signal in such a way to have constructive interference at the intended users, that is represented by the grey shared regions. The goal here is to push the received points deeper into the detection region in order to improve the intended users' detection accuracy. However, we design Eve's received signal to lie in the strapped region (boundary of the detection region), whose width is controlled by the parameter δ. The lower the value of δ, the sharper the strapped region, thus the higher the probability of falling into a different region after noised adds up, resulting in higher BER at Eve. It is worth noticing that this particular design makes Eve's detection decisions to be mostly based on noise.

For a downlink MU-MISO system, with N_{t} transmit antennas and K users, the aforementioned precoder design problem can be formulated as a power minimization problem as

$$\mathbf{x}(\mathbf{d},\mathbf{H},\mathbf{H}_{\mathrm{e}},\gamma,\delta)=\arg\min_{x}\|\mathbf{x}\|^2 \tag{5.11}$$

subject to

$$\mathrm{Re}\{\mathbf{h}_j\mathbf{x}\}\trianglelefteq\sigma_z\sqrt{\gamma_j}\mathrm{Re}\{d_j\},\quad j\in\{1\dots K\} \tag{5.12}$$

$$\mathrm{Im}\{\mathbf{h}_j\mathbf{x}\}\trianglelefteq\sigma_z\sqrt{\gamma_j}\mathrm{Im}\{d_j\},\quad j\in\{1\dots K\} \tag{5.13}$$

$$\mathrm{Re}\{\mathbf{h}_{\mathrm{e},i}\mathbf{x}\}\lesseqgtr\mp\delta,\quad i\in\{1\dots M\} \tag{5.14}$$

$$\mathrm{Im}\{\mathbf{h}_{\mathrm{e},i}\mathbf{x}\}\lesseqgtr\mp\delta,\quad i\in\{1\dots M\} \tag{5.15}$$

where $\mathbf{h}_j\mathbf{x}$ is the j-th user's noiseless received signal, $\mathbf{h}_{\mathrm{e},i}\mathbf{x}$ is the noiseless received signal at Eve's i-th receive antenna, $\trianglelefteq$ denotes the correct detection region [ACO17], δ is the distance parameter controlling the width of the strapped region, the operator $\lesseqgtr\mp$ refers to $\leq+$ and $\geq-$ simultaneously, while Re and Im denotes real and imaginary parts, respectively. The above problem[4] is non-convex because of the non-convexity of target region of Eve's received signal.The physical meaning of constraints in the PLS scheme are of two types. CI at the legitimate users, achieved by constraints (5.12) and (5.13). The CI effect results in increased power gains at the legitimate users. However, constraints (5.14) and (5.15) are intended to have destructive interference at Eve to increase the uncertainty during symbol detection. Particularly, making his received signal lie at the boundary regions, so that the noise will move it in either direction of the detection regions and hence increase the BER at the latter. We note that the quality of service of the users, exhibited by constraints (5.12) and (5.13), does not affect constraints (5.14) and (5.15) of the boundary regions.

Hence, in problem (5.11), the non-convex constraints are the ones related to the Eve, they are as follows

$$\mathrm{Re}\{\mathbf{h}_{\mathrm{e},i}\mathbf{x}\}\lesseqgtr\mp\delta, i\in\{1\dots M\} \tag{5.16}$$

$$\mathrm{Im}\{\mathbf{h}_{\mathrm{e},i}\mathbf{x}\}\lesseqgtr\mp\delta, i\in\{1\dots M\}. \tag{5.17}$$

Given these constraints, the feasibility region of Eve's received signal is non-convex, as shown in Figure 5.3 (the strapped region). In the following, we propose four convex imple-

[4]We note that the formulation in (5.11) is valid for a generic multi-level constellation, such as M-QAM, however it can be tailored to other constellations as APSK.

mentations of the problem in (5.11), with varying security and energy efficiency trade-offs.

PLS - Square Scheme

Similar to the idea in [XRS21], we take the intersection of the vertical boundary region and the horizontal one, characterized by constraints (5.16) and (5.17), respectively. The intersection of the two form a square, hence the name. With this, the problem in (5.11) becomes convex and could be solved efficiently using convex solvers such as CVX. This scheme designs Eve's received signal to lie in the square whose center is the origin and side is 2δ. When noise is added, the received signal at Eve will lie on any of the 4 detection regions (in case of QPSK), thus providing high security. Algorithm 1 explains the process of signal design of the PLS - Square scheme.

Algorithm 1 PLS - Square scheme

Input: $\mathbf{d}, \mathbf{H}, \mathbf{H}_{\mathrm{e}}, \gamma, \sigma_z^2, \delta$;

1: **while** solve problem (5.11) **as follows:**
2: Satisfy CI constraints in (5.12) and (5.13)
3: Satisfy constraints (5.14) and (5.15) simultaneously
4: **end while**

Output: $\mathbf{x}_s$

We note that Algorithm 1 is executed for every symbol slot.

PLS - Two-Steps Nearest Scheme

In this scheme, as its name implies, the transmit signal $\mathbf{x}$ is designed in two steps. In the first step, we aim to determine the region in which Eve's received signal would lie when CISPM scheme is used (we name the transmit signal inhere $\mathbf{x}_{CI}$). Once we identify the coordinates of Eve's received signal, $\mathbf{h}_{\mathrm{e}}\mathbf{x}_{CI}$, we feed this information into the second problem as an input. Herein, we execute problem (5.11) with the formulation of nearest. Namely, depending on where Eve's signal would land, we design it to fall into the nearest boundary region, either vertical one or horizontal one. We note that, when considering Gray mapping, the detection regions factorize when using this scheme, i.e., the bits are transmitted in real and imaginary parts independently and are also detected in the same way, leading to probabilities factorization due to the underlying statistical independence. Thus, this scheme is optimal with respect to the BER.

The detailed steps of the procedure are found in Algorithm 2.

Algorithm 2 PLS - Two-steps nearest

Input: $\mathbf{d}, \mathbf{H}, \gamma, \sigma_z^2$; ▷ Step 1

Solve problem (5.10)

Output: $\mathbf{x}_{CI}$

Input: $\mathbf{d}, \mathbf{H}, \mathbf{H}_\mathrm{e}, \gamma, \sigma_z^2, \delta, \mathbf{x}_{CI}$; ▷ Step 2

while solve problem (5.11) **as follows:**
    Satisfy CI constraints in (5.12) and (5.13)
    if $|\mathrm{Re}\{\mathbf{h}_{\mathrm{e},i}\mathbf{x}_{CI}\}| < |\mathrm{Im}\{\mathbf{h}_{\mathrm{e},i}\mathbf{x}_{CI}\}|$ **then**
        Satisfy constraint (5.14)
    else
        Satisfy constraint (5.15)
    end if
end while

Output: $\mathbf{x}_n$

where condition $|\mathrm{Re}\{\mathbf{h}_{\mathrm{e},i}\mathbf{x}_{\mathbf{CI}}\}| < |\mathrm{Im}\{\mathbf{h}_{\mathrm{e},i}\mathbf{x}_{CI}\}|$ implies that Eve's received signal is closer to the vertical boundary region, hence in this scheme we design Eve's received signal to lie in the vertical boundary region (constraint characterized by (5.14)). Otherwise, we design Eve's received signal to lie in the horizontal boundary region instead (represented by constraint (5.15)).

PLS - Two-Steps Farthest Scheme

This scheme can be considered as opposite of the PLS - Two-steps nearest scheme, i.e., instead of picking the closest boundary region, it always chooses the farthest one. Intuitively, this scheme would provide higher security gains than the nearest scheme (as it pushes the constellation point even farther, leading to higher chances of falling into the wrong region), however, it would consume more transmit power, the bigger the introduced deviation by the constraint, the higher power required to move it. Below we formulate the algorithm for the PLS - Two-steps farthest scheme.

Algorithm 3 PLS scheme - Two-steps farthest

Input: $\mathbf{d}, \mathbf{H}, \gamma, \sigma_z^2$; ▷ Step 1

Solve problem (5.10)

Output: $\mathbf{x}_{CI}$

Input: $\mathbf{d}, \mathbf{H}, \mathbf{H}_{\mathrm{e}}, \gamma, \sigma_z^2, \delta, \mathbf{x}_{CI}$; ▷ Step 2

while Solve problem (5.11) **as follows:**

Satisfy CI constraints in (5.12) and (5.13)

if $|Re\{\mathbf{h}_{\mathrm{e},i}\mathbf{x}_{\mathbf{CI}}\}| > |Im\{\mathbf{h}_{\mathrm{e},i}\mathbf{x}_{CI}\}|$ **then**

Satisfy constraint (5.14)

else

Satisfy constraint (5.15)

end if

end while

Output: $\mathbf{x}_f$

where condition $|\mathrm{Re}\{\mathbf{h}_{\mathrm{e},i}\mathbf{x}_{CI}\}| > |\mathrm{Im}\{\mathbf{h}_{\mathrm{e},i}\mathbf{x}_{CI}\}|$ indicates that Eve's received constellation point is closer to the horizontal boundary region, and since this is farthest scheme, we design Eve's received signal to fall into vertical boundary region (constraint characterized by (5.14)). However, if the condition is not fulfilled, we design Eve's received signal to lie in the horizontal boundary region instead (represented by constraint (5.15)).

PLS - Two-Steps 6-Lines Scheme

This scheme is also a two-step scheme, however, this particular scheme is designed for higher order modulation such as 16-QAM, where the detection regions are numerous. Figure 5.4 depicts the six lines that define the 6 boundary regions (the region where Eve's received signal should lie) for this scheme. This scheme consist of designing Eve's received signal to lie in one of these lines by choosing the closest boundary region to it. Below we define the different constraints characterizing the 6 boundary regions:

$$\mathrm{Re}\{\mathbf{h}_{\mathrm{e},i}\mathbf{x}\} \lesseqgtr \mp\delta, \quad i \in \{1 \dots M\} \tag{5.18}$$

$$\mathrm{Re}\{\mathbf{h}_{\mathrm{e},i}\mathbf{x}\} + th_{\mathrm{re}} \lesseqgtr \mp\delta, \quad i \in \{1 \dots M\} \tag{5.19}$$

$$\mathrm{Re}\{\mathbf{h}_{\mathrm{e},i}\mathbf{x}\} - th_{\mathrm{re}} \lesseqgtr \mp\delta, \quad i \in \{1 \dots M\} \tag{5.20}$$

$$\mathrm{Im}\{\mathbf{h}_{\mathrm{e},i}\mathbf{x}\} \lesseqgtr \mp\delta, \quad i \in \{1 \dots M\} \tag{5.21}$$

$$\mathrm{Im}\{\mathbf{h}_{\mathrm{e},i}\mathbf{x}\} + th_{\mathrm{im}} \lesseqgtr \mp\delta, \quad i \in \{1 \dots M\} \tag{5.22}$$

$$\mathrm{Im}\{\mathbf{h}_{\mathrm{e},i}\mathbf{x}\} - th_{\mathrm{im}} \lesseqgtr \mp\delta, \quad i \in \{1 \dots M\} \tag{5.23}$$

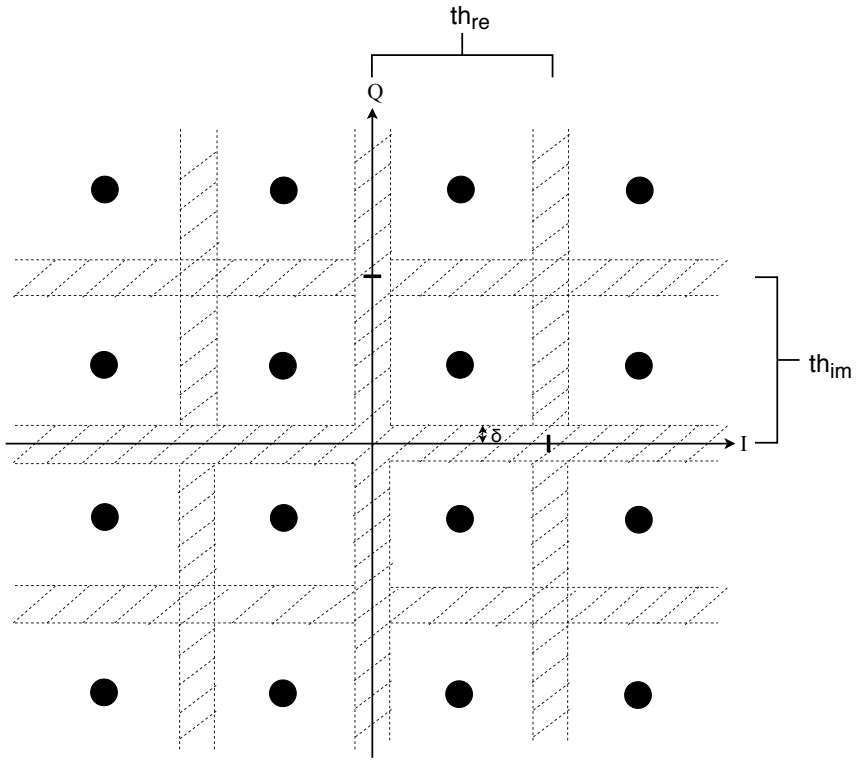


Figure 5.4: 16-QAM constellation showing the 6 lines boundary regions.

where th_{re} and th_{im} denotes the threshold that determines the boundary regions of both real and imaginary parts, respectively, as depicted in Figure (5.4).

Naturally, this scheme would require less power than all the aforementioned PLS schemes, considering that it introduces the smallest deviation of the constellation point. Algorithm 4 shows the details of this scheme.

Algorithm 4 PLS scheme - Two-steps 6lines

Input: $\mathbf{d}, \mathbf{H}, \gamma, \sigma_z^2$; ▷ Step 1

1: Solve problem (5.10)

Output: $\mathbf{x}_{CI}$

Input: $\mathbf{d}, \mathbf{H}, \mathbf{H}_{\text{e}}, \gamma, \sigma_z^2, \delta, \mathbf{x}_{CI}$; ▷ Step 2

2: **while** Solve problem (5.11) **as follows:**

3: Satisfy CI constraints in (5.12) and (5.13)

4: Determine the closest line to $\mathbf{h}_{\text{e},i}\mathbf{x}_{CI}$

5: Apply the corresponding constraint from (5.18,5.19,5.20,5.21,5.22, and 5.23)

Output: $\mathbf{x}_{6l}$

We note that the above-mentioned implementations of the optimization problem in (5.11) are convex, and thus their global optimum can be obtained using standard convex optimiza-

Table 5.2: Performance of different classifiers for countermeasure SLP-based dataset.

Classifier	**Symbol detection accuracy**
Support Vector Machines	0.28
Gradient Boosting Machine	0.3
Logistic Regression	0.28
K-Nearest Neighbors	0.26
XGBoost	0.28
Light GBM	0.27

tion tools [BV04].

5.3.3 Attack Example on PLS Scheme

As demonstrated in Section 5.2.4, a sophisticated eavesdropper that employs ML can predict a precoded signal to a user with high accuracy. However, in the case of the PLS scheme (we used Random scheme implementation in this experiment), the prediction accuracy is quite low compared to the CISPM scheme. In fact, these values are close to $\frac{1}{4}$, which is the lower bound when Eve has no side information and is randomly picking symbols out of the QPSK set. The main reason for this behavior is because we are forcing Eve's received signal to lie at the boundary region. Hence the received signal at the latter will be randomly distorted because of noise, thus making it very difficult for the ML algorithm to map the received signals to the pilot symbols. Similarly, we compared many classifiers, where their prediction accuracy is summarised in Table 5.2.

5.4 Simulation Results

In the numerical results, we define the considered performance metrics. First, the total transmit power by the BS antennas is defined as $P_{\text{tot}} = ||\mathbf{x}||^2$. In the simulations, we take the average of the above quantity over a large number of symbol slots, i.e., $E_{\mathbf{d}_n,\mathbf{H}}[P_{\text{tot}}]$, to obtain the frame-level total transmit power, which is then averaged over a large number of channel realizations. We also compute the effective BER at the Eve, by detecting the received signal at the latter. In addition, we compute the BER at the intended users in order to investigate the impact of the countermeasure scheme on the intended user performance. We note that we employ Gray mapping in the numerical simulations.

Consequently, we define the metric that we call effective rate, $\overline{R}_{\mathrm{a}}$, that quantifies the error-free part of the total rate, and can be written as

$$\overline{R}_{\mathrm{a}} = WR_c(|1 - 2BER_{\mathrm{a}}|) \tag{5.24}$$

where W is the bandwidth, R_c is the rate (equals 2 in the case of QPSK), $|\cdot|$ is the absolute value, and BER_{a} is the effective BER, where "a" can be either intended user or Eve. Since a BER value of 0.5 implies no information is communicated, it is verified in (5.24) that the effective rate returns 0 when the BER equals 0.5 and full rate when the BER is 0.
Similar to the performance metric used in [KMA+19], we define the secure rate as

$$R_{\mathrm{sec}} = \overline{R}_{\mathrm{int}} - \overline{R}_{\mathrm{eve}} \tag{5.25}$$

where $\overline{R}_{\mathrm{int}}$ is the effective rate at the intended user and $\overline{R}_{\mathrm{eve}}$ is the one at the Eve.

Last but not least, we define a new metric that we call Energy Efficiency for Secure Transmission (EEST) η_{eest} that combines both the secure bits transmitted and the transmit power consumed, so it will be [secure bits/s], and defined as

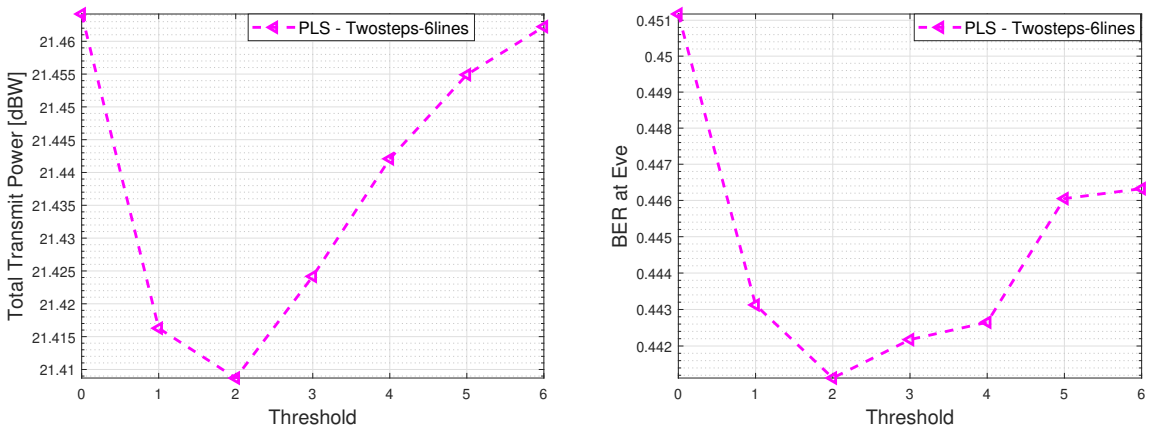
$$\eta_{\mathrm{eest}} = \frac{R_{\mathrm{sec}}[\mathrm{bit/s}]}{E_{\mathbf{d}_n,\mathbf{H}}[P_{\mathrm{tot}}]}. \tag{5.26}$$

We note that we have devised this metric to compare the overall performance of the PLS schemes, by taking into account simultaneously security and transmit power. The EEST increases either by increasing the secure rate and/or by decreasing the power consumption, and thus, higher values of it indicate either better security and/or lower power consumption. Therefore, if an A scheme provides higher EEST than another scheme B, then scheme A is providing higher security with respect to its consumed power. We note that the metric for secure transmission is the secure rate, in the numerator of the EEST, the higher the secure rate, the more secure the scheme will be. Particularly, we can observe a very high EEST that was due to very low power consumption and little security. As a result, there is no value of it that can guarantee secure transmission.

In the below figures, we used the SVM classifier for the ML-based attack as it possesses

one of the highest prediction accuracies, thus making the Eve as sophisticated as possible. Moreover, all results have been averaged over 100 channel realisations and using 1000 symbols in each realizations, in order to give an accurate performance analysis of the proposed precoders. To train the SVM classifier, we have used the following parameters: "onevone coding" without optimizing the hyperparameters and 1000 symbols for each constellation point for training. The simulated MU-MISO system comprises of $N_{\mathrm{t}} = 15$ antennas at the BS, $K = 6$ single-antenna intended users, $M = \{1, 3, 6\}$ number of antennas at Eve depending on the simulation, $\sigma_z^2 = 1Watt$, the bandwidth $W = 20Mhz$, Gray mapping, and $\sigma_e^2 = \{0, 0.2, \ldots, 1\}$ depending on the figure. In the numerical results, we assume that the users and Eve have the same coherence time length. Particularly, we use a coherence time of 1000 symbols which is long enough for Eve to have sufficient pilots for training. However, this does not apply to the intended users since they use conventional detection methods.

We note that Matlab was used as the main software for simulations, embedded with CVX as the convex optimization solver. We note that since the proposed schemes and the benchmark scheme pertain to the same class of optimization convex problems, they are of comparable complexity.



(a) Total transmit power vs. threshold

(b) BER at Eve vs. threshold

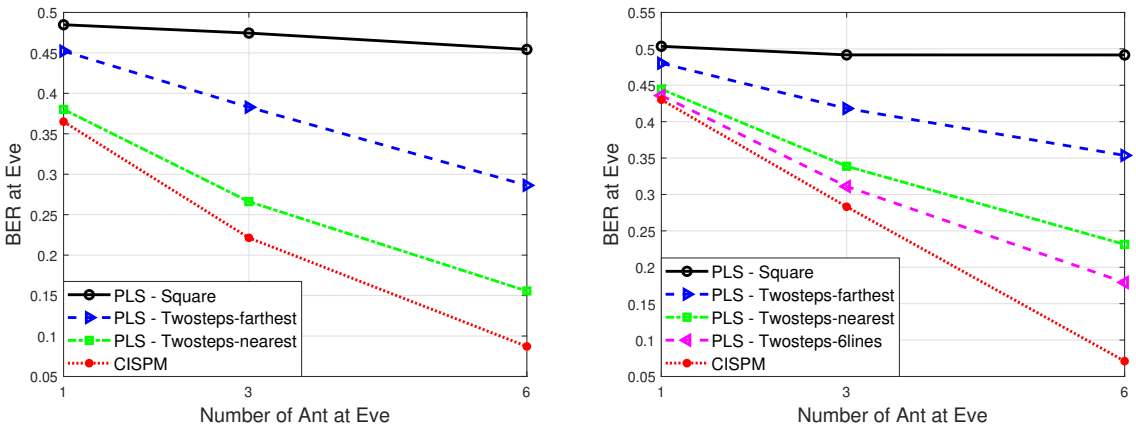
Figure 5.5: 16-QAM with $N_{\mathrm{t}} = 15$, $K = 6$, 15 dB target SINR, and $M = 1$.

For the numerical results, a value for the thresholds $\mathrm{th}_{\mathrm{re}}$ and $\mathrm{th}_{\mathrm{im}}$ has to be chosen. For that, we investigate the impact of the thresholds on both the total transmit power and the BER at Eve. We note that these threshold values are relevant only for the Two-steps 6-lines scheme, where we set $\mathrm{th}_{\mathrm{re}} = \mathrm{th}_{\mathrm{im}}$ because of the symmetry of the 16-QAM constellation. As

shown in Figure 5.5, part (a) plots the total transmit power vs. the threshold while part (b) depicts the BER at Eve vs. the threshold.

As expected, and in accordance with the SLP schemes behavior in this section, higher security (BER at Eve) comes at the cost of higher power consumption. In Figure 5.5a, the transmit power shows a minimum, corresponding to the position of the threshold close to the original position of the received symbols, i.e., the constraints are not too stringent, hence the consequent power saving. Before reaching this minimum, we observe a decrease of the power, due to the constraints becoming less and less strict. However, after reaching the minimum, the total power starts increasing because of constraints getting more stringent, but going in the opposite direction. As for the behavior of the BER at Eve depicted in Figure 5.5b, the stricter the constraints, the more distant are the boundary regions from the original position of the received symbols, the more likely for Eve to make wrong detection decisions, hence the increase in the BER at Eve.

Finally, we have picked the value of $\text{th}_{\text{re}} = \text{th}_{\text{im}} = \frac{6}{\sqrt{(2)}}$ as a trade-off value between the total power consumed and security. But after all, varying the threshold barely affects the performance, as shown in Figure 5.5, values of transmit power and BER vary slightly with the change of the threshold. We note that the suggested fixed value of the threshold is for the selected parameters used in this section.

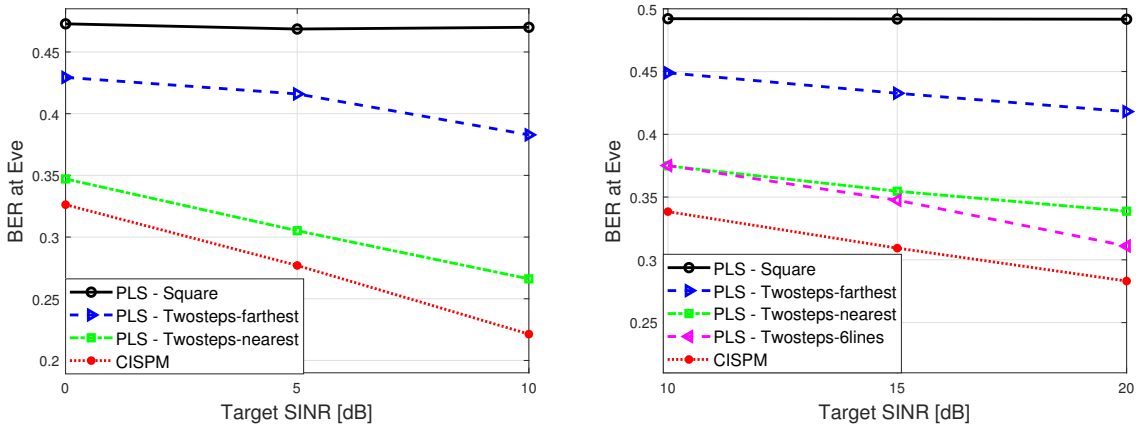


(a) QPSK with $\gamma = 10$dB (b) 16-QAM with $\gamma = 20$dB

Figure 5.6: BER vs. number antennas at Eve, with $N_{\text{t}} = 15$, $K = 6$, and $\delta = 0.1$

Figure 5.6 plots the BERs at Eve as a function of the number of antennas at Eve, for case of QPSK and 16-QAM modulations. We compare the benchmark scheme, CISPM, with the

PLS schemes. The parameters used in the simulation are $N_t = 15$, $K = 6$, $\delta = 0.1$, and 10dB of target SINR at intended user. In both QPSK and 16-QAM, all PLS schemes outperform the CISPM one with PLS Square providing the highest security gains. We also observe that the more antennas at Eve, the higher the prediction accuracy (more samples of same signal), and ultimately the lower the BER. However, we notice that PLS - Square scheme is not affected much by this increase in the number of antennas at Eve, and it is due to nature of this scheme, in particular, it randomly assigns Eve's received signal to either the horizontal boundary region or vertical one, and hence making it super hard for the ML engine to find a relationship, as it is practically impossible to predict something random, thus it provides the highest security. On the other hand, the other PLS schemes still provide better security gains than the CISPM one, but not as good as the Random one, mainly because in their design is inherently deterministic, hence the ML engine would often find ways to find the relationship between the observed precoded symbols and the actual symbols intended for a specific user, thus the decrease of the BER as the number of antennas at Eve increase, i.e., more training data. We also observe that in the case of 16-QAM.

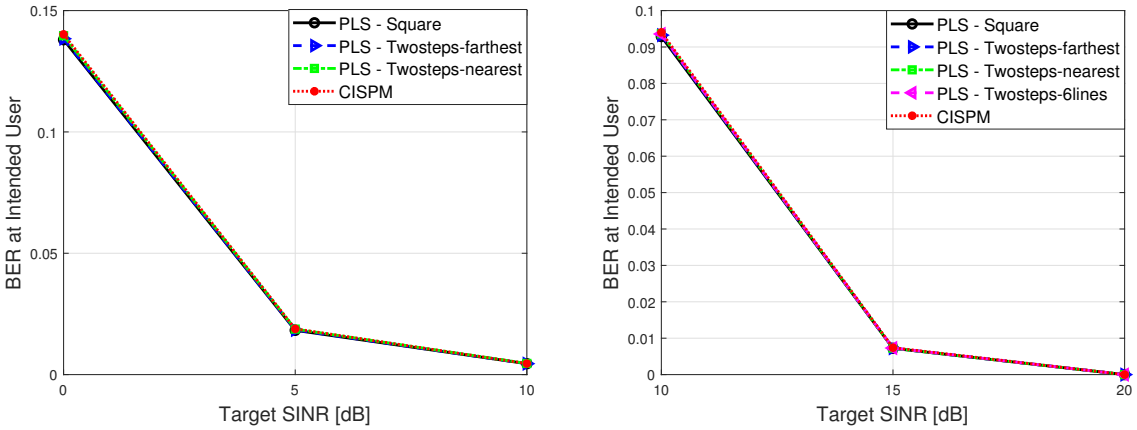


(a) QPSK (b) 16-QAM

Figure 5.7: BER vs. target SINR, with $N_t = 15$, $K = 6$, $\delta = 0.1$, and $M = 3$.

Figure 5.7 depicts the BERs at Eve as a function of the target SINR at the intended user, that we set to the same value for all users for simplicity, for case of QPSK and 16-QAM modulations. We compare the benchmark CISPM scheme with the PLS schemes. The parameters used in the simulation are $N_t = 15$, $K = 6$, $\delta = 0.1$, and $M = 3$. Similarly, for all precoding schemes, the higher the target SINR at the intended user, the lower the BER

at the Eve, except for the PLS - Square scheme, where its performance is not affected by the target SINR due it its invulnerability to ML-based attack, given the employed randomness in the design. We also observe that all PLS schemes outperform the CISPM one as they include some Eve-related constraints in their formulation, with the same behavior in both QPSK and 16-QAM.

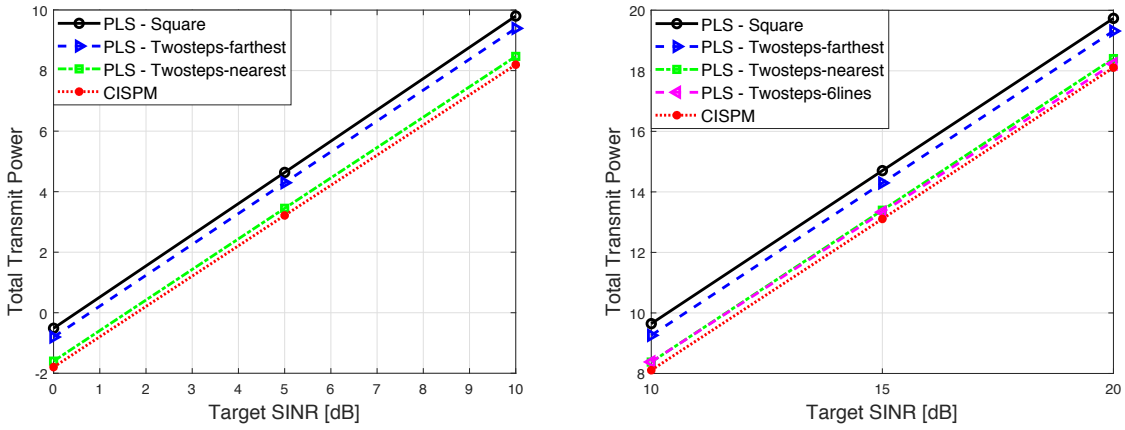


(a) QPSK (b) 16-QAM

Figure 5.8: BER at intended user vs. γ, with $N_{\mathrm{t}} = 15$, $K = 6$, $\delta = 0.1$, and $M = 3$.

Fig 5.8 plots the BERs at intended user as a function of the target SINR at the intended user, that we set to the same value for all users for simplicity, for case of QPSK and 16-QAM modulations. We compare the benchmark CISPM scheme with the PLS schemes. The parameters used in the simulation are $N_{\mathrm{t}} = 15$, $K = 6$, $\delta = 0.1$, and $M = 3$. As expected, for all the schemes and both modulations, the BER at the intended user decreases as the target SINR at the latter increases. That is, higher target SINR implies higher transmit power, hence better SNR at intended, that leads to an improved BER. We also note that all the schemes have the same performance at the intended user, a match is observed among all schemes, both PLS and CISPM, consequently, we can conclude that despite the security gains offered by the PLS schemes, their use does not impact the performance at the intended user.

Figure 5.9 show the total transmit power, in dB, as a function of the target SINR, that we set to the same value for all users for simplicity, for case of QPSK and 16-QAM modulations. We compare the benchmark CISPM with the PLS schemes. The parameters used in the simulation are $N_{\mathrm{t}} = 15$, $K = 6$, $\delta = 0.1$, and $M = 3$. As expected, the power consumption

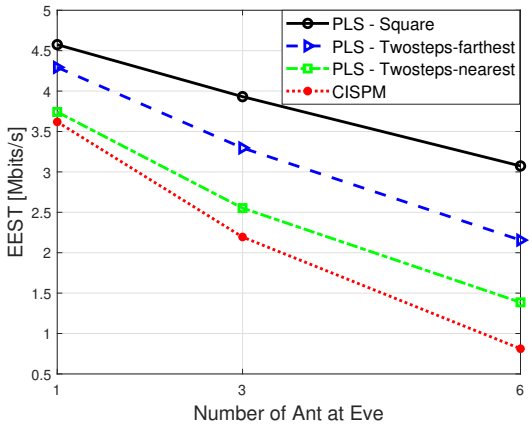


(a) QPSK (b) 16-QAM

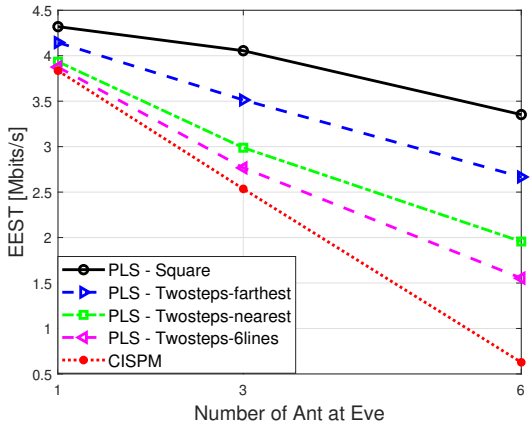
Figure 5.9: Total transmit power vs. target SINR, with $N_{\mathrm{t}} = 15$, $K = 6$, $\delta = 0.1$, and $M = 3$.

increases linearly with target SINR, with benchmark scheme consuming a bit less power than the PLS ones. In particular, the PLS - Square scheme consumes more than PLS - Twosteps-farthest which consumes more than PLS - Twosteps-nearest. This is intuitive in the sense that the Twosteps-nearest consumes less than the Twosteps-farthest as it incur a smaller deviation of the target received signal. To illustrate more, this behavior is due to the fact that, the more we constrain our signal design problem, the more power is required. Moreover, the more antennas at the Eve, the higher the transmit power for the PLS schemes, i.e. number of constraints increase with the number of antennas at Eve. For 16-QAM alone, the Twosteps-6lines consuming the least among all PLS schemes, as it requires the smallest deviation of the Eve's received constellation point, i.e., it moves it to the closest line among the 6 lines (boundary regions). We also observe that the power consumption difference between the two scheme is only of 1 dB in the case of 3 antennas at Eve. Thus, only a small additional power consumption is required to provide such high security.

Figure 5.10 plots the energy efficiency for secure transmission η_{eest}, in [Secure bits/s], as a function of the number of antennas at Eve, for case of QPSK and 16-QAM modulations. We compare the benchmark scheme with the PLS schemes. The parameters used in the simulation are $N_{\mathrm{t}} = 15$, $K = 6$, $\delta = 0.1$, and 10 dB target SINR for QPSK and 20 dB for 16-QAM. In the case of QPSK, it turns out that, when taking into account both total transmit power and secure rate, PLS schemes still outperform CISPM scheme. This is due to the fact that the difference in power consumption is relatively smaller than the difference in secure bits, hence



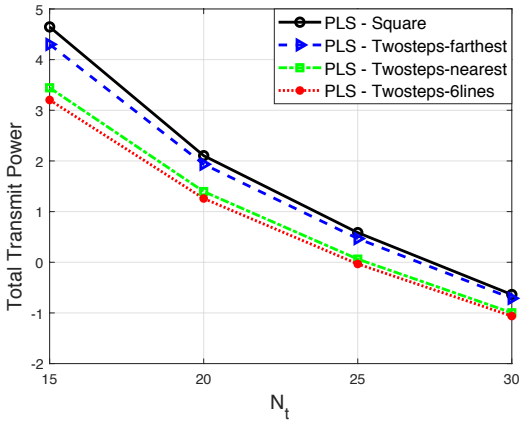
(a) QPSK with $\gamma = 10$ dB



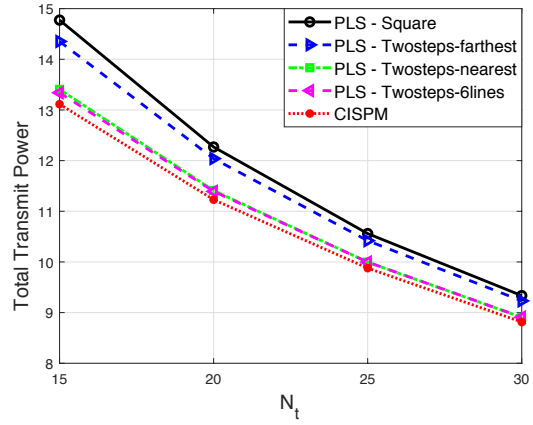
(b) 16-QAM with $\gamma = 20$ dB

Figure 5.10: Rate/power efficiency vs. number of antennas at Eve, with $N_\mathrm{t} = 15$, $K = 6$, and $\delta = 0.1$.

keeping the same ranking, with PLS Square scheme providing the highest efficiency. Similarly, for the case of 16-QAM, we observe the same pattern, PLS schemes outpeforming the CISPM scheme where the order of EEST performance amongst PLS schemes is maintained the same as the order of BER performance in Figure 5.6 with PLS Square scheme achieving the highest efficiency.



(a) QPSK with $\gamma = 5$ dB

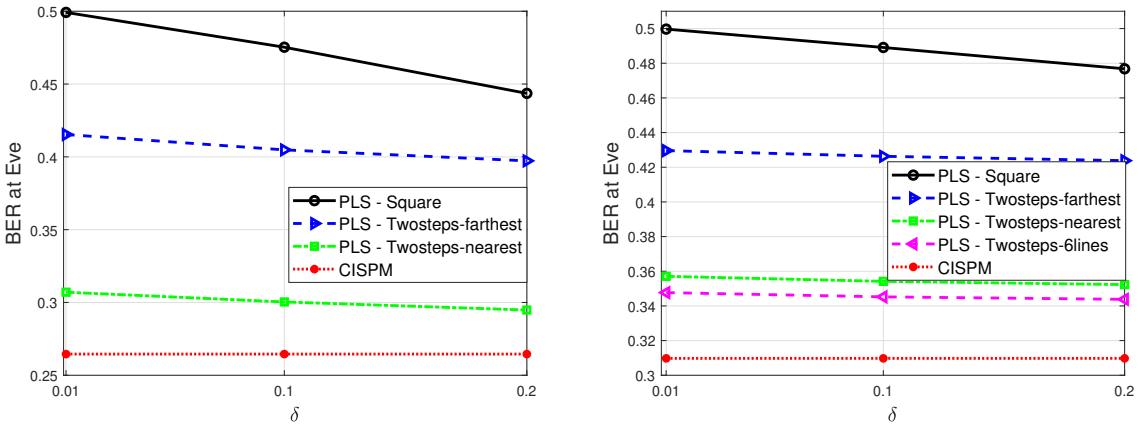


(b) 16-QAM with $\gamma = 15$ dB

Figure 5.11: Total transmit power vs. N_t, with $K = 6$, $\delta = 0.1$, and $M = 3$.

Figure 5.11 depicts the total transmit power, in dB, as a function of the number of antennas N_t, for a fixed target SINR of 5 dB for QPSK and 15 dB for 16-QAM. We compare the CISPM scheme with the PLS schemes. The parameters used in the simulation are $N_\mathrm{t} =$

15, $K = 6$, $\delta = 0.1$, and $M = 3$. We observe that, for all the schemes, the power consumption decreases with the number of antennas. As a result, increasing the number of antennas, Nt, leads to higher power gains at the receivers, hence the less required power by the transmitter. In other words, increasing N_{t} leads to stronger inter-user interference, hence higher power gains. On the other hand, PLS schemes consumes more power than CISPM scheme, where increasing the number of antennas at Eve leads to even higher power consumption, i.e., more antennas at Eve imply more constraints given that the constraints are applied on a per-antenna basis, hence the more power required.



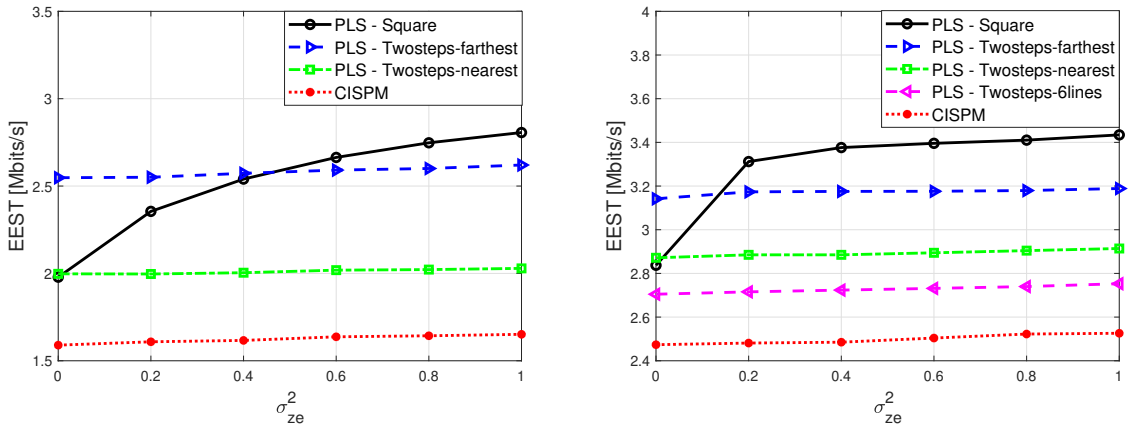
(a) QPSK with $\gamma = 5$ dB (b) 16-QAM with $\gamma = 15$ dB

Figure 5.12: BER vs. δ, with $N_{\mathrm{t}} = 15$, $K = 6$, and $M = 3$.

Figure 5.12 represent the BER at Eve, as a function of δ, for a fixed target SINR of 5 dB for case of QPSK and 15 dB for 16-QAM. We compare the CISPM scheme with the PLS schemes. The parameters used in the simulation are $N_{\mathrm{t}} = 15$, $K = 6$, and $M = 3$. Similarly, PLS schemes outperform CISPM scheme, with PLS Square providing highest security gains. We observe that for all PLS schemes, the BER at Eve decreases as δ gets smaller. This can be explained intuitively as follows. The larger the δ, the thicker the boundary region, hence the more chances for constellations points to fall into a deeper position inside the detection region, in this case, noise will have a smaller chance on pushing this to the opposite region as opposed to the case where the boundary region is very thin, thus the smaller the δ, the higher the BER (more security gains). We note that same pattern is observed for both constellations, QPSK and 16-QAM.

In all of the below results, we considered a noisy channel of Eve. In the below simulation,

however, we investigate the case where Eve has different SNR levels. For instance, in the case where Eve is very close to the BS, the received signal at the later will be strong, and vice-versa in the case when Eve is far from the BS. Similarly, if Eve uses sophisticated radio-frequency (RF) hardware, then the noise variance at the latter might be negligible.



(a) QPSK with $\gamma = 10$ dB

(b) 16-QAM with $\gamma = 20$ dB

Figure 5.13: Rate/power efficiency vs. σ^2_{ze}, with $N_{\mathrm{t}} = 15$, $K = 6$, $\delta = 0.1$, and $M = 3$.

Figure 5.13 plots η_{eest}, in [Secure bits/s], as a function of the noise variance at Eve σ^2_{ze}, for the case of QPSK and 16-QAM modulations. We compare the benchmark scheme with the PLS schemes. The parameters used in the simulation are $N_{\mathrm{t}} = 15$, $K = 6$, $\delta = 0.1$, and 10 dB target SINR for QPSK and 20 dB for 16-QAM. For both QPSK and 16-QAM, it is observed that,in the case where Eve has almost noiseless receive signal, the PLS - two-step-farthest scheme outperforms the PLS - Square scheme in both BER and η_{eest}, and even consumes less power. This behavior can be explained in the sense that the boundary schemes strength lies in the assumption that the noise at Eve would push it to either detection regions with almost equal probability, therefore in the case of noiseless Eve, noise is no longer there to move the receive signal, and by design the PLS - two-step-farthest scheme feasible region is bigger than the PLS - Square scheme, therefore it is harder for the ML algorithm to track the Two-steps-farthest than the Square scheme. However, in the case of noisy Eve, the Square scheme performs better in BER and η_{eest} because noise will push it not only to the opposite detection region but also the other neighbouring ones, hence the increase in security. Concurrently, in all schemes we observe that η_{eest} increases with the increase of σ^2_{ze} as a result of the noise that further distorts Eve's received signals away. More noise makes it harder for the ML

algorithm to track the mapping, hence the higher BER at Eve and consequently higher η_{eest}.

5.5 Summary

In this chapter, we investigated the problem of one-shot MIMO detection robust to inaccurate CSIT in MU-MISO systems without channel coding. In this setting, we proposed a new ML-based attack that permits a sophisticated eavesdropper to detect a message in a downlink MU-MISO system with a decent accuracy. The Eve learns patterns from the sent precoded pilots and predicts data symbols accordingly, where this sophisticated Eve employs several antennas and has ability to detect multi-level modulation schemes. We showed that this vulnerability is valid even when conventional SLP based precoding is employed. Still, these conventional precoders, such as CISPM scheme, have the advantage of not requiring the knowledge of Eve's channel. As a countermeasure to this attack, we propose novel SLP-based precoders. In general, the Square scheme provides the highest security gains and also computationally wise, it consumes almost half of the computation time than the other two-steps schemes, however it consumes the highest transmit power. However, as shown in the numerical results, depending on the modulation used, the number of antennas at Eve and the noise power at the latter, both Two-steps-farthest and nearest can outperform the Square scheme. Therefore, the proposed PLS schemes provide different trade-offs between security, computation time, and transmit power, which would give the BS options to choose the most suitable scheme depending on level of security required and/or transmit power needed and parameters used. Notably, despite all the security gains offered by the PLS schemes, their use does not affect the performance at the intended user. Numerical results validate both the attack as well as the countermeasures, where the proposed PLS precoders achieve drastic security gains at the expense of only a small additional power consumption at the transmitter. Future research topics would be to extend this chapter to the case of non-perfect CSI and also where the channel to the Eve is unknown to the BS.

Chapter 6

Multi-Antenna Data-Driven Eavesdropping Attacks and Symbol-Level Precoding Countermeasures for Coded MU-MISO Systems

In this chapter, we investigate MIMO detection at a multi-antenna eavesdropper robust to CSIT deterioration in coded systems. Particularly, we consider secure communications in wireless MU-MISO systems with channel coding in the presence of a multi-antenna Eve who is part of the MU-MISO system. In this setting, we exploit ML tools to design soft and hard decoding schemes by using precoded pilot symbols as training data. The proposed ML frameworks allow an Eve to determine the transmitted message with high accuracy. We thereby show that MU-MISO systems are vulnerable to such eavesdropping attacks even when relatively secure transmission techniques are employed, such as SLP. To counteract this attack, we propose two novel SLP-based schemes that increase the bit-error rate at Eve by impeding the learning process. We design these two security-enhanced schemes to meet different requirements regarding runtime, security, and power consumption. Simulation results validate both the ML-based eavesdropping attacks as well as the countermeasures, and show that the gain in security is achieved without affecting the decoding performance at

the intended users.

The rest of the chapter is organized as follows: Section 6.1 describes the system model. In Section 6.2, we introduce the ML-based attacks, whereas in Section 6.3, we propose our novel SLP-based schemes as countermeasures to this attack. Simulation results are discussed in Section 6.4, followed by the summary in Section 6.5.

6.1 System Model

We consider coded single-cell MU-MISO downlink system which channel coding, which, as depicted in Figure 5.1, consists of a BS equipped with N_t transmit antennas, K single-antenna users such that $K \leq N_\mathrm{t}$, and one multi-antenna Eve with M antennas. The received coded signal by the k-th user at the symbol slot n can be expressed as

$$y_k[n] = \mathbf{h}_k \mathbf{x}_\mathrm{d}[n] + z_k[n], \tag{6.1}$$

where $\mathbf{x}_\mathrm{d}[n] \in \mathbb{C}^{N_\mathrm{t} \times 1}$ is the transmitted coded vector from the N_t transmit antennas, $\mathbf{h}_k \in \mathbb{C}^{1 \times N_\mathrm{t}}$ is the channel between the transmit BS antennas and the k-th user, and $z_k[n] \in \mathbb{C}$ is the AWGN at the k-th user with variance σ_z^2.

The above model can be rewritten in a more compact form by gathering the received signals at all users in vector $\mathbf{y}[n] \in \mathbb{C}^{K \times 1}$ as

$$\mathbf{y}[n] = \mathbf{H}^\mathrm{T} \mathbf{x}_\mathrm{d}[n] + \mathbf{z}[n], \tag{6.2}$$

where $\mathbf{H} = [\mathbf{h}_1^\mathrm{T} \dots \mathbf{h}_K^\mathrm{T}] \in \mathbb{C}^{N_\mathrm{t} \times K}$ defines the MU-MISO system channel matrix and $\mathbf{z}[n] \in \mathbb{C}^{K \times 1}$ collects the independent AWGN components of all users.

Likewise, the received signal at Eve, $\mathbf{y}_\mathrm{e}[n] \in \mathbb{C}^{M \times 1}$, can be expressed as

$$\mathbf{y}_\mathrm{e}[n] = \mathbf{H}_\mathrm{e}^\mathrm{T} \mathbf{x}_\mathrm{d}[n] + \mathbf{z}_\mathrm{e}[n], \tag{6.3}$$

where $\mathbf{H}_\mathrm{e} = [\mathbf{h}_{\mathrm{e},1}^\mathrm{T} \dots \mathbf{h}_{\mathrm{e},M}^\mathrm{T}] \in \mathbb{C}^{N_\mathrm{t} \times M}$ is the channel matrix between the BS and Eve with $\mathbf{h}_{\mathrm{e},i} \in \mathbb{C}^{1 \times N_\mathrm{t}}$ representing the channel between the BS antennas and Eve's ith receive antenna, and $\mathbf{z}_\mathrm{e}[n] \in \mathbb{C}^{M \times 1}$ assembles the independent AWGN components at the M antennas of variance σ_e^2 each.

We note that the pilot symbols, also being referred to as reference signals, are an integral

part of communication systems that are known entities to all parties. In particular, they are commonly used for CSI and SINR estimation. Specifically, non-precoded pilot symbols are used for CSI estimation while precoded pilot signals are intended for SINR estimation [ETS14]. In this work, we are interested in the latter case, precoded pilot symbols, which uses the same MCS used for precoding the data. In this context, we define N as the number of precoded pilot symbols used within a frame. We also note that these N pilot symbols are interleaved with data symbols in a frame that fits within the channel coherence time T. In this setting, we define the input data symbols intended for the K users as $\mathbf{d} \in \mathbb{R}^{K\times 1}$, with d_k being the symbol intended for user k.

In the case of block-level precoding, we define η as the mean power. For the SLP case, we define γ_k as the target SINR for the k-th user with $\boldsymbol{\gamma} = [\gamma_1 \dots \gamma_K] \in \mathbb{R}^{K\times 1}$ representing the target SINR for all users. For ease of notation, we drop the time index n in the remainder of the chapter.

6.2 ML-Based Attacks

In this section, we will propose two ML eavesdropping attacks, where a multi-antenna Eve uses precoded pilot symbols as training data to accurately hard/soft decode the transmitted symbols. We start by presenting the motivation of our work along with the adversarial model. Next, we present the ML frameworks for the proposed soft and hard decoding schemes. We note that our proposed ML framework is valid in all cases where the transmitter sends also pilot symbols, which is actually the case for a standard downlink MU-MISO system, the one considered in this chapter.

6.2.1 Motivation

To motivate our work, we study the received signal at Eve when the BS sends precoded pilot signals to the intended users. We investigate the case when the BS uses a conventional block-level precoder, i.e., ZF [TG06] as well as the case of a conventional SLP precoder, i.e., the CISPM approach in [ACO15c].

To that end, we first examine a special case scenario for illustration purposes. Afterwards, we present a more general scenario that represents a typical downlink MU-MISO system.

In the illustrative special case scenario, we consider the following toy example: an MU-

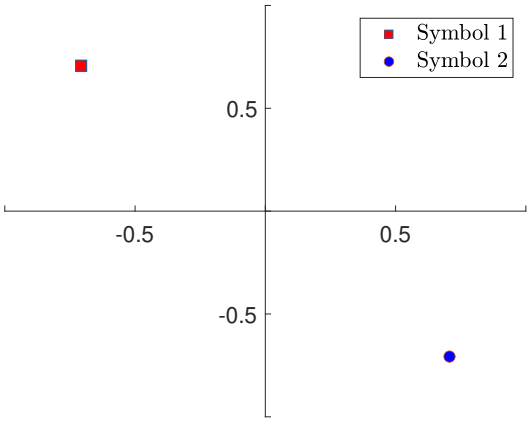


Figure 6.1: Symbols used in the two pilot signals intended for user k.

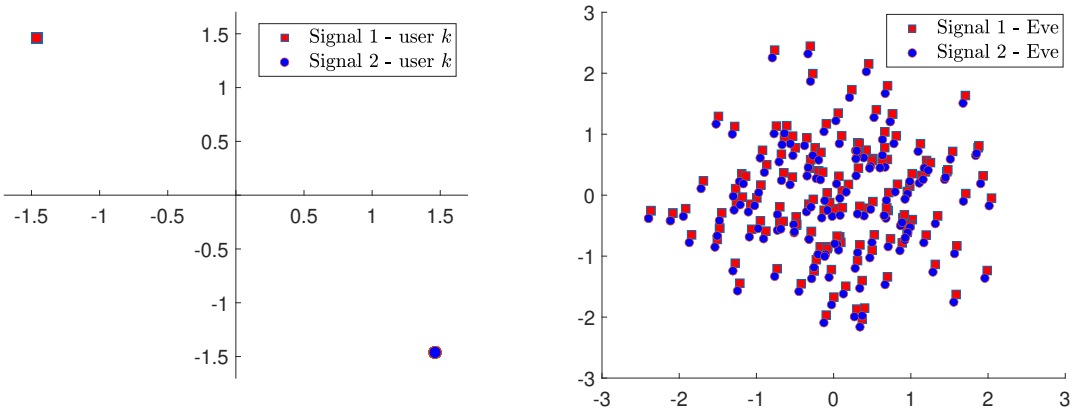
MISO system with $N_{\mathrm{t}} = 15$, $K = 6$, $\sigma_z^2 = 1$, one channel realization, and QPSK as a modulation scheme, where the BS sends to each user two precoded pilot signals of $N = 150$ symbols each. In this setting, a multi-antenna Eve attempts to eavesdrop a specific user k.

To better understand the example visually, the BS sends the same symbols to user k while it sends pseudo-random sequences to the remaining users. The two symbols constructing the two pilot signals intended for user k are plotted in Figure 6.1.

When the BS precodes the aforementioned pilot signals with ZF precoding of mean power of 5 dB, the noiseless received signals at user k and the Eve are respectively given as in Figure 6.2(a) and Figure 6.2(b), respectively. We note that the channel to the K users and Eve were generated randomly.

As depicted in Figure 6.2(a), the received signal at user k shows no inter-user interference as it was cancelled by the ZF precoder. However, the received signal at Eve is spread due to the inter-user interference effect, as Eve's channel is different from user's k channel. Still, we can observe a precoding pattern that applies to both received signals, i.e., the red squares are mostly positioned on the top right of the blue circles.

A more inherently-secure precoding scheme, which does not depict patterns of the precoding used, is SLP precoding [ACO15c]. This particular SLP scheme is designed to exploit the multi-user interference for power gains. In other words, this scheme propels the intended users' received signals deeper into the correct detection region of the desired symbol for each intended user. The corresponding optimization problem is defined as



(a) Noiseless received signals at user k (b) Noiseless received signals at Eve

Figure 6.2: Noiseless received signals at user k and Eve when the BS uses ZF precoding with a mean power of 5 dB.

$$\mathbf{x}_{\mathrm{d}}(\mathbf{d},\mathbf{H},\boldsymbol{\gamma}) = \arg\min_{\mathbf{x}} ||\mathbf{x}||^2 \tag{6.4}$$

$$\text{subject to}$$

$$\mathrm{Re}\{\mathbf{h}_k\mathbf{x}\} \trianglelefteq \sigma_z\sqrt{\gamma_k}\mathrm{Re}\{d_k\},\ \forall k$$

$$\mathrm{Im}\{\mathbf{h}_k\mathbf{x}\} \trianglelefteq \sigma_z\sqrt{\gamma_k}\mathrm{Im}\{d_k\},\ \forall k,$$

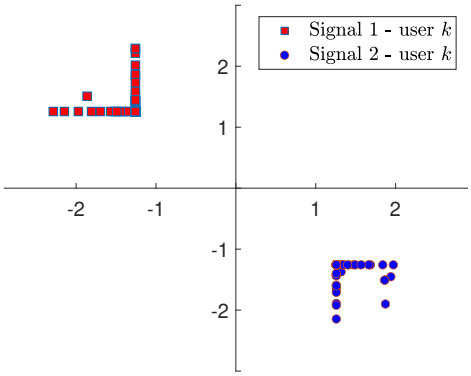
where the operator $\trianglelefteq$ denotes[1] the correct detection region [ACO17]. As shown in eq. (6.4), the CISPM scheme minimizes the transmit power while guaranteeing a certain target SINR at the intended users through constructive interference constraints.

Thus, the CISPM precoding takes inputs: the channel to the intended users, $\mathbf{H}$, the input data to be transmitted to the intended users, $\mathbf{d}$, the target SINR for all intended users, $\boldsymbol{\gamma}$, and the noise variance at the users σ_z^2.

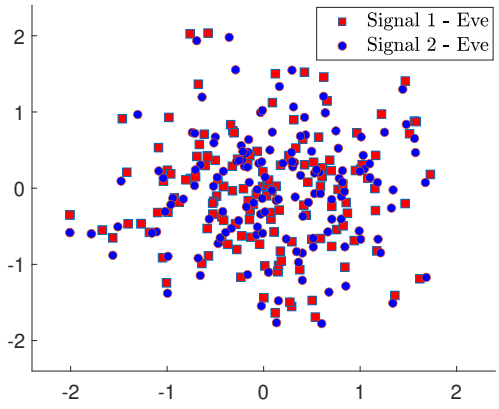
Now we consider the aforementioned toy example of transmitted pilot signals illustrated in Figure 6.1, but with the BS using CISPM precoding with a target SINR value of 5 dB for each user. The corresponding results of this example are illustrated in Figure 6.3.

Figure 6.3(a), represents the noiseless received signal at user k. As expected when using SLP precoding, the inter-user interference is transformed into power gains, resulting in deviations of the received signal deeper into the detection region while guaranteeing a specific target SINR value. The noiseless received signal at Eve when the aforementioned signals are

[1]For further detailed information, the reader should refer to [LSK+20, ASK+18].



(a) Noiseless received signals at user k



(b) Noiseless received signals at Eve

Figure 6.3: Noiseless received signals at user k and Eve when the BS uses CISPM precoding.

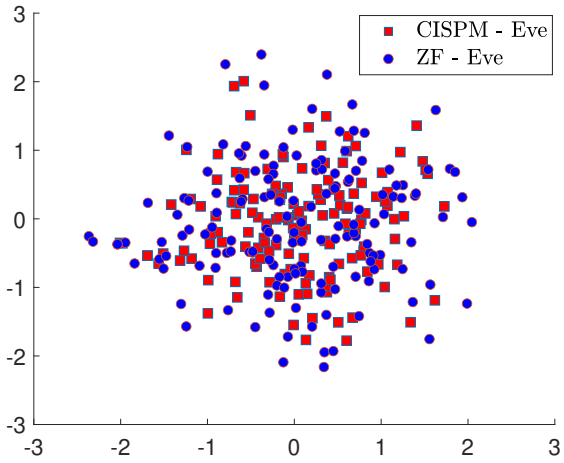


Figure 6.4: Received signal at Eve when BS sends pseu-random sequences to every user.

transmitted is plotted in Figure 6.3(b). As opposed to Figure 6.2(b) where the blue circles are always below and to the left of the red boxes, in Figure 6.3(b) there is no apparent fixed pattern.

This discrepancy is due to the randomness of the input data to transmit $\mathbf{d}$, which changes at each symbol slot resulting in $\mathbf{x}_{\mathrm{d}}$ to vary accordingly. Even though Eve is trying to eavesdrop user k whose received symbols do not change, what Eve receives provides no apparent insights about what was transmitted.

In a more general scenario representing a typical downlink MU-MISO system, the BS is not constrained to send pilot signals where one user's pilots are constructed with the same

symbol. In this case, the received signal will exhibit even more apparent randomness, as illustrated in Figure 6.4, since the actual pilots are pseudo-random sequences for all users. Hence, since Eve does not know the CSI, using conventional detection techniques to directly decode the received signals will result in a poor performance.

Thus, in this chapter we propose to use ML to model the non-linear mappings underlying this apparent randomness, so that Eve can decode with an acceptable BER. Since the symbols used for the precoded pilots are known in communication standards to all parties, we propose to use ML to leverage this knowledge and decode the transmitted data to a particular user with decent accuracy. To that end, we propose ML-based soft and hard decoding schemes that can accurately decode the transmitted signal by using the precoded pilot symbols as training data.

6.2.2 Adversarial Model

Contrary to the adversarial wiretap channel model proposed in [WS16] that considers active adversaries, herein, we consider a passive eavesdropper that can listen to the wireless medium with fading and noise effects. Specifically, we consider a passive Eve in the sense that Eve does not interfere with the communication channel, whereas an active Eve can jam as well as eavesdrop.

Concerning Eve's environment, Eve is part of the downlink MU-MISO system that comprises of the sender, i.e., the BS, intended users, and Eve. In this context, we assume that the BS knows Eve's CSI $\mathbf{H}_{\mathrm{e}}$. However, we highlight that Eve does not know $\mathbf{H}_{\mathrm{e}}$ nor $\mathbf{H}$. Since Eve is a registered user of the MU-MISO system, she knows the pilot symbols transmitted, the modulation scheme used, and the FEC parameters.

As for Eve's profile and capabilities, we consider Eve to have: 1) unlimited computation power, and 2) access to state-of-the-art machine-learning tools and algorithms. Contrary to the intended users who are single-antennas receivers, Eve is equipped with M antennas.

Next we present our proposed ML-based soft and hard decoding schemes. For a thorough explanation of our proposed ML-based decoding schemes, in the following, we use ZF precoding as an example. However, the proposed decoding frameworks are valid for any precoding techniques.

6.2.3 ML Framework for the Proposed Soft Decoding Scheme

As illustrated in Figure 6.5, the ML framework for soft-decoding encompasses two steps: 1) training phase, where the ML model is trained by using the precoded pilot symbols; 2) inference phase, where probabilities are estimated and employed to calculate the LLRs which are consequently fed to a soft decoder.

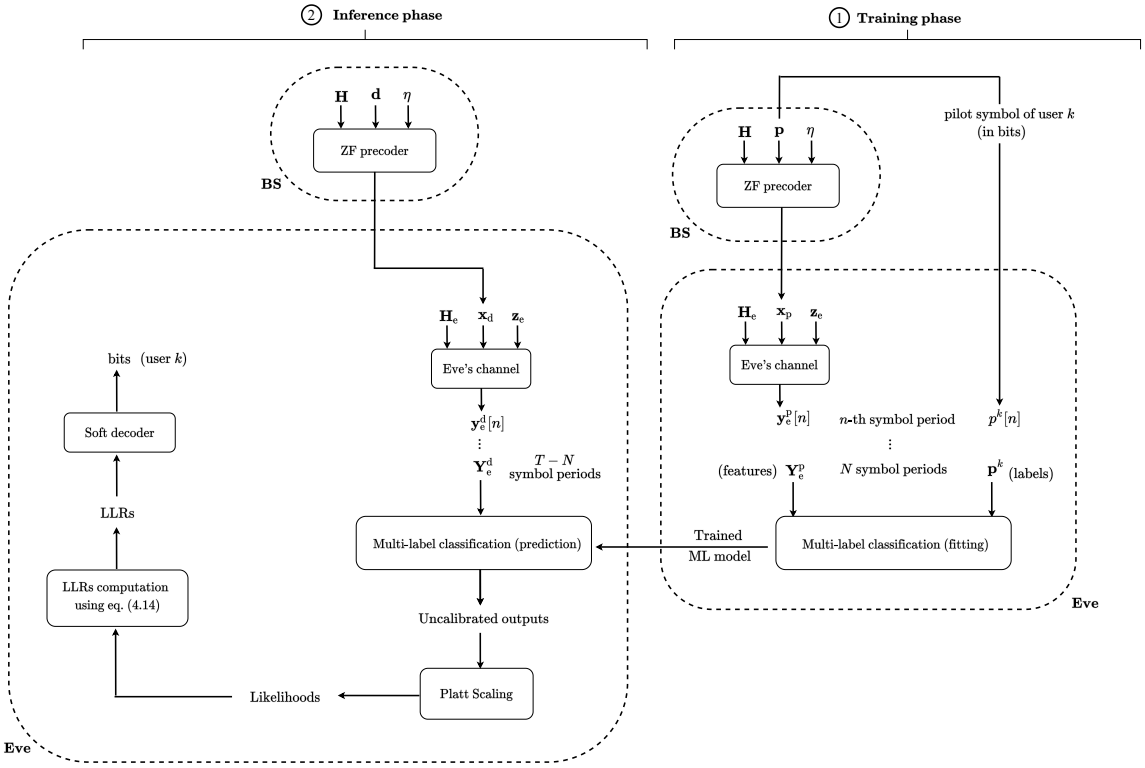


Figure 6.5: Overview of the ML-based soft decoding scheme.

Training Phase

As pointed out earlier, the BS sends the pilot symbols $\mathbf{p} \in \mathbb{C}^{K\times 1}$ as training data, which are pseudo-random sequences for all users. For one SP, the overall received pilot signal at Eve's all antennas, $\mathbf{y}_{\mathrm{e}}^{\mathrm{p}} \in \mathbb{C}^{M\times 1}$, can be written as

$$\mathbf{y}_{\mathrm{e}}^{\mathrm{p}} = \mathbf{H}_{\mathrm{e}}^{\mathrm{T}}\mathbf{x}_{\mathrm{p}} + \mathbf{z}_{\mathrm{e}}, \tag{6.5}$$

where $\mathbf{x}_{\mathrm{p}} \in \mathbb{C}^{N_{\mathrm{t}}\times 1}$ is the transmitted precoded pilot signal from the N_{t} BS's transmit antennas.

As depicted in Figure 6.5, the transmitted signal $\mathbf{x}_{\mathrm{p}}$ depends on all the users' symbols.

Thus, Eve could target any user individually by retraining the ML model according to the pilot sequences used for each user. Herein, Eve creates a mapping between the received signal $\mathbf{y}_{\mathrm{e}}^{\mathrm{p}}$ and the pilot symbols p_{s}^{k} corresponding to the targeted user k. We note that the subscript s in p_{s}^{k} stands for "soft", where the pilot symbols are represented in bits, i.e., $p_{\mathrm{s}}^{k} \in \{\text{“00”}, \text{“01”}, \text{“11”}, \text{“10”}\}$ in the case of QPSK modulation.

We note that, the higher the M, the higher the number of received signals at Eve, and thus the better is the detection performance. Basically, each antenna at Eve receives a different distorted version of the same transmitted signal $\mathbf{x}_{\mathrm{p}}$, the more received replicas of the same transmitted signal, the better the decoding performance.

Hence, the training set $\mathcal{D}_{\mathrm{s}}$ is the collection of $\{\mathbf{y}_{\mathrm{e}}^{\mathrm{p}}[n], p_{\mathrm{s}}^{k}[n]\}, n \in \{1 \dots N\}$, where $\mathbf{y}_{\mathrm{e}}^{\mathrm{p}}[n]$ represents the received pilot signal at Eve during the n-th SP, while $p_{\mathrm{s}}^{k}[n]$ is the corresponding pilot symbol of user k. Therefore, the training set $\mathcal{D}_{\mathrm{s}}$ can be written in a more compact form as

$$\mathcal{D}_{\mathrm{s}} = \{\mathbf{Y}_{\mathrm{e}}^{\mathrm{p}}, \mathbf{p}_{\mathrm{s}}^{k}\}, \tag{6.6}$$

where $\mathbf{Y}_{\mathrm{e}}^{\mathrm{p}} \in \mathbb{C}^{N \times M}$ are the received pilot symbols at Eve during N SPs and $\mathbf{p}_{\mathrm{s}}^{k} \in \mathbb{C}^{N \times 1}$ are the corresponding transmitted pilot symbols to the k-th user. Using ML terminology, $\mathbf{Y}_{\mathrm{e}}^{\mathrm{p}}$ represents the features while $\mathbf{p}_{\mathrm{s}}^{k}$ represents the labels, where both constitute the training dataset. We note that the input features in $\mathbf{Y}_{\mathrm{e}}^{\mathrm{p}}$ are complex-valued and cannot be directly processed by ML algorithms in general. Usually, this is addressed by considering real and imaginary parts separately. For non-binary modulation schemes, this ML problem is considered as a MLC problem [ZLLG18, YWF+15] as more than 1 bit is required to encode the symbols. MLC is a supervised learning problem where an observation, i.e, a scalar or a vector of features, is associated with multiple labels. Hence, as depicted in Figure 6.5, the training dataset is fed to the MLC fitting module which will in turn output a trained ML model, that will subsequently be used in the inference phase.

Inference Phase

As depicted in Figure 6.5, in each SP, the BS sends the symbols $\mathbf{d} \in \mathbb{C}^{K \times 1}$ to the K users after precoding them using the same precoding scheme employed in the previous phase. The received signals at Eve in each SP, $\mathbf{y}_{\mathrm{e}}^{\mathrm{d}} \in \mathbb{C}^{M \times 1}$, can be written as

$$\mathbf{y}_{\mathrm{e}}^{\mathrm{d}} = \mathbf{H}_{\mathrm{e}}^{\mathrm{T}}\mathbf{x}_{\mathrm{d}} + \mathbf{z}_{\mathrm{e}}, \tag{6.7}$$

where $\mathbf{x}_{\mathrm{d}} \in \mathbb{C}^{N_{\mathrm{t}} \times 1}$ represents the transmitted precoded data from the N_{t} transmit antennas, intended for all the users during one SP. If we assume that there are T symbols in one coherence time, $\mathbf{Y}_{\mathrm{e}}^{\mathrm{d}} \in \mathbb{C}^{(T-N)\times M}$ represents the collection of all received signals at Eve during one coherence time of the transmitted $(T - N)$ data symbols. In ML nomenclature, $\mathbf{Y}_{\mathrm{e}}^{\mathrm{d}}$ is commonly being referred to as the test/evaluation dataset.

In principle, the goal of classification is to predict labels. In this context, however, we are not interested in the labels (hard outputs) but rather in the corresponding probabilities (soft outputs) to be used subsequently for LLR computation, as previously described in Section 4.2.1. As depicted in Figure 6.5, once the LLRs are computed, Eve can simply feed the computed LLRs to the soft decoder to obtain the transmitted message to user k.

We note that, for the soft decoding scheme, the multi-antenna eavesdropping attack was modeled as a classification problem, given the nature of the pilot data. However, since the LLRs are continuous, using regression instead of classification to model the eavesdropping attack might yield better detection performance. To evaluate this claim, in Chapter 8 we re-reformulate the eavesdropping attack using regression and present the corresponding performance evaluation results.

6.2.4 ML Framework for the Proposed Hard Decoding Scheme

The proposed ML-based hard decoding scheme also comprises of two phases: 1) training phase, where the ML model is trained by using the precoded pilot symbols, 2) inference phase, where the module directly predicts the transmitted symbols to a particular user, which are in turn mapped into bits to finally be fed to a hard decoder to obtain the transmitted bits to user k.

Training Phase

As depicted in Figure 6.6, for each SP, the BS first sends pilot symbols $\mathbf{p} \in \mathbb{C}^{K\times 1}$ to the K users, which after precoding become the signal $\mathbf{x}_{\mathrm{p}} \in \mathbb{C}^{N_{\mathrm{t}}\times 1}$. Eve receives $\mathbf{y}_{\mathrm{e}}^{\mathrm{p}}[n] \in \mathbb{C}^{M\times 1}$ from all its antennas at the n-th SP, as in eq. (6.5).

For Eve to eavesdrop user k, it creates a mapping between the received signal $\mathbf{y}_{\mathrm{e}}^{\mathrm{p}}$ and the

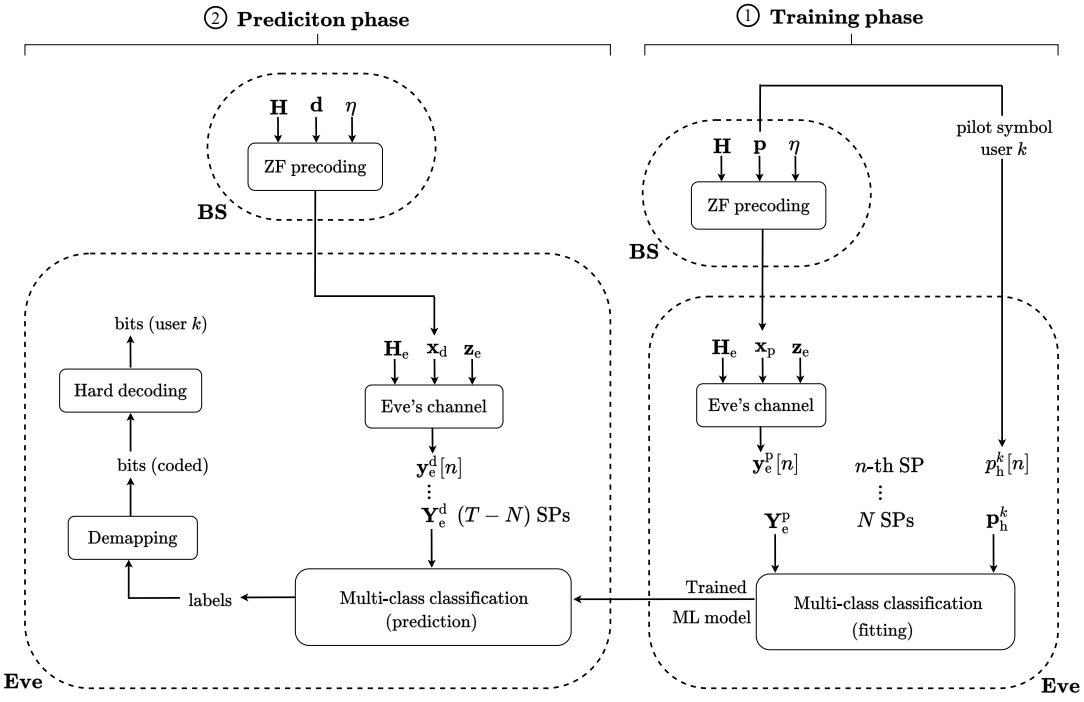


Figure 6.6: Overview of the ML-based hard decoding scheme.

pilot symbols p_{h}^k corresponding to the targeted user k. We note that the subscript h in p_{h}^k stands for "hard", which defines the decimal representation of the pilot symbols, for instance, $p_{\mathrm{h}}^k \in \{0, 1, 2, 3\}$ in the case of QPSK modulation.

Thus, the training set $\mathcal{D}_{\mathrm{h}}$ is the collection of $\{\mathbf{y}_{\mathrm{e}}^{\mathrm{p}}[n], p_{\mathrm{h}}^k[n]\}, n \in \{1 \dots N\}$, where $p_{\mathrm{h}}^k[n]$ is the pilot symbol of user k corresponding to the received pilot signal $\mathbf{y}_{\mathrm{e}}^{\mathrm{p}}[n]$ at the n-th SP. Thus,

$$\mathcal{D}_{\mathrm{h}} = \{\mathbf{Y}_{\mathrm{e}}^{\mathrm{p}}, \mathbf{p}_{\mathrm{h}}^k\}, \tag{6.8}$$

where $\mathbf{Y}_{\mathrm{e}}^{\mathrm{p}} \in \mathbb{C}^{N \times M}$ are the received pilot signals at Eve and $\mathbf{p}_{\mathrm{h}}^k \in \mathbb{C}^{N \times 1}$ are the transmitted pilot symbols, represented in decimals, to the k-th user, both during N SPs.

For non-binary modulation schemes, this problem is a single-label multi-class classification (MCC) problem. MCC is a supervised learning problem where an observation, i.e, a scalar or a vector of features, is associated with a single-label with multiple classes. Considering a modulation order M_{o}, the label space is $\{0 \dots M_{\mathrm{o}} - 1\}$ with M_{o} total classes. Thus, as depicted in Figure 6.6, the training dataset $\mathcal{D}_{\mathrm{h}}$ is fed to the MCC fitting module that in sequence outputs a trained ML model, which will be used thereafter in the inference phase.

Inference Phase

As depicted in Figure 6.6, the BS transmits $\mathbf{d} \in \mathbb{C}^{K\times 1}$ data symbols to the K users in each SP in the form of the precoded signal $\mathbf{x}_{\mathrm{d}} \in \mathbb{C}^{N_{\mathrm{t}}\times 1}$. We note that the precoding at BS herein uses the same precoding scheme employed in the training phase. The corresponding received signals at Eve is $\mathbf{y}_{\mathrm{e}}^{\mathrm{d}} \in \mathbb{C}^{M\times 1}$, as detailed in eq. (6.7).

Considering a coherence time of T symbols, there are $(T-N)$ symbols dedicated to data transmission. Thus $\mathbf{Y}_{\mathrm{e}}^{\mathrm{d}} \in \mathbb{C}^{(T-N)\times M}$ represents the collection of all received data signals at Eve during one coherence time, which constitutes the test/evaluation dataset.

Contrary to the proposed soft decoding scheme, herein we are interested in predicting the labels, i.e., hard outputs. As depicted in Figure 6.6, to obtain the labels, we feed the test dataset and the trained ML model to the MCC prediction module. We note that the labels are in the form of decimals, i.e., the same nature of the labels used in the training phase. Once the labels are predicted, they will be first mapped into bits and then fed to the hard decoder for decoding to finally obtain the bits transmitted to user k.

We note that the proposed soft and hard decoding schemes are also valid for other constellations, including higher-order quadrature amplitude modulation (QAM). For instance, in 16-QAM, 4 bits are needed to represent the symbols as opposed to 2 in the QPSK case. Consequently, for the soft scheme in Figure. 6.5, the pilot symbols of user k will contain 4 bits instead and the prediction module will generate 4 uncalibrated outputs as a result. However, for the hard scheme in Figure. 6.6, the pilot symbols of user k will be represented using 16 classes, and therefore the prediction module will output a label in the set $\{0, 1, \ldots, 15\}$. Therefore, the proposed soft and hard decoding frameworks are valid for any constellation, where the modulation order determines the number of bits used in pilot/data symbols and also defines the number of labels/classes employed, all the rest of the processing remain unchanged.

6.3 Countermeasure: PLS Schemes

In this section, we propose novel security-enhanced SLP schemes that yield high BER at Eve. Similar to [MST+19b] and [MST+20], the idea is to design the transmitted signal $\mathbf{x}_{\mathrm{d}}$ to have constructive interference at the intended users, while at the same time, increasing the BER at Eve through destructive interference at the latter. We note that the CSI to Eve is

available at the BS.[2]

6.3.1 PLS Random Scheme

We design this scheme to have constructive interference at the intended users and destructive interference at Eve. To that end, we align the transmitted signal to the corresponding detection regions of the intended users using their CSI while we force Eve's received signal to lie at the boundaries of the detection regions using Eve's CSI.

In order to illustrate the idea, consider the QPSK constellation in Figure 5.3. The aim here is to propel the received signals deeper into the detection regions, which are illustrated by the gray squares (actually, those regions are unbounded).

On the other hand, we design Eve's received signal to lie in the strapped region, which is centered at the boundary of the detection regions. The strapped region's width is governed by the parameter δ. The smaller the δ, the tighter the strapped region, leading to higher probability of landing into the opposite region when noise adds up, resulting in higher BER at Eve.

Inspired by the boundary scheme presented in Chapter 5, Section 5.3.1, we propose to embed randomness in Eve's received signal by randomly selecting the boundary region, either horizontal or vertical, as depicted in Figure 5.3. We decompose the non-convex strapped region in Figure 5.3 into two convex regions: the vertical part and the horizontal one. And at each SP and for each antenna at Eve, we *randomly* choose between the two regions, so that on average, Eve's received signals would lie evenly on both regions, assuming the symbol distribution is equiprobable. We refer to this scheme as "PLS random scheme".

The optimization problem for this scheme can be formulated as

$$\mathbf{x}_{\mathrm{d}}(\mathbf{d},\mathbf{H},\mathbf{H}_{\mathrm{e}},\boldsymbol{\gamma},\delta,\mathbf{b}) = \arg\min_{\mathbf{x}} \|\mathbf{x}\|^2 \tag{6.9}$$

subject to

$$\mathrm{Re}\{\mathbf{h}_k\mathbf{x}\} \trianglelefteq \sigma_z\sqrt{\gamma_k}\mathrm{Re}\{d_k\},\ \forall k \tag{6.10}$$

$$\mathrm{Im}\{\mathbf{h}_k\mathbf{x}\} \trianglelefteq \sigma_z\sqrt{\gamma_k}\mathrm{Im}\{d_k\},\ \forall k \tag{6.11}$$

$$b_i\mathrm{Re}\{y_{\mathrm{e}}^i\}+(1-b_i)\mathrm{Im}\{y_{\mathrm{e}}^i\} \trianglerighteq \delta, \forall i \tag{6.12}$$

[2] We follow the same methodology used in the relevant PLS work in [LM13, FS11, LCMC11], which assumed the knowledge of Eve's CSI at the transmitter. However, we intend to extend our work to a more general eavesdropper, by considering no Eve's CSI at the BS [KMW18].

where $\mathbf{h}_k\mathbf{x}$ is the k-th user's noiseless received signal, $y_{\mathrm{e}}^i = \mathbf{h}_{\mathrm{e},i}\mathbf{x}$ is the noiseless received signal at Eve's i-th antenna, $b_i \in \{0, 1\}$ is the realization of a binary RV that represents the boundary region to use for Eve's i-th antenna, $\mathbf{b}$ is the vector collecting the b_i realizations, which select the corresponding strapped sub-region in Figure 5.3 — either vertical ($b_i = 1$) or horizontal ($b_i = 0$) —, of the M antennas at Eve, $\trianglelefteq$ defines the correct[3] detection region, and $\delta > 0$ is the distance parameter controlling the width of the boundary region. This problem is convex and can be solved efficiently using standard optimization toolboxes such as CVX [BV04]. Algorithm 5 details the process of signal design of the PLS random scheme, which is executed for every symbol slot.

Algorithm 5 PLS random scheme

Input: $\mathbf{d}, \mathbf{H}, \mathbf{H}_{\mathrm{e}}, \boldsymbol{\gamma}, \sigma_z^2, \delta, \mathbf{b}$;
Solve problem (6.9) **as follows:**
Satisfy CI constraints in (6.10) and (6.11)
For Each antenna i at Eve,
If $b_i = 1$, push y_{e}^i to the vertical boundary region
Else, push y_{e}^i to the horizontal boundary region
Output: $\mathbf{x}_{\mathrm{d}} = 0$

As demonstrated in Algorithm 5, the PLS random scheme designs Eve's received signal to have no pattern as to which boundary region it will fall into, thus it will be extremely difficult for a ML-based engine to find relationships between the known pilot symbols and Eve's received signals.

The advantages of this scheme are: 1) enhance PLS as the detection decisions at the Eve will be mostly made on the basis of noise, and 2) save the transmit power since this scheme involves only a small deviations of the targeted received constellation points at Eve.

6.3.2 PLS Eve-Min-Power Scheme

Targeting a computationally simpler scheme, herein we take the basic SLP scheme for constructive interference at the intended users, the CISPM, and change the cost function to achieve PLS. Specifically, we do not include any PLS related constraints. To attain security with the CISPM scheme, in addition to minimizing the transmit power, we also minimize the power of Eve's noiseless received signal simultaneously, hence the name "PLS Eve-min-power".

[3] For further detailed information, the reader should refer to [LSK+20, ASK+18, SACO18].

Compared to the PLS random scheme that has the strict constraint of forcing Eve's received signal to lie in the boundary region, in the PLS Eve-min-power, we allow Eve's signal to lie anywhere in the plane but as close to zero as possible, whatever the degrees of freedom at the transmitter allow. This scheme can be formulated as follows:

$$\mathbf{x}_{\mathrm{d}}(\mathbf{d},\mathbf{H},\mathbf{H}_{\mathrm{e}},\boldsymbol{\gamma})=\arg\min_{\mathbf{x}_{\mathrm{d}}} ||\mathbf{x}_{\mathrm{d}}|| + ||\mathbf{H}_{\mathrm{e}}\mathbf{x}_{\mathrm{d}}|| \tag{6.13}$$

subject to

$$\mathrm{Re}\{\mathbf{h}_k\mathbf{x}_{\mathrm{d}}\} \trianglelefteq \sigma_z\sqrt{\gamma_k}\mathrm{Re}\{d_k\},\ \forall k \tag{6.14}$$

$$\mathrm{Im}\{\mathbf{h}_k\mathbf{x}_{\mathrm{d}}\} \trianglelefteq \sigma_z\sqrt{\gamma_k}\mathrm{Im}\{d_k\},\ \forall k. \tag{6.15}$$

As shown in (6.13), the PLS Eve-min-power scheme minimizes the sum of the transmit power and the noiseless received power at Eve, while guaranteeing a certain target SINR at the intended users through constructive interference constraints (6.14) and (6.15). Similarly, this problem is convex and can be solved efficiently using standard optimization toolboxes.

We note that this PLS scheme leads to higher security than the CISPM, as it makes it considerably harder for an ML-enabled Eve to correctly detect the low-power signal induced by minimizing the power at Eve.

We note that, low-complexity and computationally efficient SLP design was developed for practical and real-time implementations [KMA+19, HKD+21]. For instance, an FPGA-accelerated design of computationally efficient SLP for high-throughput communication systems was proposed in [KMA+19], which enables real-time operation and provides a high symbol throughput for multiple receive terminals. In [HKD+21], a low-complexity FPGA design for SLP was proposed for MU-MISO downlink communication systems, by developing an approximate yet computationally-efficient closed-form solution to alleviate the excessive complexity incurred by the SLP design.

To validate the proposed schemes, we have conducted numerical simulations according to the same methodology as [KMW18]. Future extensions of our work include adapting and optimizing the SLP technique for real-time validation similar to [KMA+19, HKD+21]. To that end, since the formulations of the proposed and benchmark SLP precoding schemes are convex, in the numerical results we use the CVX modeling framework to solve the SLP precoders' underlying optimization problems. CVX's employed solvers rely on primal-dual interior point methods to solve the problems. However, the time complexity analysis of the SLP precoders

employed in this paper depends heavily on the solver used and its implementation, which makes it challenging to obtain a closed-form big $\mathcal{O}$ representation of it. Nevertheless, we resort to *runtime* [SDW19] analysis of the implementations, where we measure the total time it takes for the algorithm to solve the optimization problem (in milliseconds). We present the runtime analysis of the proposed and benchmark schemes in the following section.

6.4 Simulation Results

To make this section more comprehensive, we split it into three parts: 1) Parameters, metrics, and benchmarks where we defined the simulations' setting, 2) selection of ML algorithms for Eve attack in which we experiment with several algorithms and select the most performing, and 3) comparisons and insights to assess the performance of our proposed schemes in terms of security, power consumption, and runtime.

6.4.1 Parameters, Metrics, and Benchmarks

As a benchmark to the proposed SLP-based PLS schemes, we employed the CISPM [ACO15c] and ZF precoding [TG06] schemes. We note that in the following simulations, for a fair comparison, we set $\eta = \gamma_k$ such that all the examined schemes have the same transmit power.

Regarding the metrics used to evaluate the different schemes, we use the BER at Eve to assess the security offered by a particular decoding scheme for a given precoding design. The lower the BER at Eve, the lower the security and vice versa. In a similar way, we also evaluate the frame-error rate (FER) at Eve since we have channel coding in the system, which is defined as the ratio of frames in error (one altered bit suffices to make the entire frame erroneous) to the total number of frames received. We also evaluate the BER/FER at the intended user to examine the impact of using the PLS schemes on the intended user's performance. Finally, we define the total transmit power by the BS antennas as $P_{\text{tot}} = \|\mathbf{x}_{\text{d}}\|^2$. In the simulations, we take the average of the above quantity over a large number of symbol slots to obtain the frame-level total transmit power, which is then averaged over a large number of channel realizations.

In the following simulations, we use QPSK modulation with Gray mapping. For the PLS random scheme, we set $\delta = \sigma_z^2/10$ to make sure that the noise will push Eve's received signal

Table 6.1: Channel coding parameters used for the simulations

Parameters	**Values**
Code rates	⅓, ¼
Constraint length	7
Frame Size	150
Number of frames	100
Trace-back length	96
Decoder decision technique	Hard, Soft

to either of the neighboring detection regions, i.e., to cause higher error rates at Eve. For simplicity, we consider unitary noise variance σ_z^2. As for the channel coding part, we use convolutional coding [Yua10] and Viterbi decoding [Vit06] with the parameters in Table 6.1. We note that low coding rates are chosen in order to consider a worst case eavesdropping scenario, where Eve can take advantage of the redundancy to correct as much errors as possible.

6.4.2 Selection of ML Algorithms for the Eve Attack

For the MLC modules used for the ML-based soft decoding scheme, in this simulation, we adopt problem transformation methods that remodel our MLC problem into single-label problem(s). Since our labels are bits, the MLC problem will be decomposed into k binary classifiers, where $k = \log_2 M_o$ is the number of bits constructing each symbol.

Herein, we use two transformation methods, BR [ZLLG18] and CC [YWF+15]. BR is the most simple and efficient method to solve MLC problems, which trains the k binary classifiers independently; its only drawback is that it does not consider labels correlation. CC, however, takes into account the correlation between labels by using the outputs of the previously trained classifiers as features for the subsequent ones in the chain, except for the first classifier. We refer to these soft-decoding implementation by "Soft - BR" and "Soft - CC" accordingly.

Concerning the MCC module used for the ML-based hard decoding scheme, it does not require any transformation or specific approach. It can be solved using any classifier. We refer to this scheme subsequently by "Hard". It is worth mentioning that the ML-based decoding schemes apply only to Eve, whereas the intended users employ conventional (not ML-based) soft and hard decoding techniques.

Table 6.2: Prediction accuracy of our proposed ML-based decoding schemes with several classifiers when using ZF, CIPSM, PLS random, and PLS Eve-min-power precoding schemes.

Classifiers	ZF		CISPM		PLS random		PLS Eve-min-power	
	Soft - CC	Hard	Soft - CC	Hard	Soft - CC	Hard	Soft - CC	Hard
Gaussian_NB	0.8338	0.8378	0.8271	0.8298	0.5431	0.5429	0.5708	0.5712
Log_Reg blue!25	0.9375	0.9374	0.8935	0.8929	0.5427	0.5433	0.5732	0.5732
SVM	0.9370	0.9361	0.8925	0.8931	0.5429	0.5426	0.5697	0.5718
R_Forest	0.8831	0.8798	0.8538	0.8491	0.5382	0.5354	0.5669	0.5650
KNN	0.9047	0.8994	0.8623	0.8570	0.5265	0.5245	0.5474	0.5454
Decision_Tree	0.8101	0.7951	0.7738	0.7571	0.5178	0.5207	0.5372	0.5374
Extra_Trees	0.9017	0.8970	0.8669	0.8644	0.5387	0.5377	0.5670	0.5679
LightGBM	0.8998	0.8958	0.8655	0.8611	0.5359	0.5356	0.5645	0.5629
XGB	0.8983	0.8946	0.8633	0.8609	0.5326	0.5349	0.5598	0.5605

To make Eve as sophisticated as possible, we experiment with several state-of-the-art classifiers and choose the one with the best performance. In Table 6.2, we compare the prediction accuracy of the proposed soft[4] and hard decoding schemes, considering ZF and CISPM precoding as well as the proposed PLS precoding schemes. The parameters used for this simulation are: $N_{\mathrm{t}} = 15$, $K = 6$, $M = 9$, and $\eta = \gamma_k = 6$ dB. We note that these results represent the averaged results over 100 different channel realizations. We also note that this accuracy applies before channel decoding, i.e., by comparing the ML predicted labels to the actual coded transmitted symbols to user k.

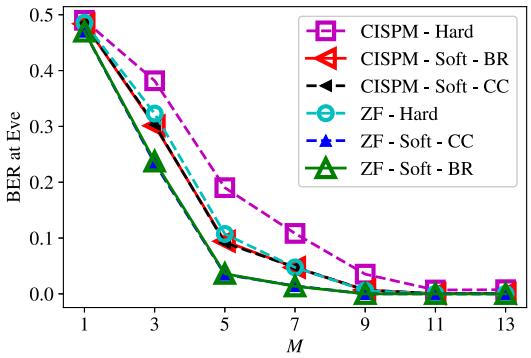
As observed in Table 6.2, the logistic regression classifier achieves the highest prediction accuracy among all the precoding and decoding schemes. Therefore, in the following simulations, we use this classifier in our proposed ML-based decoding schemes.

6.4.3 Comparison and Insights

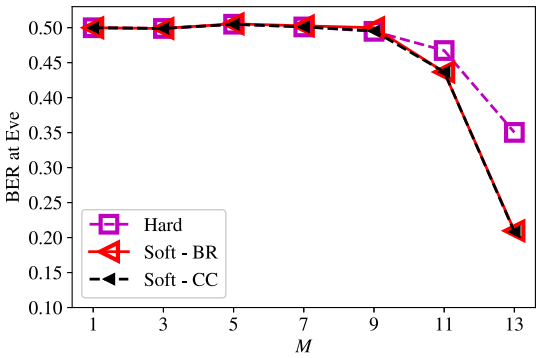
We note that a BER value of 0.5 indicates full confusion. Regarding the target FER values at the intended users, it varies depending on the application scenario. For instance, eMBB in 5G requires an FER on the order of 10^{-3} while massive machine-type communications (mMTC) require only 10^{-1} [PTSD18]. In this section, we will validate the eavesdropping attacks by showing the FER at Eve to be in the order of the intended users' FER values.

Figure 6.7 depicts the coded BER at Eve as a function of its number of antennas M. We compare the proposed hard and soft decoding schemes. The parameters used in the simulation are: $r = 1/3$, $N_{\mathrm{t}} = 15$, $K = 6$, and $\eta = \gamma_k = 6$ dB. Figure 6.7(a) represents the non-secure precoding schemes, ZF and CISPM. For our hard and soft decoding schemes, we observe that the more antennas at Eve, the lower is the BER, i.e., the more antennas at Eve, the higher the prediction accuracy (more versions of the same signal, hence more

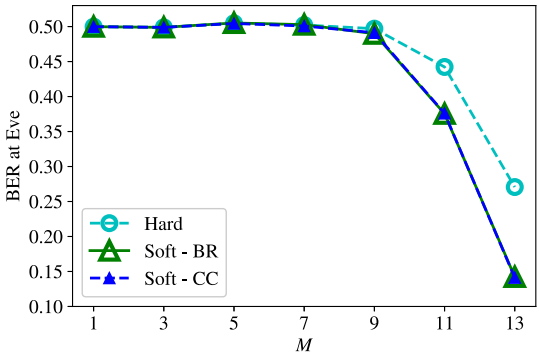
[4]In the table, we did not show results for Soft - BR to avoid redundancy, as its results were almost the same as Soft - CC.



(a) ZF and CISPM schemes



(b) PLS random scheme



(c) PLS Eve-min-power scheme

Figure 6.7: BER at Eve vs. number of antennas at Eve, with $r = 1/3$, $N_t = 15$, $K = 6$, and $\eta = \gamma_k = 6$ dB.

features used for training and inference), leading to lower BER. In addition, our Soft technique outperforms the Hard one, where Soft - BR and Soft - CC are equivalent. Moreover, with 9 antennas at Eve, the BER at Eve is so low to the point that it could be compared to an intended user's decoding performance, leading to a big vulnerability in systems that use ZF and CISPM precoding. In addition, as expected, CISPM is more secure than ZF as predicted in Section 6.2. As for the case of the PLS random scheme in Figure 6.7(b), we observe the same pattern as in Figure 6.7(a), the more antennas at Eve, the lower is the BER, with Soft decoding outperforming the Hard one. However, when M values are lower or equal to 9, the BER at Eve is at 0.5, indicating total equivocation. In fact, even for higher values than 9, the BER at Eve is still very high compared to ZF and CISPM schemes, i.e., PLS random scheme is offering a significant security gain when compared to the latter ones. Similarly, for the PLS Eve-min-power scheme in Figure 6.7(c), we observe the same behavior as for PLS

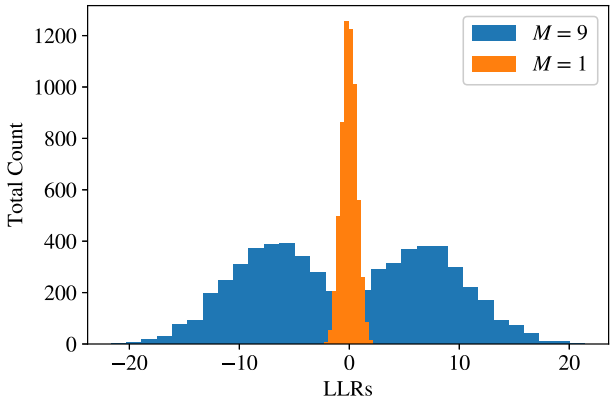
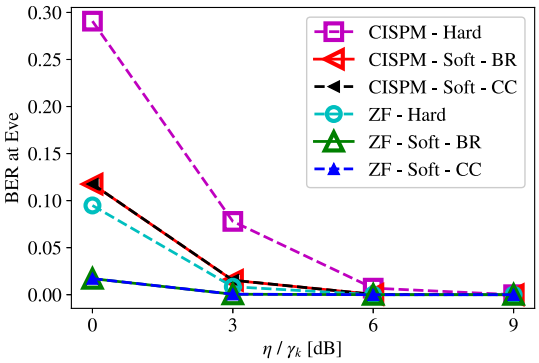


Figure 6.8: LLRs distribution of the Soft - CC decoding scheme with $r = 1/3$ and $M \in \{1, 9\}$.

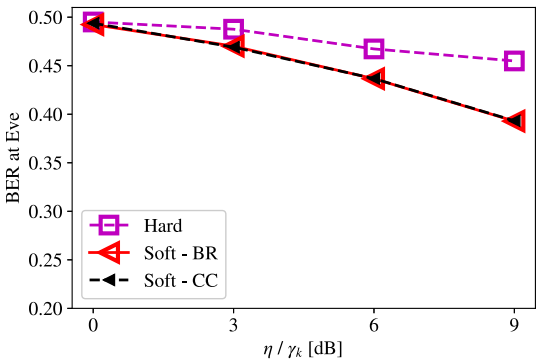
random scheme, with the PLS Eve-min-power scheme BER going lower than PLS random, i.e., PLS Eve-min-power is less secure than PLS random. However, when compared to non-secure precoding schemes, ZF and CISPM, PLS Eve-min-power still offers a drastic security gain. To summarise, as shown in Figure 6.7, the BER at Eve when using the proposed PLS schemes is close to 1/2 in the case when $M \leq 9$, which represents a worst-case reception scenario where the interference and noise effects fully dominate the useful signal. However, when Eve uses $M = 13$ antennas, the BER is close to 0.2 when using PLS random scheme. Even in this case, if Eve does not have a reference to compare to, i.e., Eve does not have any prior knowledge about the received data, then Eve does not know where this 20% error occurred, thus it is extremely difficult to make sense of this data.

Figure 6.8(a) depicts the distribution of the estimated LLRs for Soft - CC decoding scheme. We note that for this particular plot, we used 1000 symbols to obtain a smooth histogram. We recall that a probability of 0.5 indicates that the predictor is not sure whether the predicted bit should be 0 or 1; a value close to 1 means the predictor is very sure that it is a 1, whereas a probability close to 0 indicates the opposite, i.e., it is very sure that it is not a 1. As expected, the LLRs values for the case of 9 antennas at Eve are distributed mostly away from 0, indicating high quality LLRs. Namely, the corresponding probabilities are mostly different than 0.5, thus yielding a high prediction accuracy. Using a higher number of antennas entails more features that can be employed in both training and inference phases, therefore leading to higher accuracy when estimating the likelihoods. However, when Eve uses only 1 receive antenna, the LLRs are close to 0 as their probabilities are close to 0.5,

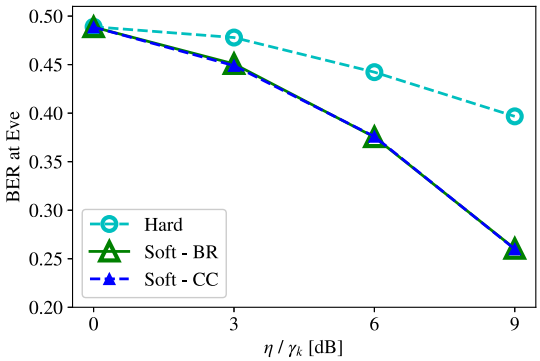
implying poor quality LLRs. Therefore, we conclude that higher number of antennas at Eve leads to higher prediction accuracy, and therefore lower BER.



(a) ZF and CISPM schemes



(b) PLS random scheme

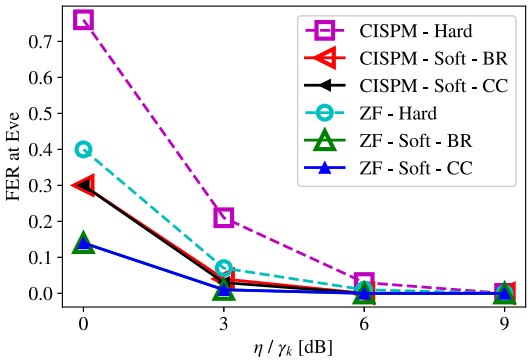


(c) PLS Eve-min-power scheme

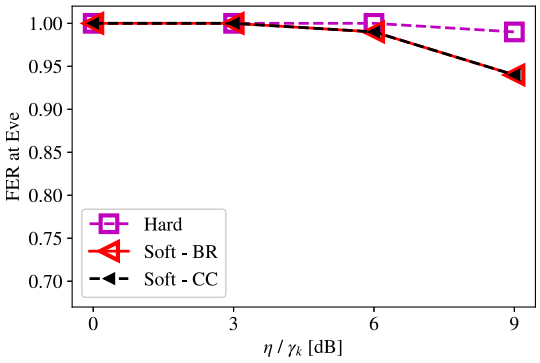
Figure 6.9: BER at Eve vs. η/γ_k [dB], with $r = 1/3$, $N_t = 15$, $K = 6$, and $M = 11$.

Figure 6.9 depicts the coded BER at Eve as a function of η/γ_k [dB], which we set to the same value for all users for simplicity. The parameters used in the simulation setup are: $r = 1/3$, $N_t = 15$, $K = 6$, and $M = 11$. Concerning ZF and CISPM schemes in Figure 6.9(a), with the proposed soft and hard decoding approaches we notice that the higher the values of η/γ_k, the lower the BER. Particularly, higher η/γ_k values leads to higher transmit power, which cause higher received power at Eve, thus better decoding performance. Moreover, Soft decoding is outperforming the Hard one. Additionally, CISPM precoding is more secure than ZF, i.e. BER at Eve for CISPM is higher than the one of ZF. As for the use of the PLS random scheme, in Figure 6.9(b), we notice that, similarly, the higher the values of η/γ_k, the lower the BER. We also observe that soft decoding is the most performing with the difference being that using PLS random scheme offers much higher security compared to

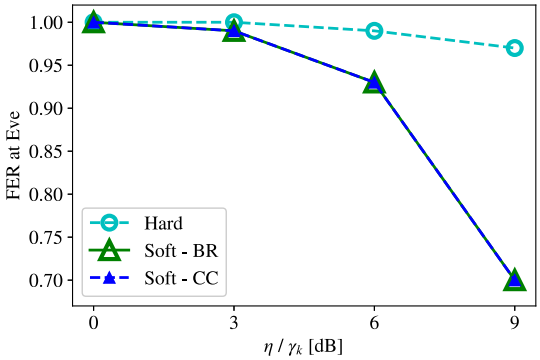
ZF and CISPM scheme, with high BER values at Eve even when using 11 antennas at Eve. Concerning the PLS Eve-min-power scheme in Figure 6.9(c), it depicts the same behavior as PLS random. However, the latter scheme is more secure because of its incurred randomness in the precoding design, while PLS Eve-min-power scheme is designed with Eve's channel in the objective function that lowers the received power at Eve.



(a) ZF and CISPM scheme



(b) PLS random scheme



(c) PLS Eve-min-power scheme

Figure 6.10: FER at Eve vs. η/γ_k [dB], with $r = 1/4$, $N_t = 15$, $K = 6$, and $M = 11$.

Figure 6.10 depicts the coded FER at Eve as a function of η/γ_k [dB]. The parameters used in the simulation are: $r = 1/4$, $N_t = 15$, $K = 6$, and $M = 11$. In Figure 6.10(a), for ZF and CISPM precoding, we notice that the higher the values of η/γ_k, the lower the FER, which is due to the increase of the transmit power. Particularly, soft decoding outperforms hard decoding, with FER values decently low, which validates the eavesdropping attack for ZF and CISPM precoding schemes, with CISPM being more secure. In Figure 6.10(b) however, when we use the PLS random scheme, we notice that the higher the values of η/γ_k, the lower the FER. Particularly, the high FER values validate the security of the PLS random scheme.

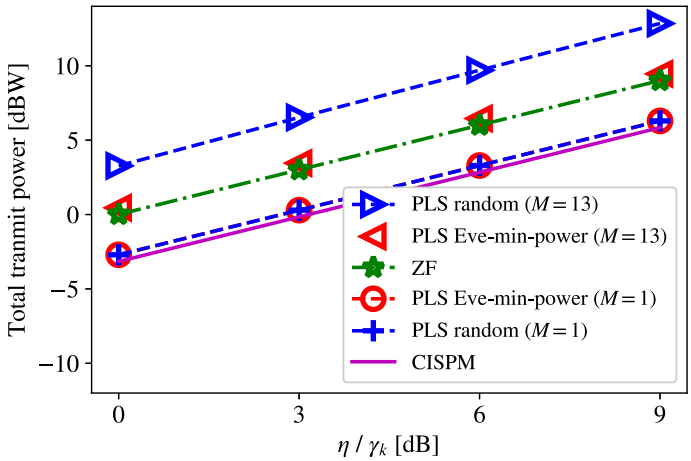


Figure 6.11: Total transmit power [dBW] vs. target SINR [dB], with $N_{\mathrm{t}} = 15$, $K = 6$, and $M \in \{1, 13\}$.

Lastly, when using the PLS Eve-min-power scheme, in Figure 6.10(c), we observe the same behavior as in the case of the PLS random scheme, with FER values at Eve lower than ones for the PLS random approach. This validates the high security exhibited by the PLS random scheme, which outperforms the PLS Eve-min-power scheme's security performance. Yet, the FER values for the PLS Eve-min-power are still very high compared to the non-secure schemes, i.e., ZF and CISPM schemes.

Figure 6.11 shows the total transmit power P_{tot}, in dBW, as a function of η/γ_k, which we set to the same value for all users for simplicity. We compare ZF and CISPM schemes with the PLS ones. The parameters used in the simulation are: $N_{\mathrm{t}} = 15$, $K = 6$, and $M = \{1, 13\}$. As explained above, the higher the η/γ_k values, the higher is the transmit power, for all the schemes. For the CISPM scheme, we observe that P_{tot} is lower than the target SINR at the intended users, which is due to the constructive interference turning into power gains at the receivers, hence less transmit power is required to attain the desired SINR value. However for ZF precoding, as predicted, P_{tot} is the same as the mean power η, as it has been set. As for the PLS schemes, P_{tot} depends on the number of antennas at Eve M, higher M leads to higher P_{tot}. This increase is due to the fact that more antennas at Eve imply more constraints in the case of PLS random that result in the observed big increase in P_{tot} for $M = 13$. To elaborate more, this behavior is due to the fact that, the more we constrain our signal design problem, the more power is required to solve it. However, in the case of PLS Eve-min-power, the higher M, the higher the power consumption, i.e., the more antennas at

Eve, the Eve-related part of the objective function tends to have higher values due to the higher degrees of freedom at the Eve's side, thus higher power consumption; even for $M = 13$, PLS Eve-min-power does not consume as much power as PLS random, its consumption is in fact equivalent to ZF scheme in P_{tot}. However, for $M = 1$, the two proposed PLS schemes consume the same power.

Lastly, we investigate the impact of pilot overhead on our proposed ML-based attacks and countermeasures. We refer to [LJ10] for some practical pilot overhead values. Figure 1 in [LJ10] shows the ergodic spectral efficiency as a function of pilot overhead in a high velocity setting, i.e., where Doppler spectrum is Clarke-Jakes[5] with a maximum normalized frequency $f_D = 0,02$, e.g, the user/Eve is at 100 Km/h speed in a WiMAX system. We observe that maximum spectral efficiency is achieved when using 0.1 pilot overhead. Thus, we evaluate our proposed Soft - CC decoding scheme using such value and lower.

Figure 6.12 plots the coded BER at Eve using Soft - CC as a function of the pilot overhead. The parameters used in the simulation are: $r = 1/3$, $N_{\text{t}} = 15$, $K = 6$, $M = 11$, $\eta = \gamma_k = 6$ dB, and a frame size of 900 symbols. When the BS uses the non-secure schemes, ZF and CISPM, we observe that the higher the pilot overhead, the lower the BER, in particular, with a pilot overhead value of 0.1 the BER at Eve is as low as 10^{-3}, which is a sufficiently small BER that threatens the communication security. Thus, this validates again our ML-based attack even with pilot overhead values as low as 0.1. However, when the BS uses PLS schemes, the BER at Eve remains high, which again validates our countermeasures.

Lastly, in Fig. 6.13 we plot the runtime per SP [ms] as a function of the number of antennas at Eve M of the proposed and benchmark SLP schemes. The parameters used in the simulation are: $N_{\text{t}} = 15$, $K = 6$, and $\gamma_k = 6$ dB, and a frame size of 900 symbols. We observe that for both PLS Eve-min-power and CISPM schemes, the runtime does not depend on M because both schemes do not have Eve-related constraints. And as expected, the PLS Eve-min-power scheme's runtime is a bit higher than the CISPM's one because of the extra Eve-related term in the PLS Eve-min-power scheme's objective function in eq. (6.13). However, the runtime of PLS random increases with M and is higher than the runtime of PLS Eve-min-power. This increase is due to the Eve-related constraints in eq. (6.12) of the PLS random scheme, where in addition to the CI constraints in eqs. (6.10) and (6.11), there

[5]We note that Jakes-Clark model requires a uniform angular distribution of scattering objects around the users in addition to rich scattering.

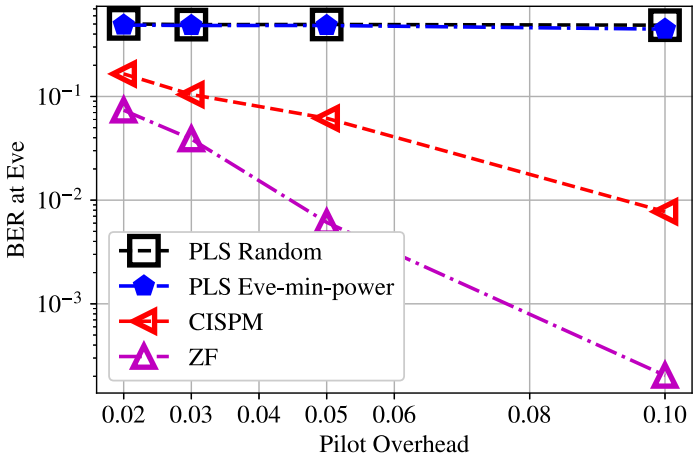


Figure 6.12: BER at Eve using Soft - CC vs. the pilot percentage, with $r = 1/3$, $N_t = 15$, $K = 6$, $M = 11$, $\eta = \gamma_k = 6$ dB, and a frame size of 900 symbols.

will be M Eve-related constraints. Thus, the higher M, the more constraints in the PLS random optimization problem, and therefore the higher the runtime.

We conclude this section by summarizing the insights from the numerical results.

- Soft decoding schemes always outperform hard decoding, i.e., soft values carry extra information that is used by the decoder to better estimate the original data.
- Soft – CC and Soft – BR performance is the same because of the lack of label-correlation, due to the random nature of data to be transmitted.
- Our proposed soft decoding schemes operates well with pilot overhead values as low as 0.1.
- CISPM precoding is more secure than ZF because the precoding pattern changes at each symbol-period while ZF precoding is fixed throughout the whole coherence time.
- Proposed PLS schemes are much more secure than CISPM and ZF, with PLS random being the most secure because of its induced randomness in the signal design that makes it harder for Eve to learn the precoding pattern.
- Tremendous security gains are offered by the PLS Eve-min-power scheme when compared to the benchmark SLP scheme, CISPM, at the expense of only a marginal extra runtime.

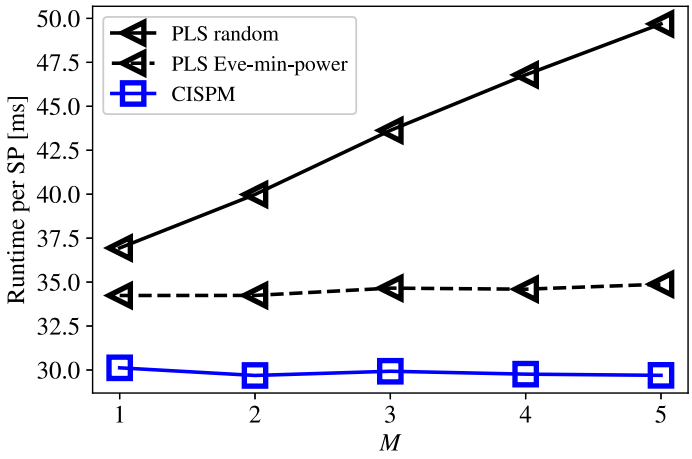


Figure 6.13: Runtime per SP [ms] vs. M of proposed and benchmark schemes with $N_{\mathrm{t}} = 15$, $K = 6$, $\gamma_k = 6$ dB, and a frame size of 900 symbols.

- PLS random scheme offers higher security than PLS Eve-min-power, however, its runtime increases linearly with the number of antennas at Eve.
- Logistic regression is the most performing classifier amongst the tested state-of-the-art classifiers.
- The system parameters that directly affect the BER/FER at Eve are: the number of antennas at Eve, the total transmit power, and the coding rate.
- PLS schemes offer significant security gains compared to ZF and CISPM precoding schemes at the expense of additional power consumption at the transmitter.
- More importantly, these security gains are achieved without affecting the performance at the intended users.

6.5 Summary

In this chapter, we investigated one-shot ML-based MIMO detection for multi-antenna Eve in coded systems. We proposed ML-based decoding schemes for a multi-antenna Eve in the context of a FEC-enabled MU-MISO systems. The proposed eavesdropping attacks use precoded pilot symbols as training data and enable an Eve to soft/hard decode a message with high accuracy. As a countermeasure to these attacks, we proposed two novel security-enhanced SLP precoders that seek to obstruct the learning process at Eve. Numerical results

validated both the attacks as well as the countermeasures, where the soft decoding scheme always outperforms the hard decoding one. In addition, our proposed PLS schemes outperform ZF and CISPM precoding in security at the expense of additional power consumption at the transmitter, with PLS random scheme offering the highest security. Thus, the proposed PLS schemes provide different trade-offs between security, runtime, and power consumption, which would give the BS the option to select the most suited scheme depending on the required criteria. Notably, despite all the security gains offered by our proposed PLS schemes, their use does not affect the performance at the intended user. Future research topics would be to extend this work to the case of imperfect CSI and also where the channel to Eve is unknown to the BS.

Chapter 7

Conclusions and Future Works

This chapter lists the main conclusions, discusses the underlying limitations of the proposed schemes and methodologies, and proposes some future works.

7.1 Conclusions

This thesis addressed detection challenges in precoded MIMO systems, considering a variety of scenarios and systems. In the context of SU-MIMO systems with channel coding, one-shot ML-based MIMO detection schemes robust to CSIT deterioration were proposed. A similar detection approach was developed for multi-antenna eavesdropping in coded and uncoded MU-MISO systems. Specifically, in this thesis, two detection-related problems were investigated: 1) MIMO detection robust to CSIT deterioration in SU-MIMO systems, and 2) MIMO detection for eavesdropping attacks in MU-MISO systems. The proposed detection schemes leverage ML tools and exploit the downlink pilots as training data. For the first problem, under severe CSIT deterioration, the proposed ML-based MIMO detectors can achieve good detection performance and outperform the two-step MMSE detector, i.e., that first estimates the CSIR using the downlink pilots and then performs detection. For the second problem, numerical results validated the effectiveness of the proposed eavesdropping attacks, where an Eve can decode the transmitted data with good accuracy. To counteract these attacks, several SLP-based schemes were proposed to enhance PLS. The proposed schemes were designed to meet varying runtime, security, and power consumption trade-offs, to provide the BS with options to choose the most suitable scheme depending on the desired criteria. Hence, the title of this thesis, ML for *MIMO detection* and *eavesdropping with SLP countermeasures*.

7.2 Limitations and Future Works

Finally, we provide some perspective on the obtained results by outlining some shortcomings of the proposed MIMO detectors and suggest possible improvements.

7.2.1 Outlook for Online Learning

Although the proposed MIMO detection approaches function in the online regime, where the detectors are optimized for each channel realization, this approach requires training the ML-model for each coherence time. In particular, the detection performance of the online approach is limited in the case when pilot data is scarce due to short coherence time. Thus, one-shot MIMO detection in offline learning should be investigated in order to 1) support scenarios where the coherence length is short and 2) when the computation power at the receiver is limited and does not permit to re-train the ML-model whenever the channel changes.

7.2.2 Limitation of using Eve's CSI at the BS

In the proposed countermeasures to the eavesdropping attacks, the underlying SLP-based schemes are designed by assuming that the BS knows the channel to Eve, which is not the case in general. Therefore, it is important to investigate the case when the BS does not know Eve's CSI. Particularly, to counteract the proposed eavesdropping attacks, precoding schemes to enhance the PLS without requiring Eve's CSI should be investigated to consider a more general Eve.

7.2.3 Outlook for Deep Learning

The proposed ML-based detectors employ non-neural networks based algorithms, we did not consider deep learning (DL) despite its high performance in several areas. The main reason for not using DL-based approaches in our design are 1) DL requires considerable amount of training data, which is not available in our case due to the limited pilots available in each coherence time and 2) DL typically requires to train millions of parameters, which makes it prohibitively expensive in our application scenario. Nevertheless, when offline learning is considered, i.e., the detector is trained for a multitude of channel realizations at once, the underlying pilot data might be large enough to train DL models. Thus, it might be

interesting to consider DL algorithms for one-shot MIMO detection, especially in the offline training regime.

7.2.4 Using Interleaving with Channel Coding

In this thesis, we employed channel coding without interleaving, where we used convolutional coding and Viterbi decoding. Interleaving improves the detection performance further, thus employing interleaving with channel coding will further enhance the eavesdropping attack's and the MIMO detection schemes' performance.

7.2.5 Using ML Algorithms that Directly Process Complex Numbers

In the proposed ML frameworks for eavesdropping and MIMO detection, for the features employed for training and inference, we considered real and imaginary parts separately because the employed ML algorithms' implementations did not support complex-valued data. Nevertheless, there is a growing interest in building NN-based algorithms using complex numbers and exploring the benefits of directly processing complex data [BQL21]. Thus, it is interesting to investigate complex-based ML and DL algorithms, which might bring better performance due to the join processing.

7.2.6 Using More Benchmarks in Precoding and Detection

In this thesis, with regards to the proposed countermeasures, we have compared the proposed SLP precoders' performance to precoding schemes that are not secure by design. Therefore, the next step of this work is to compare the proposed schemes to state-of-the-art secure schemes in terms of security gains, power consumption, and complexity. In the same direction, the benchmarks used for comparison with the proposed ML detectors are not robust by design, thus it is interesting to compare the proposed ML detectors to robust detectors and assess the resulting complexity-performance trade-offs.

—

Chapter 8

Appendix: Regression-Based ML Framework for Soft Decoding

In this chapter, we re-design the ML framework for soft decoding in Section 6.2.3 using regression instead of classification.

We revise that supervised learning is applicable when the training set is labeled, i.e., the input data (X) is associated to an output label (Y), where the ML algorithm learns the mapping function f from the input to the output, $Y = f(X)$. The aim herein is to approximate the mapping function, without underfitting nor overfitting, to predict well for new input data (X). In our setting, the input variable (X) represents the Eve's received pilot signal $\mathbf{y}_{\mathrm{e}}^{\mathrm{p}}$, detailed in (6.5). In Section 6.2.3, we chose the variable (Y) to represent the pilot vector $\mathbf{p}$, which constitute discrete values, hence the use of classification. Motivated by the numerical nature of LLRs, we convert variable (Y) to LLRs to train the model on the estimated LLRs instead. This transformation results in the considered supervised learning problem being regarded as a *regression* problem, as the output variable is of real/continuous nature.

As illustrated in Figure 8.1, the regression-based ML framework for soft-decoding encompasses two steps: 1) training phase, where the ML model is trained by using the estimated LLRs, and 2) inference phase, where the LLRs are directly predicted and used by the soft decoder to obtain the transmitted bits.

In the training phase, the difference with the classification-based ML framework presented in Section 6.2.3 is in training the ML model using LLRs instead of bits. As depicted in

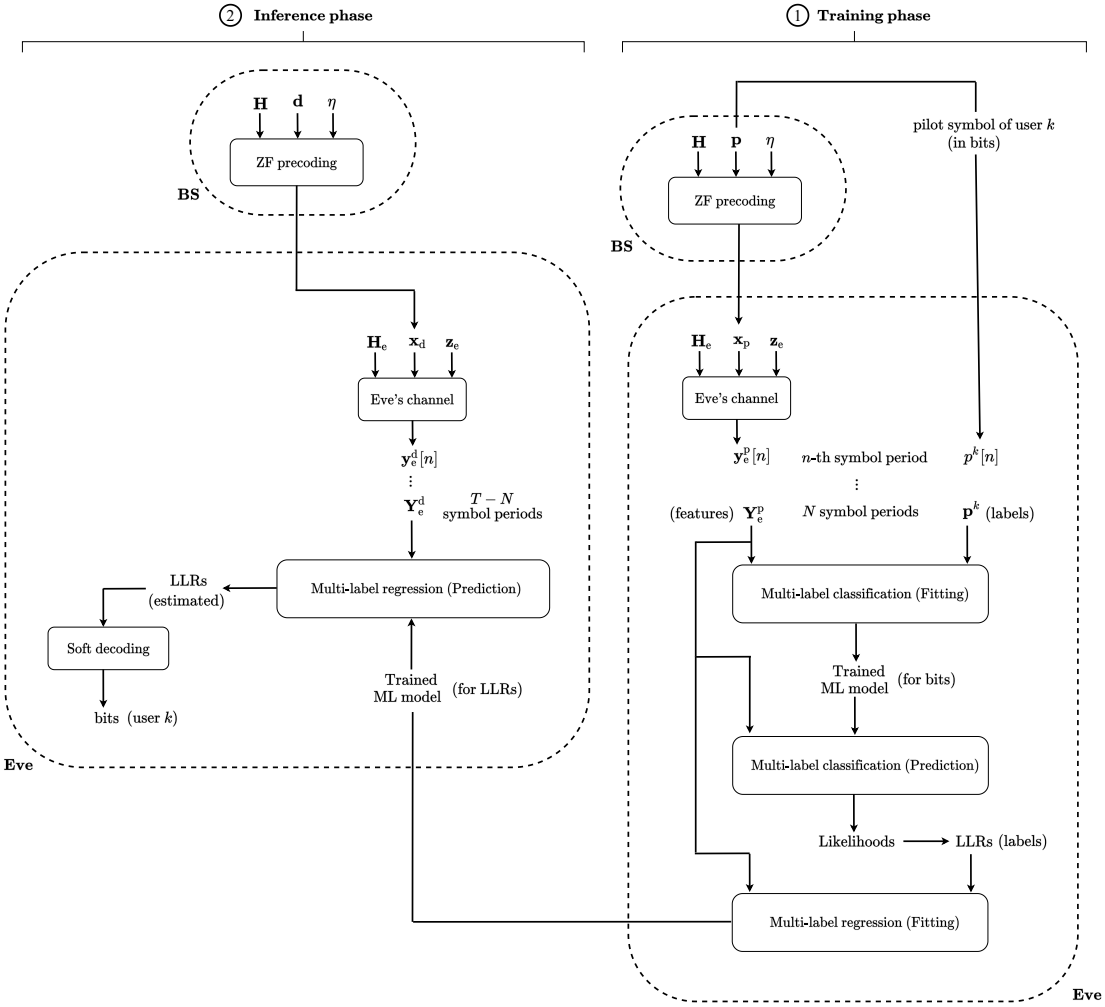


Figure 8.1: Overview of the regression-based ML framework for soft decoding.

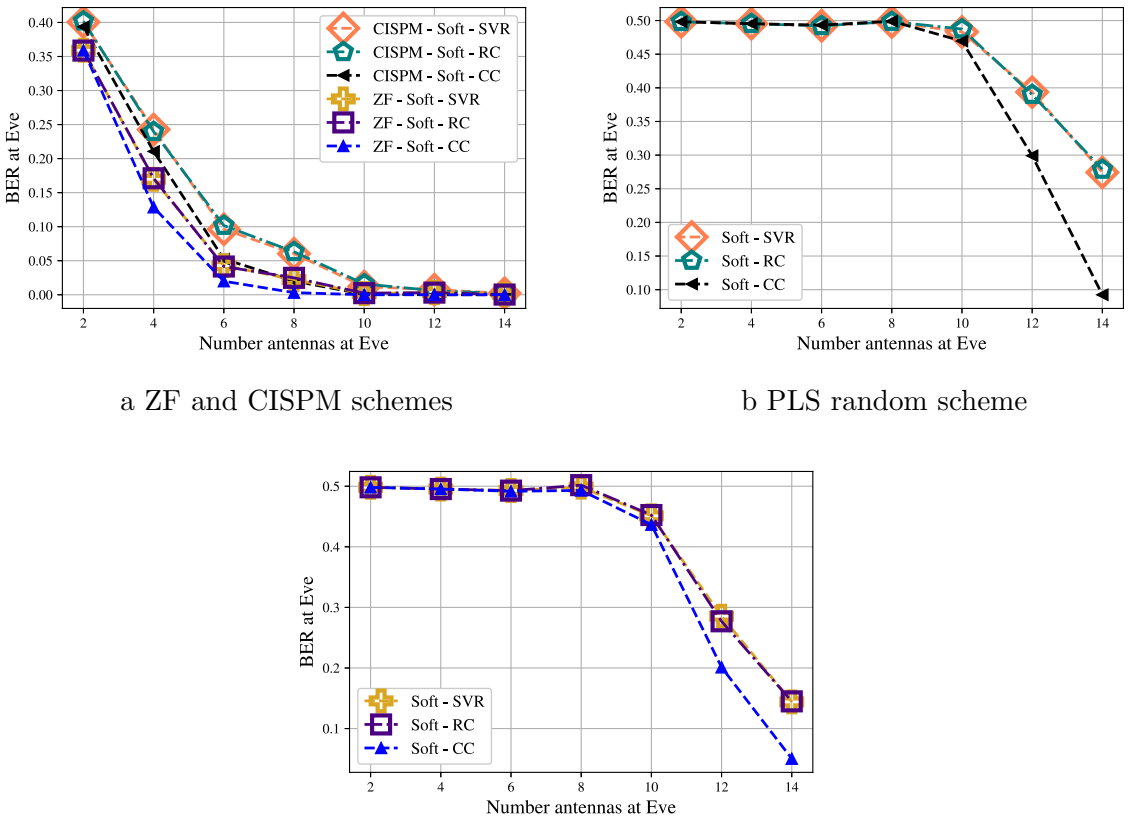
Figure 8.1, the LLRs corresponding to the training set's features are estimated by employing the classification-based approach. In particular, via employing the bits-based ML model in Section 6.2.3 to estimate the LLRs. The outcome of this phase is a trained multi-label regression model for LLRs.

In the inference phase, as depicted in Figure 8.1, the trained ML model for LLRs is employed by the multi-label regression prediction module to directly infer the LLRs, which are fed to a soft decoder to obtain the transmitted bits to user k.

Similarly to Section 4.3, herein, we employ a lightweight implementations of the multi-label regression algorithms, the support vector regression (SVR) [AK15] and the regressor chain (RC) [RM19] approaches. In fact, SVR is similar to the BR approach in the sense that both methods train multiple single-label modules independently and combining the produced

outputs into a multi-label output. On the other hand, RC is equivalent to the CC method where both take into account the correlation between the labels, by using the outputs of the previously trained modules as features for the next ones in the chain.

Using the same channel coding parameters in Section 6.4, in the following, we present the BER/FER results at Eve that includes the SVR and RC approaches and compare their performance to the CC method.



a ZF and CISPM schemes

b PLS random scheme

c PLS Eve-min-power scheme

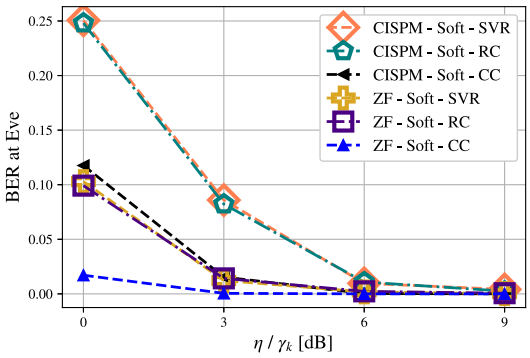
Figure 8.2: BER at Eve vs. number of antennas at Eve, with $r = 1/3$, $N_t = 15$, $K = 6$, and $\eta = \gamma_k = 6$ dB.

Figure 8.2 depicts the coded BER at Eve as a function of M. The parameters used in the simulation are: $r = 1/3$, $N_t = 15$, $K = 6$, and $\eta = \gamma_k = 6$ dB. Figure 8.2a represents the non-secure precoding schemes, ZF and CISPM. Similar to Figure 6.7, we observe that the more antennas at Eve, the lower is the BER. In particular, SVR and CC approaches depict the same BER performance as the labels are random in nature, i.e., no label correlation between them. Nevertheless, we observe that the CC technique outperforms both the SVR and RC

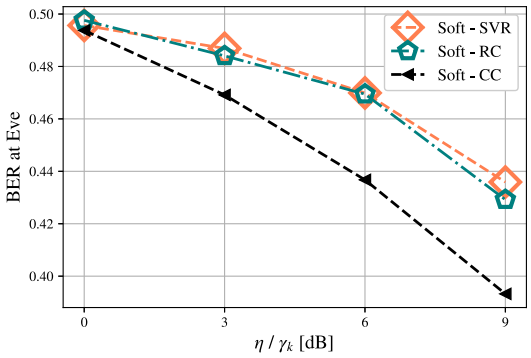
approaches. This is due to the regression-based modeling that trains the model on LLRs that are estimated, which incurs an additional loss in performance. However, the classification-based approach in Chapters 4 and 6 performs training on the actual labels, thus the better prediction accuracy. Moreover, with $M = 10$, the BER at Eve is so low that it could be compared to an intended user's decoding performance, leading to an important eavesdropping vulnerability in systems that employ ZF and CISPM precoding schemes. Additionally, as shown in Section 6.4, the ZF scheme leads to lower BER than the CISPM precoding, which makes the CISPM scheme more secure. For the PLS random scheme in Figure 8.2b, we observe the same behavior as in Figure 8.2a, higher M leads to lower BER, where SVR and RC are equivalent while CC outperforming both. However, when M is lower than 10, the BER at Eve is at 0.5, indicating full confusion at Eve. In fact, even when $M > 10$, the BER at Eve is still very high when compared to ZF and CISPM schemes. Similarly, for the PLS Eve-min-power scheme in Figure 8.2c, we observe that the PLS Eve-min-power scheme's BER values are lower than the PLS random values, making the latter scheme more secure. However, PLS Eve-min-power scheme still offers considerable security gains when compared to the non-secure precoding schemes, ZF and CISPM.

Figure 8.3 depicts the coded BER at Eve as a function of η/γ_k [dB], which we set to the same value for all users. The parameters used in the simulation setup are: $r = 1/3$, $N_\mathrm{t} = 15$, $K = 6$, and $M = 11$. Regarding ZF and CISPM schemes in Figure 8.3a, similar to Figure 6.9, we remark that the higher the values of η/γ_k, the lower the BER, as higher η/γ_k values generate higher transmit power, which in turn results in higher received power at Eve, thus better decoding performance. Similarly, the same observations remain: CISPM precoding is more secure than ZF, SVR and RC lead to the same performance, and the CC approach outperforms SVR and RC. For the PLS random scheme in Figure 8.3b, we witness the same behavior with regard to η/γ_k. In particular, the CC approach clearly outperforms the SVR and RC techniques. Still, even with $\eta/\gamma_k = 9$ dB and with such a high number of antennas at Eve, $M = 11$, the BER values are still very high when using the PLS Random scheme. In Figure 8.3c, the PLS Eve-min-power scheme depicts the same behavior as PLS random. Even if the former scheme BER values are lower than the latter, the BER performance at the former is still very poor, approximating 0.33 for the SVR and RC approaches even when using $M = 11$ and $\eta/\gamma_k = 9$ dB.

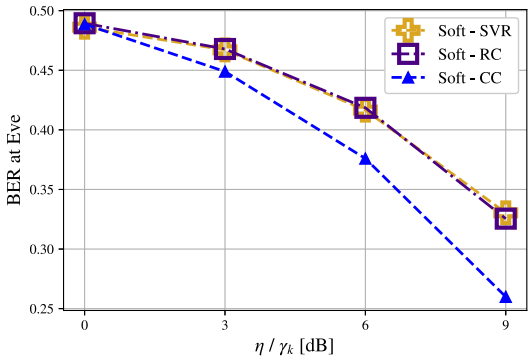
Figure 8.4 depicts the coded FER at Eve as a function of η/γ_k [dB]. The parameters used



a ZF and CISPM schemes



b PLS random scheme

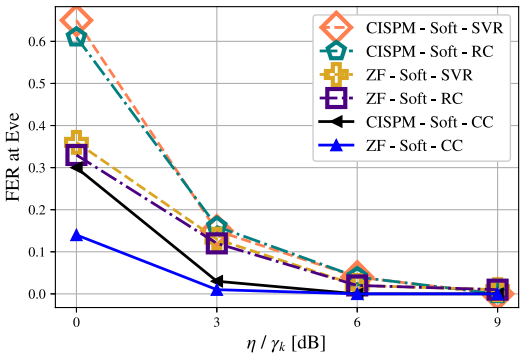


c PLS Eve-min-power scheme

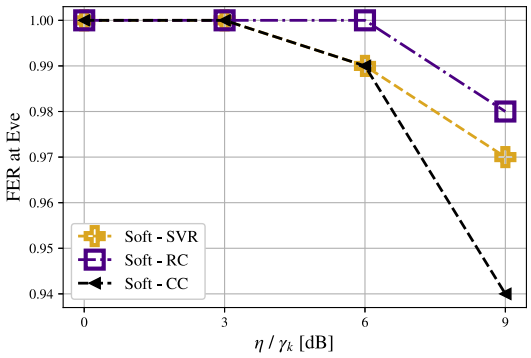
Figure 8.3: BER at Eve vs. η/γ_k [dB], with $r = 1/3$, $N_t = 15$, $K = 6$, and $M = 11$.

in the simulation are: $r = 1/4$, $N_t = 15$, $K = 6$, and $M = 11$. Figure 8.4a plots the FER at Eve when using ZF and CISPM precoding schemes. Similar to Figure 8.3, we observe that: 1) the higher the η/γ_k, the lower the FER, 2) the CC achieves lower FER than the SVR and RC approaches, and 3) the SVR and RC yield the same decoding performance. The same behavior is noted in Figures 8.4b and 8.4c, for the PLS random and Eve-min-power schemes, respectively, with the exception that the FER values are still very high even when $\eta/\gamma_k = 9$ dB. Specifically, for the former scheme, FER values are above 0.94 but 0.7 for the latter. Again, this validates the security edge of the PLS random scheme over the PLS Eve-min-power scheme. Though, the PLS Eve-min-power successfully obstructs the eavesdropping attacks, i.e., Eve could correctly decode no more than 30% of the transmitted data.

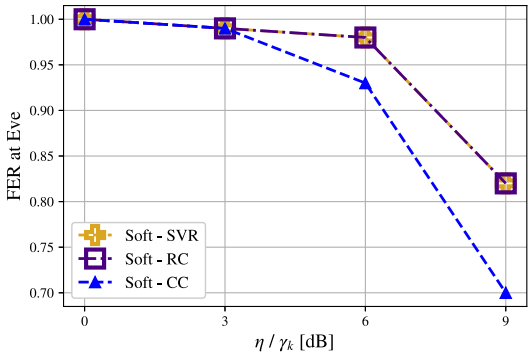
We conclude this appendix by summarizing the insights from the numerical results of using the SVR and RC multi-label regression algorithms.



a ZF and CISPM scheme



b PLS random scheme



c PLS Eve-min-power scheme

Figure 8.4: FER at Eve vs. η/γ_k [dB], with $r = 1/4$, $N_t = 15$, $K = 6$, and $M = 11$.

- The SVR and RC approaches apply only to the soft decoding ML framework, where the LLRs are estimated and used to train the underlying multi-label regression models.
- These approaches are equivalent in performance because of the random nature of the labels employed, i.e., there is no inter-label correlation to be exploited to further enhance the accuracy of the RC approach.
- These approaches fall behind the CC method in BER/FER performance, as they both rely on the estimated LLRs for training.
- The BER/FER values tend to 0 when employing the ZF and CISPM schemes at the BS, however, the values remain very high when the BS uses the proposed PLS schemes.

—

—

Bibliography

[ACH08] M. Abuthinien, S. Chen, and L. Hanzo. Semi-blind joint maximum likelihood channel estimation and data detection for MIMO systems. *IEEE Signal Process. Lett.*, 15:202–205, 2008.

[ACO14] M. Alodeh, S. Chatzinotas, and B. Ottersten. A multicast approach for constructive interference precoding in MISO downlink channel. *IEEE Int. Symposium Inf. Theory*, pages 2534–2538, June 2014.

[ACO15a] M. Alodeh, S. Chatzinotas, and B. Ottersten. Constructive interference through symbol level precoding for multi-level modulation. In *IEEE Global Commun. Conf. (GLOBECOM)*, pages 1–6, 2015.

[ACO15b] M. Alodeh, S. Chatzinotas, and B. Ottersten. Constructive Multiuser Interference in Symbol Level Precoding for the MISO Downlink Channel. *IEEE 16th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, page 36–40, 2015.

[ACO15c] M. Alodeh, S. Chatzinotas, and B. Ottersten. Constructive Multiuser Interference in Symbol Level Precoding for the MISO Downlink Channel. *IEEE Trans. Signal Process.*, 63(9):2239–2252, May 2015.

[ACO16] M. Alodeh, S. Chatzinotas, and B. Ottersten. Energy-efficient symbol-level precoding in multiuser MISO based on relaxed detection region. *IEEE Trans. Wireless Comm.*, 15(5):3755–3767, May 2016.

[ACO17] M. Alodeh, S. Chatzinotas, and B. Ottersten. Symbol-level multiuser MISO precoding for multi-level adaptive modulation. In *IEEE Trans. Wireless Commun.*, volume 16, pages 5511–5524, Aug 2017.

[ADM+07] E. Alberty, S. Defever, C. Moreau, R. De Gaudenzi, A. Ginesi, R. Rinaldo, G. Gallinaro, and A. Vernucci. Adaptive coding and modulation for the DVB-S2 standard interactive applications: Capacity assessment and key system issues. *IEEE Wireless Commun.*, 14(4):61–69, 2007.

[AEVZ02] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger. Closest point search in lattices. *IEEE Trans. Inf. Theory*, 48(8):2201–2214, 2002.

[AK15] Mariette Awad and Rahul Khanna. Support vector regression. In *Efficient Learning Machines*, chapter 4, pages 67–80. Apress Press, Berkeley, CA, USA, 2015.

[AM17] P. V. Amadori and C. Masouros. Constant envelope precoding by interference exploitation in phase shift keying-modulated multiuser transmission. *IEEE Trans. Wireless Commun.*, 16(1):538–550, Jan 2017.

[ASK+18] M. Alodeh, D. Spano, A. Kalantari, C. G. Tsinos, D. Christopoulos, S. Chatzinotas, and B. Ottersten. Symbol-level and multicast precoding for multiuser multiantenna downlink: A state-of-the-art, classification, and challenges. *IEEE Commun. Surveys Tuts.*, 20(3):1733–1757, thirdquarter 2018.

[BBO14] Emil Björnson, Mats Bengtsson, and Björn Ottersten. Optimal multiuser transmit beamforming: A difficult problem with a simple solution structure [lecture notes]. *IEEE Signal Process. Mag.*, 31(4):142–148, 2014.

[BM13] B. N. Bharath and C. R. Murthy. Channel training signal design for reciprocal multiple antenna systems with beamforming. *IEEE Trans. Vehicular Tech.*, 62(1):140–151, 2013.

[BO99] Mats Bengtsson and Björn Ottersten. Optimal downlink beamforming using semidefinite optimization. In *37th Annual Allerton Conference on Communication, Control, and Computing*, pages 987–996, 1999.

[BO01] M. Bengtsson and B. Ottersten. *Optimal and suboptimal transmit beamforming.* in Handbook of Antennas in Wireless Communications, CRC Press, Jan 2001.

[BQL21] Joshua Bassey, Lijun Qian, and Xianfang Li. A survey of complex-valued neural networks. arXiv preprint arXiv:2101.12249, 2021.

[Bur19] Andriy Burkov. *The Hundred-Page Machine Learning Book.* Andriy Burkov, 2019. [Online]. Available: http://themlbook.com/.

[BV04] S. Boyd and L. Vandenberghe. *Convex optimization.* Cambridge Univ. Press, 2004.

[Cis20] Cisco public. Cisco Annual Internet Report (2018-2023) White Paper. *Cisco Public Inf.*, Mar 2020.

[CSH10] S. Chen, S. Sugiura, and L. Hanzo. Semi-blind joint channel estimation and data detection for space-time shift keying systems. *IEEE Signal Process. Lett.*, 17(12):993–996, 2010.

[DB09] M. Ding and S. D. Blostein. MIMO minimum total MSE transceiver design with imperfect CSI at both ends. *IEEE Trans. Signal Proces.*, 57(3):1141–1150, 2009.

[DCHB18] Sebastian Dörner, Sebastian Cammerer, Jakob Hoydis, and Stephan ten Brink. Deep learning based communication over the air. *IEEE J. Selected Topics Signal Process.*, 12(1):132–143, 2018.

[Den14] Li Deng. A tutorial survey of architectures, algorithms, and applications for deep learning. *APSIPA Trans. Signal Inf. Process.*, 3, Jan 2014.

[DGAA13] G. Dartmann, X. Gong, W. Afzal, and G. Ascheid. On the duality of the max?min beamforming problem with per-antenna and per-antenna-array power constraints. *IEEE Trans. Vehicular Tech.*, 62(2):606–619, Feb 2013.

[DGCO19] S. Domouchtsidis, C. G. Tsinos, S. Chatzinotas, and B. Ottersten. Symbol-level precoding for low complexity transmitter architectures in large-scale antenna array systems. *IEEE Trans. Wireless Commun.*, 18(2):852–863, 2019.

[dMM08] Rodrigo de Miguel and Ralf. R. Muller. Convex precoding for vector channels in high dimensions. In *IEEE Int. Zurich Seminar Commun.*, pages 120–123, 2008.

[DTCO21] S. Domouchtsidis, C. G. Tsinos, S. Chatzinotas, and B. Ottersten. Joint symbol level precoding and combining for MIMO-OFDM transceiver architectures based on one-bit DACs and ADCs. *IEEE Trans. Wireless Communi.*, pages 1–1, 2021.

[ETS14] ETSI. Digital Video Broadcasting (DVB); Second generation framing structure, channel coding and modulation systems for Broadcasting, Interactive Services, News Gathering and other broadband satellite applications; Part 2: DVB-S2 Extensions (DVB-S2X). Deliverable 302.307-2, European Telecommunications Standards Institute (ETSI), 10 2014. Version 1.1.1.

[FLLL21] Y. Fan, A. Li, X. Liao, and V. C. M. Leung. Secure interference exploitation precoding in MISO wiretap channel: Destructive region redefinition with efficient solutions. *IEEE Trans. Inf. Forensics Security*, 16:402–417, 2021.

[Fos96] G. J. Foschini. Layered space-time architecture for wireless communication in fading environments when using multi-element antennas. *Bell Labs Tech. J.*, pages 14–51, 1996.

[FS11] S. A. A. Fakoorian and A. L. Swindlehurst. Solutions for the MIMO gaussian wiretap channel with a cooperative jammer. *IEEE Trans. Signal Process.*, 59(10):5013–5022, 2011.

[FW03] R.F.H. Fischer and C.A. Windpassinger. Improved MIMO precoding for decentralized receivers resembling concepts from lattice reduction. In *IEEE Global Telecommun. Conf. (GLOBECOM)*, volume 4, pages 1852–1856 vol.4, 2003.

[Gal06] Robert Gallager. *Chapter 8, course materials for 6.450 Principles of Digital Communications I.* MIT OpenCourseWare, Fall 2006.

[GHH+10] D. Gesbert, S. Hanly, H. Huang, S. Shamai Shitz, O. Simeone, and W. Yu. Multi-cell MIMO cooperative networks: A new look at interference. *IEEE J. Sel. Areas in Commun.*, 28(9):1380–1408, 2010.

[GJJV03] A. Goldsmith, S.A. Jafar, N. Jindal, and S. Vishwanath. Capacity limits of MIMO channels. *IEEE J. Sel. Areas Commun.*, 21(5):684–702, 2003.

[GLT+20] G. Gui, M. Liu, F. Tang, N. Kato, and F. Adachi. 6G: Opening new horizons for integration of comfort, security and intelligence. *IEEE Wireless Commun.*, pages 1–7, 2020.

[GM14] A. Garcia-Rodriguez and C. Masouros. Power-efficient Tomlinson-Harashima precoding for the downlink of multi-user MISO systems. *IEEE Trans. Commun.*, 62(6):1884–1896, 2014.

[GN06] Zhan Guo and P. Nilsson. Algorithm and implementation of the K-best sphere decoding for MIMO detection. *IEEE J. Sel. Areas Commun.*, 24(3):491–503, 2006.

[GN08] Satashu Goel and Rohit Negi. Guaranteeing secrecy using artificial noise. *IEEE Trans. Wireless Commun.*, 7(6):2180–2189, 2008.

[GSK05] M. Guillaud, D. T. M. Slock, and R. Knopp. A practical method for wireless channel reciprocity exploitation through relative calibration. In *In Proc. 8th Int. Symp. Signal Process. Appl., 2005.*, volume 1, pages 403–406, 2005.

[HDY+12] Geoffrey Hinton, Li Deng, Dong Yu, George E. Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick Nguyen, Tara N. Sainath, and Brian Kingsbury. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Process. Mag.*, 29(6):82–97, 2012.

[HFA19] J. M. Hamamreh, H. M. Furqan, and H. Arslan. Classifications and applications of physical layer security techniques for confidentiality: A comprehensive survey. *IEEE Commun. Surveys Tuts.*, 21(2):1773–1828, Secondquarter 2019.

[HKD+21] Alireza Haqiqatnejad, Jevgenij Krivochiza, Juan Carlos Merlano Duncan, Symeon Chatzinotas, and Björn Ottersten. Design optimization for low-complexity FPGA implementation of symbol-level multiuser precoding. *IEEE Access*, 9:30698–30711, 2021.

[HKO20] A. Haqiqatnejad, F. Kayhan, and B. Ottersten. Robust SINR-constrained symbol-level multiuser precoding with imperfect channel knowledge. *IEEE Trans. Signal Process.*, 68:1837–1852, 2020.

[HM72] H. Harashima and H. Miyakawa. Matched-transmission technique for channels with intersymbol interference. *IEEE Trans. Commun.*, 20(4):774–780, 1972.

[HPS05] B.M. Hochwald, C.B. Peel, and A.L. Swindlehurst. A vector-perturbation technique for near-capacity multiantenna multiuser communication-part II: perturbation. *IEEE Trans. Commun.*, 53(3):537–544, 2005.

[HWJL18] Hengtao He, Chao-Kai Wen, Shi Jin, and Geoffrey Ye Li. A model-driven deep learning network for MIMO detection. arXiv preprint arXiv:1809.09336, 2018.

[HZRS16] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *IEEE Conf. Computer Vision Pattern Recognition (CVPR)*, pages 770–778, 2016.

[JGMS15] Charles Jeon, Ramina Ghods, Arian Maleki, and Christoph Studer. Optimality of large MIMO detection via approximate message passing. In *IEEE Int. Symp. Inf. Theory (ISIT)*, pages 1227–1231, 2015.

[JHL17] Y. Jeon, S. Hong, and N. Lee. Blind detection for MIMO systems with low-resolution ADCs using supervised learning. In *2017 IEEE Int. Conf. Commun. (ICC)*, pages 1–6, 2017.

[JO08] Joakim Jalden and Björn Ottersten. The diversity order of the semidefinite relaxation detector. *IEEE Trans. Inf. Theory*, 54(4):1406–1422, 2008.

[JUN05] M. Joham, W. Utschick, and J. A. Nossek. Linear transmit processing in MIMO communications systems. *IEEE Trans. Signal Process.*, 53(8):2700–2712, Aug. 2005.

[JZ15] Rolf Johannesson and Kamil Sh. Zigangirov. *Fundamentals of Convolutional Coding*. Wiley-IEEE Press, 2015.

[JZR+17] C. Jiang, H. Zhang, Y. Ren, Z. Han, K. Chen, and L. Hanzo. Machine learning paradigms for next-generation wireless networks. *IEEE Wireless Commun.*, 24(2):98–105, Apr. 2017.

[KAHF20] M. Khani, M. Alizadeh, J. Hoydis, and P. Fleming. Adaptive neural signal detection for massive MIMO. *IEEE Trans. Wireless Commun.*, 19(8):5635–5648, 2020.

[Kay93] S. M. Kay. *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. Prentice Hall, 1 edition, 1993.

[KMA+19] J. Krivochiza, J. Merlano Duncan, S. Andrenacci, S. Chatzinotas, and B. Ottersten. FPGA acceleration for computationally efficient symbol-level precoding in multi-user multi-antenna communication systems. *IEEE Access*, 7:15509–15520, 2019.

[KMCO21] S. Kisseleff, W. A. Martins, S. Chatzinotas, and B. Ottersten. Symbol-level precoding with constellation rotation in the finite block length regime. *IEEE Commun. Lett.*, pages 1–1, 2021.

[KMW18] M. R. A. Khandaker, C. Masouros, and K. Wong. Constructive interference based secure precoding: A new dimension in physical layer security. *IEEE Trans. Inf. Forensics Security*, 13(9):2256–2268, Sept. 2018.

[KSM+16] A. Kalantari, M. Soltanalian, S. Maleki, S. Chatzinotas, and B. Ottersten. Directional modulation via symbol-level precoding: A way to enhance security. *IEEE J. Sel. Topics Signal Process.*, 10(8):1478–1493, Dec. 2016.

[KYK14] D. Kwon, W. Yeo, and D. K. Kim. A new precoding scheme for constructive superposition of interfering signals in multiuser MIMO systems. *IEEE Commun. Lett.*, 18(11):2047–2050, Nov 2014.

[Lar09] E. G. Larsson. MIMO detection methods: How they work [lecture notes]. *IEEE Signal Process. Mag.*, 26(3):91–95, 2009.

[LCCK13] H. Liang, R. Y. Chang, W. Chung, and S. Kuo. A reduced-complexity blind detector for MIMO system using K-means clustering algorithm. In *IEEE 77th Vehicular Tech. Conf. (VTC Spring)*, pages 1–5, 2013.

[LCK16] H. Liang, W. Chung, and S. Kuo. Coding-aided K-Means clustering blind transceiver for space shift keying MIMO systems. *IEEE Trans. Wireless Commun.*, 15(1):103–115, 2016.

[LCMC11] W. Liao, T. Chang, W. Ma, and C. Chi. QoS-based transmit beamforming in the presence of eavesdroppers: An optimized artificial-noise-aided approach. *IEEE Trans. Signal Process.*, 59(3):1202–1216, 2011.

[LCW17] Y. Liu, H. Chen, and L. Wang. Physical layer security for next generation wireless networks: Theories, technologies, and challenges. *IEEE Commun. Surveys Tuts.*, 19(1):347–376, 2017.

[LDL11] Ya-Feng Liu, Yu-Hong Dai, and Zhi-Quan Luo. Coordinated beamforming for MISO interference channel: Complexity analysis and efficient algorithms. *IEEE Trans. Signal Process.*, 59(3):1142–1157, 2011.

[LJ10] A. Lozano and N. Jindal. Optimum pilot overhead in wireless communication: A unified treatment of continuous and block-fading channels. *European Wireless Conf.*, pages 725–732, 2010.

[LLLS20] R. Liu, M. Li, Q. Liu, and A. L. Swindlehurst. Secure symbol-level precoding in MU-MISO wiretap systems. *IEEE Trans. Inf. Forensics Security*, 15:3359–3373, 2020.

[LM13] Q. Li and W. Ma. Spatially selective artificial-noise aided transmit optimization for MISO multi-eves secrecy rate maximization. *IEEE Trans. Signal Process.*, 61(10):2704–2717, 2013.

[LM16] K. L. Law and C. Masouros. Constructive interference exploitation for downlink beamforming based on noise robustness and outage probability. In *IEEE Int. Conf. Acoustics, Speech and Signal Process. (ICASSP)*, pages 3291–3295, 2016.

[LM17a] A. Li and C. Masouros. Exploiting constructive mutual coupling in P2P MIMO by analog-digital phase alignment. *IEEE Trans. Wireless Commun.*, 16(3):1948–1962, 2017.

[LM17b] A. Li and C. Masouros. Mutual coupling exploitation for point-to-point MIMO by constructive interference. In *IEEE Int. Conf. Commun. (ICC)*, pages 1–6, 2017.

[LM18a] K. L. Law and C. Masouros. Symbol error rate minimization precoding for interference exploitation. *IEEE Trans. Commun.*, 66(11):5718–5731, 2018.

[LM18b] A. Li and C. Masouros. Interference exploitation precoding made practical: Optimal closed-form solutions for PSK modulations. *IEEE Trans. Wireless Commun.*, 17(11):7661–7676, 2018.

[LM18c] Y. Liu and W. Ma. Symbol-level precoding is symbol-perturbed ZF when energy efficiency is sought. In *IEEE Int. Conf. Acoustics, Speech Signal Process. (ICASSP)*, pages 3869–3873, 2018.

[LML+20] A. Li, C. Masouros, X. Liao, Y. Li, and B. Vucetic. Multiplexing more data streams in the MU-MISO downlink by interference exploitation precoding. In *IEEE Wireless Commun. Netw. Conf. (WCNC)*, pages 1–6, 2020.

[LMLV19] A. Li, C. Masouros, Y. Li, and B. Vucetic. Interference exploitation precoding for multi-level modulations. In *IEEE Int. Conf. Acoustics, Speech Signal Process. (ICASSP)*, pages 4679–4683, 2019.

[LMS+10] Zhi-quan Luo, Wing-kin Ma, Anthony Man-cho So, Yinyu Ye, and Shuzhong Zhang. Semidefinite relaxation of quadratic optimization problems. *IEEE Signal Process. Mag.*, 27(3):20–34, 2010.

[LMV+21] A. Li, C. Masouros, B. Vucetic, Y. Li, and A. L. Swindlehurst. Interference exploitation precoding for multi-level modulations: Closed-form solutions. *IEEE Trans. Commun.*, 69(1):291–308, 2021.

[Lo99] T. K. Y. Lo. Maximum ratio transmission. *IEEE Trans. Commun.*, 47(10):1458–1461, Oct. 1999.

[LSK+20] A. Li, D. Spano, J. Krivochiza, S. Domouchtsidis, C. G. Tsinos, C. Masouros, S. Chatzinotas, Y. Li, B. Vucetic, and B. Ottersten. A tutorial on interference exploitation via symbol-level precoding: Overview, state-of-the-art and future directions. *IEEE Commun. Surveys Tuts.*, 22(2):796–839, Secondquarter 2020.

[LT02] Lizhong Zheng and D. N. C. Tse. Communication on the Grassmann manifold: a geometric approach to the noncoherent multiple-antenna channel. *IEEE Trans. Inf. Theory*, 48(2):359–383, 2002.

[MA07] C. Masouros and E. Alsusa. A novel transmitter-based selective-precoding technique for DS/CDMA systems. *IEEE Signal Process. Lett.*, 14(9):637–640, 2007.

[MA09] C. Masouros and E. Alsusa. Dynamic linear precoding for the exploitation of known interference in MIMO broadcast systems. *IEEE Trans. Wireless Commun.*, 8(3):1396–1404, 2009.

[Mas11] C. Masouros. Correlation rotation linear precoding for MIMO broadcast communications. *IEEE Trans. Signal Process.*, 59(1):252–262, 2011.

[Mas18] C. Masouros. Harvesting signal power from constructive interference in multiuser downlinks. *Wireless Inf. Power Transfer: A New Paradigm for Green Commun., Springer*, page 87–122, 2018.

[MGM08] Ralf R. Muller, Dongning Guo, and Aris L. Moustakas. Vector precoding for wireless MIMO systems and its replica analysis. *IEEE J. Sel. Areas Commun.*, 26(3):530–540, 2008.

[MH06] T. L. Marzetta and B. M. Hochwald. Fast transfer of channel state information in wireless systems. *IEEE Trans. Signal Process.*, 54(4):1268–1278, 2006.

[MMCO21] Abderrahmane Mayouche, Wallace A. Martins, Symeon Chatzinotas, and Björn Ottersten. Data-driven precoded MIMO detection robust to channel estimation errors. *IEEE Open J. Commun. Soc.*, 2:1144–1157, 2021.

[MMT+21a] Abderrahmane Mayouche, Wallace Martins, Christos Tsinos, Symeon Chatzinotas, and Bjorn Ottersten. Multi-antenna data-driven eavesdropping attacks and symbol-level precoding countermeasures. *IEEE Open J. Vehicular Tech.*, pages 1–1, 2021.

[MMT+21b] Abderrahmane Mayouche, Wallace A. Martins, Christos G. Tsinos, Symeon Chatzinotas, and Björn Ottersten. A novel learning-based hard decoding scheme and symbol-level precoding countermeasures. In *IEEE Wireless Commun.Netw. Conf. (WCNC)*, pages 1–6, 2021.

[MP17] J. Ma and L. Ping. Orthogonal AMP. *IEEE Access*, 5:2020–2033, 2017.

[MRS+13] C. Masouros, T. Ratnarajah, M. Sellathurai, C. B. Papadias, and A. K. Shukla. Known interference in the cellular downlink: a performance limiting factor or a source of green signal power? *IEEE Commun. Mag.*, 51(10):162–171, 2013.

[MSR12] C. Masouros, M. Sellathurai, and T. Ratnarajah. Interference optimization for transmit power reduction in Tomlinson-Harashima precoded MIMO downlinks. *IEEE Trans. Signal Process.*, 60(5):2470–2481, 2012.

[MST+19a] Abderrahmane Mayouche, Danilo Spano, Christos G. Tsinos, Symeon Chatzinotas, and Björn Ottersten. Machine learning assisted PHYSEC attacks and SLP countermeasures for multi-antenna downlink systems. *in Proc. IEEE Global Comm. Conf. (GLOBECOM)*, pages 1–6, 2019.

[MST+19b] Abderrahmane Mayouche, Danilo Spano, Christos G. Tsinos, Symeon Chatzinotas, and Björn Ottersten. SER-constrained symbol-level precoding for physical-layer security. *IEEE Conf. Commun. Netw. Security*, pages 1–5, June 2019.

[MST+20] A. Mayouche, D. Spano, C. G. Tsinos, S. Chatzinotas, and B. Ottersten. Learning-assisted eavesdropping and symbol-level precoding countermeasures for downlink MU-MISO systems. *IEEE Open J. Comm. Soc.*, 1:535–549, 2020.

[MVZJ18] V. Mavroeidis, K. Vishi, M. D. Zych, and A. Josang. The impact of quantum computing on present cryptography. *Int. J. Adv. Comput. Sci. Appl. (IJACSA)*, 9(3):405–414, Mar. 2018.

[MZ15a] C. Masouros and G. Zheng. Exploiting known interference as green signal power for downlink beamforming optimization. *IEEE Trans. Signal Process.*, 63(14):3628–3640, 2015.

[MZ15b] C. Masouros and G. Zheng. Power efficient downlink beamforming optimization by exploiting interference. In *IEEE Global Commun. Conf. (GLOBECOM)*, pages 1–6, 2015.

[NBB16] Eliya Nachmani, Yair Be'ery, and David Burshtein. Learning to decode linear codes using deep learning. In *54th Annual Allerton Conf. Commun., Control, and Computing (Allerton)*, pages 341–346, 2016.

[NCL+20] S. X. Ng, A. Conti, G. Long, P. Muller, A. Sayeed, J. Yuan, and L. Hanzo. Guest editorial advances in Quantum communications, computing, cryptography, and sensing. *IEEE J. Sel. Areas Commun.*, 38(3):405–412, Mar. 2020.

[NNT+20] L. V. Nguyen, D. T. Ngo, N. H. Tran, A. L. Swindlehurst, and D. H. N. Nguyen. Supervised and semi-supervised learning for MIMO blind detection with low-resolution ADCs. *IEEE Trans. Wireless Commun.*, 19(4):2427–2442, 2020.

[Ott96] B. Ottersten. Array processing for wireless communications. In *In Proc. 8th Workshop Stat. Signal Array Process.*, pages 466–473, 1996.

[PDM17] Alberto Del Pia, Santanu S. Dey, and Marco Molinaro. Mixed-integer quadratic programming is in NP. *Math. Program.*, 162(1-2):225–240, Mar. 2017.

[PHS05] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst. A vector-perturbation technique for near-capacity multiantenna multiuser communication-part I: channel inversion and regularization. *IEEE Trans. Commun.*, 53(1):195–202, 2005.

[Pla99] John C. Platt. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In *Advances in Large Margin Classifiers*, pages 61–74. MIT Press, 1999.

[PTSD18] P. Popovski, K. F. Trillingsgaard, O. Simeone, and G. Durisi. 5G wireless network slicing for eMBB, URLLC, and mMTC: A communication-theoretic view. *IEEE Access*, 6:55765–55779, 2018.

[RM16] L. Rose and M. Maso. Receiver-centric inter-cell interference cancellation in D2D-assisted networks. In *IEEE Globecom Workshops (GC Wkshps)*, pages 1–7, 2016.

[RM19] Jesse Read and Luca Martino. Probabilistic regressor chains with Monte Carlo methods. arXiv preprint arXiv:1907.08087, 2019.

[RPHF09] Jesse Read, Bernhard Pfahringer, Geoff Holmes, and Eibe Frank. Classifier chains for multi-label classification. In *Machine Learning and Knowledge Discovery in Databases*, pages 254–269. Springer Berlin Heidelberg, 2009.

[RRZ19] M. Rebato, L. Rose, and M. Zorzi. Performance assessment of MIMO precoding on realistic mmWave channels. In *IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, pages 1–6, 2019.

[SACO18] D. Spano, M. Alodeh, S. Chatzinotas, and B. Ottersten. Symbol-level precoding for the nonlinear multiuser MISO downlink channel. *IEEE Trans. Signal Process.*, 66(5):1331–1345, Mar. 2018.

[SB14] S. Shalev-Shwartz and S. Ben-David. *Understanding Machine Learning: From Theory to Algorithms.* Cambridge Univ. Press, 2014.

[SBC19] W. Saad, M. Bennis, and M. Chen. A vision of 6G wireless systems: Applications, trends, technologies, and open research problems. *IEEE Netw.*, 34(3):134–142, May/June 2019.

[Sch08] Tim Schenk. *RF Imperfections in High-rate Wireless Systems.* Springer Netherlands, 1 edition, 2008.

[SDW19] N. Samuel, T. Diskin, and A. Wiesel. Learning to detect. *IEEE Trans. Signal Process.*, 67(10):2554–2564, 2019.

[SF13] A. Schenk and R. F. H. Fischer. Noncoherent detection in massive MIMO systems. In *17th Int. ITG Workshop Smart Antennas (WSA)*, pages 1–8, 2013.

[SJU08] David A Schmidt, Michael Joham, and Wolfgang Utschick. Minimum mean square error vector precoding. *European Trans. Telecommun.*, 19(3):219–231, 2008.

[Smi04] G.S. Smith. A direct derivation of a single-antenna reciprocity relation for the time domain. *IEEE Trans. Antennas Propagation*, 52(6):1568–1577, 2004.

[SPBL20] O. Sholev, H. H. Permuter, E. Ben-Dror, and W. Liang. Neural network MIMO detection for coded wireless communication with impairments. In *2020 IEEE Wireless Commun. Netw. Conf. (WCNC)*, pages 1–8, 2020.

[SSH04] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt. Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels. *IEEE Trans. Signal Process.*, 52(2):461–471, Feb 2004.

[TDCO20] C. G. Tsinos, S. Domouchtsidis, S. Chatzinotas, and B. Ottersten. Symbol level precoding with low resolution DACs for constant envelope OFDM MU-MIMO systems. *IEEE Access*, 8:12856–12866, 2020.

[TEG04] Taesang Yoo, Eunchul Yoon, and A. Goldsmith. MIMO capacity with channel uncertainty: does feedback help? In *in Proc. IEEE Global Comm. Conf. (GLOBECOM)*, volume 1, pages 96–100 Vol.1, 2004.

[Tel99] E. Telatar. Capacity of multi-antenna gaussian channels. *Eur. Trans. Telecomm. ETT*, 10(6):585–596, Nov 1999.

[TG06] Taesang Yoo and A. Goldsmith. On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming. *IEEE J. Sel. Areas in Commun.*, 24(3):528–541, Mar. 2006.

[TK07] Grigorios Tsoumakas and Ioannis Katakis. Multi-label classification: An overview. *Int. J. Data Warehousing Mining (IJDWM)*, 3(3):1–13, July 2007.

[TKCO18] C. G. Tsinos, A. Kalantari, S. Chatzinotas, and B. Ottersten. Symbol-level precoding with low resolution dacs for large-scale array MU-MIMO systems. In *IEEE Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, pages 1–5, 2018.

[Tom71] M. Tomlinson. New automatic equaliser employing modulo arithmetic. *Electronics Lett.*, 7(5):138–139, 1971.

[Ver98] S. Verdu. *Multiuser detection.* Cambridge Univ. Press, 1998.

[Vit06] A. J. Viterbi. A personal history of the Viterbi algorithm. *IEEE Signal Process. Magazine*, 23(4):120–142, July 2006.

[WES08] A. Wiesel, Y. C. Eldar, and S. Shamai. Zero-forcing precoding and generalized inverses. *IEEE Trans. Signal Process.*, 56(9):4409–4418, Sep. 2008.

[WFVH04] C. Windpassinger, R.F.H. Fischer, T. Vencel, and J.B. Huber. Precoding in multiantenna and multiuser communications. *IEEE Trans. Wireless Commun.*, 3(4):1305–1316, 2004.

[Win87] J. Winters. On the capacity of radio communication systems with diversity in a rayleigh fading environment. *IEEE J. Sel. Areas Commun.*, 5(5):871–878, 1987.

[WKX+18] Y. Wu, A. Khisti, C. Xiao, G. Caire, K. Wong, and X. Gao. A survey of physical layer security techniques for 5G wireless networks and challenges ahead. *IEEE J. Sel. Areas in Commun.*, 36(4):679–695, Apr. 2018.

[WLXY13] Hui-Ming Wang, Miao Luo, Xiang-Gen Xia, and Qinye Yin. Joint cooperative beamforming and jamming to secure AF relay systems with individual power constraint and no eavesdropper's CSI. *IEEE Signal Process. Letters*, 20(1):39–42, 2013.

[WML+20] Z. Wei, C. Masouros, F. Liu, S. Chatzinotas, and B. Ottersten. Energy- and cost-efficient physical layer security in the era of IoT: The role of interference. *IEEE Commun. Mag.*, 58(4):81–87, 2020.

[WS16] P. Wang and R. Safavi-Naini. A model for adversarial wiretap channels. *IEEE Trans. Inf. Theory*, 62(2):970–983, 2016.

[WTZ+20] W. Wang, J. Tang, N. Zhao, X. Liu, X. Y. Zhang, Y. Chen, and Y. Qian. Joint precoding optimization for secure SWIPT in UAV-aided NOMA networks. *IEEE Trans. Commun.*, 68(8):5028–5040, 2020.

[WWN15] Hui-Ming Wang, Chao Wang, and Derrick Wing Kwan Ng. Artificial noise assisted secure transmission under training and feedback. *IEEE Trans. Signal Process.*, 63(23):6285–6298, 2015.

[WYX12] Hui-Ming Wang, Qinye Yin, and Xiang-Gen Xia. Distributed beamforming for physical-layer security of two-way relay networks. *IEEE Trans. Signal Process.*, 60(7):3532–3545, 2012.

[WZX15] Hui-Ming Wang, Tongxing Zheng, and Xiang-Gen Xia. Secure MISO wiretap channels with multiantenna passive eavesdropper: Artificial noise vs. artificial fast fading. *IEEE Trans. Wireless Commun.*, 14(1):94–106, 2015.

[XRS21] Q. Xu, P. Ren, and A. L. Swindlehurst. Rethinking secure precoding via interference exploitation: A smart eavesdropper perspective. *IEEE Trans. Inf. Forensics Security*, 16:585–600, 2021.

[YH15] S. Yang and L. Hanzo. Fifty years of MIMO detection: The road to large-scale MIMOs. *IEEE Commun. Surveys Tuts.*, 17(4):1941–1988, 2015.

[YL07] W. Yu and T. Lan. Transmitter optimization for the multi-antenna downlink with per-antenna power constraints. *IEEE Trans. Signal Process.*, 55(6):2646–2660, June 2007.

[YLH15] Y. Bengio Y. LeCun and G. Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015.

[YLJ18] Hao Ye, Geoffrey Ye Li, and Biing-Hwang Juang. Power of deep learning for channel estimation and signal detection in OFDM systems. *IEEE Wireless Commun. Lett.*, 7(1):114–117, 2018.

[Yua10] Jiang Yuan. *A practical guide to error-control coding using MATLAB*. Boston: Artech House, 2010.

[YWF+15] Z. Yu, Q. Wang, Y. Fan, H. Dai, and M. Qiu. An improved classifier chain algorithm for multi-label classification of big data analysis. *IEEE 17th Int. Conf. High Performance Comput. Commun., IEEE 7th Int. Symposium Cyberspace Safety Security, 12th Int. Conf. Embedded Softw. and Syst.*, pages 1298–1301, 2015.

[YXXL19] P. Yang, Y. Xiao, M. Xiao, and S. Li. 6G wireless communications: Vision and potential techniques. *IEEE Netw.*, 33(4):70–75, July/Aug 2019.

[YYX+17] L. You, P. Yang, Y. Xiao, S. Rong, D. Ke, and S. Li. Blind detection for spatial modulation systems based on clustering. *IEEE Commun. Lett.*, 21(11):2392–2395, 2017.

[ZE02] Bianca Zadrozny and Charles Elkan. Transforming classifier scores into accurate multiclass probability estimates. *in Proc. 8th ACM SIGKDD Int. Conf. Knowledge Discovery Data Mining*, July 2002.

[ZHL+21] J. Zhang, Y. He, Y. W. Li, C. K. Wen, and S. Jin. Meta learning-based MIMO detectors: Design, simulation, and experimental test. *IEEE Trans. Wireless Commun.*, 20(2):1122–1137, 2021.

[ZKM+14] G. Zheng, I. Krikidis, C. Masouros, S. Timotheou, D. Toumpakaris, and Z. Ding. Rethinking the role of interference in wireless networks. *IEEE Commun. Mag.*, 52(11):152–158, 2014.

[ZLLG18] Min-Ling Zhang, Yu-Kun Li, Xu-Ying Liu, and Xin Geng. Binary relevance for multi-label learning: an overview. *Springer Front. Comput. Sci.*, 12(1):191–202, Jan 2018.

[ZLZ+20] N. Zhao, Y. Li, S. Zhang, Y. Chen, W. Lu, J. Wang, and X. Wang. Security enhancement for NOMA-UAV networks. *IEEE Trans. Vehicular Tech.*, 69(4):3994–4005, 2020.

[ZO95] P. Zetterberg and B. Ottersten. The spectrum efficiency of a base station antenna array system for spatially selective transmission. *IEEE Trans. Veh. Tech.*, 44(3):651–660, 1995.

[ZWMM03] T. Zemen, J. Wehinger, C. Mecklenbrauker, and R. Muller. Iterative detection and channel estimation for MC-CDMA. In *IEEE Int. Conf. Commun. (ICC)*, volume 5, pages 3462–3466 vol.5, 2003.

[ZZZX20] Xiaodong Zhu, Xiangguo Zhang, Weiliang Zeng, and Jun Xie. Deep learning-based precoder design in MIMO systems with finite-alphabet inputs. *IEEE Commun. Lett.*, 24(11):2518–2521, 2020.