RIO

PUC

Dissertação de Mestrado

Design of Comparator Networks-Aided Low-Resolution MIMO Receivers via Convex Optimization

Reencarnación Quispe Achahuanco

Pontifícia Universidade Católica do Rio de Janeiro
Centro Técnico Científico
Departamento de Engenharia Elétrica

Rio de Janeiro, 19 de Setembro de 2025

Pontifícia
Universidade
Católica do
Rio de Janeiro

Dissertação de Mestrado

Design of Comparator Networks-Aided Low-Resolution MIMO Receivers via Convex Optimization

Reencarnación Quispe Achahuanco

Orientador: Professor Dr. Ing. Lukas Tobias Nepomuk Landau
Co-Orientador: Dr. Ahmed Mohammed Noreldien Elzakaloby Marai Ahmed

Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre em Engenharia Elétrica pelo programa de Pós-Graduação em Engenharia Elétrica de PUC-Rio, no Departamento de Engenharia Electrica.

Rio de Janeiro, 19 de setembro de 2025

Pontifícia
Universidade
Católica do
Rio de Janeiro

Design of Comparator Networks-Aided Low-Resolution MIMO Receivers via Convex Optimization

Reencarnación Quispe Achahuanco

Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre em Engenharia Eletrica da PUC-Rio. Aprovada pela Comissão examinadora abaixo:

Professor Dr. Ing. Lukas Tobias Nepomuk Landau
Orientador
Departamento de Engenharia Eletrica-PUC-Rio

Dr. Ahmed Mohammed Noreldien Elzakaloby Marai Ahmed
Co-Orientador
Departamento de Engenharia Eletrica-PUC-Rio

Professor Dr. Luis Alencar Reis Silva da Mello
Departamento de Engenharia Eletrica -PUC-Rio

Professor Dr. Alexander Beremiz Hilario Tacuri
Universidad Nacional de San Agustin (UNSA- Perú)

Dra. Clara Elizabeth Verdugo Munoz
Departamento de Engenharia Eletrica-PUC-Rio

Rio de Janeiro, 19 de setembro de 2025

Pontifícia
Universidade
Católica do
Rio de Janeiro

Todos os direitos reservados. A reprodução, total ou parcial, do trabalho é proibida sem autorização da universidade, da autora e do orientador.

Reencarnación Quispe Achahuanco
Engenheira Eletrônica pela Universidade Nacional de San Antonio Abad de Cusco em 2023.

Ficha Catalográfica

Quispe Achahuanco, Reencarnación

Design of comparator networks-aided low-resolution MIMO receivers via convex optimization / Reencarnación Quispe Achahuanco ; orientador: Lukas Tobias Nepomuk Landau ; co-orientador: Ahmed Mohammed Noreldien Elzakaloby Marai Ahmed. – 2025.
65f.: il. color.; 30 cm

Dissertação (mestrado)–Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Elétrica, 2025.
Inclui bibliografia

1. Engenharia Elétrica – Teses. 2. MIMO massivo. 3. ADCs de 1-bit. 4. Rede de comparadores. 5. Teorema de Bussgang. 6. Taxa de soma 7. Otimização convexa. I. Landau, Lukas Tobias Nepomuk. II. Ahmed, Ahmed Mohammed Noreldien Elzakaloby Marai. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. IV. Título.

CDD: 621.3

Acknowledgments

To my advisor and professor, Dr.-Ing. Lukas Tobias Nepomuk Landau, for all his support and guidance provided during the development of this thesis.

To my co-advisor, Dr. Ahmed Mohammed, for his reviews and patience.

To the evaluation committee for my thesis, Dr. Luis Silva Da Mello, Dr. Alexander Tacuri, and Dr. Clara Elizabeth Verdugo.

To my professors at UNSAAC, especially to Ing. Jorge Luis Arizaca Cusicuna and Ing. Milton Jhon Velasquez Ccuro.

To God, for always guiding and protecting me.

To my father, Antolín Quispe Orcotorio, for his unconditional love and sacrifice, for always taking care of me, for all his teachings, who thaugt me to face challenges with strength and perseverance.

To my mother, Luciana Achahuanco Rodriguez, for her unconditional love and sacrifice, for always taking care of me and teaching me kindness and compassion.

To my siblings Adonay, Eloim, Enmanuel, Marisol and Alejandro for all their unconditional support.

To my colleagues and friends from Laboratory C61: Diana, Erico, Luiz, Roberto, Jessica, Hassan, Saeed, Lu, Nayav, Mary Ann, and Iam.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Abstract

Reencarnación Quispe Achahuanco; Landau, Lukas T. N. (Advisor); Ahmed Mohammed Noreldien Elzakaloby Marai Ahmed (Co-Advisor) **Design of Comparator Networks-Aided Low-Resolution MIMO Receivers via Convex Optimization**; . 65p. Dissertação de mestrado – Departament of Electrical Engineering, Pontifícia Universidade Católica do Rio de Janeiro.

Considering low-resolution analog-to-digital converters (ADCs), multiple-input, multiple output (MIMO) is promising in terms of hardware cost and power consumption. In order to reduce the performance degradation caused by the low-resolution ADCs, additional information can be provided by comparator networks, which can be interpreted as virtual channels. The design of a comparator network corresponds to a difficult combinatorial problem. In this study, a comparator network with convex optimization is proposed that considers the sum rate as the objective function. The original problem formulation is approximated, relaxed, and expressed as a convex optimization problem. Numerical results show that the proposed optimization strategy yields a sum rate performance close to the optimal solution exhaustive search with polynomial complexity. Moreover, the simulations confirm that the sum rate significantly increases by adding comparator networks and that it can be further increased by using the proposed optimization strategy.

Keywords

Massive MIMO; 1-bit ADCs; comparator networks; Bussgang theorem; sum rate; convex optimization.

Resumo

Reencarnación Quispe Achahuanco; Landau, Lukas T. N.(Orientador); Dr. Ahmed Mohammed Noreldien Elzakaloby Marai Ahmed (Co-Orientador) **Projeto de Receptores MIMO de Baixa Resolução Auxiliados por Redes de Comparadores via Otimização Convexa** 65p. Dissertação de Mestrado- Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

Considerando conversores analógico-digitais (ADCs) de baixa resolução, o MIMO (Multiple-Input, Multiple Output) é promissor em termos de custo de hardware e consumo de energia. Para reduzir a degradação do desempenho causada pelos ADCs de baixa resolução, informações adicionais podem ser fornecidas por redes comparadoras, que podem ser interpretadas como canais virtuais. A rede de comparadores corresponde a um problema combinatório complexo. Neste estudo, é proposta uma rede de comparadores com otimização convexa que considera a taxa de soma como função objetivo. A formulação original do problema é aproximada, relaxada e expressa como um problema de otimização convexa. Os resultados numéricos mostram que a estratégia de otimização proposta produz um desempenho de taxa de soma próximo à busca exaustiva da solução ótima com complexidade polinomial. Além disso, as simulações confirmaram que a taxa de soma melhora significativamente com a adição de redes de comparadores e que pode melhorar ainda mais com o uso da estratégia de otimização proposta.

Palavras-chave

MIMO massivo; ADCs de 1-bit; rede de comparadores; teorema de Bussgang; taxa de soma; otimização convexa.

Table of contents

1 Introduction **12**
1.1 Motivation and Prior works 13
1.2 Contributions 14
1.3 Thesis Outline 14
1.4 Notation 15

2 MIMO System Model **16**
2.0.1 Comparator network design 19
2.1 Fully Comparator Network 20
2.2 Partially Connected Network 21

3 Quantization at the receiver side **23**
3.1 Quantization with 1-bit ADCs at the receiver side 23
3.2 Quantization at the receiver side with 1-bit ADCs and Comparator Networks 24

4 Sum Rate **25**
4.1 Sum Rate Lower Bound 25
4.2 Proposed Approximation of the Sum Rate with comparator aided low-resolution MIMO Receiver 26

5 Linear Detection **32**
5.1 Infinite Resolution 33
5.2 Detection MMSE with 1-bit Quantization ADCs 35
5.3 Detection Low Resolution Aware Minimum Mean Square Error (LRA-MMSE) with 1-bit ADCs and Comparator Networks 38
5.4 Computational complexity of typical Linear Detector 41

6 Comparator Selection Optimization Algorithms **44**
6.1 MMSE Greedy Search Network 44
6.2 Sequential SINR Algorithm Search 46
6.3 Proposed Convex Optimization Strategy 47

7 Numerical results **52**
7.1 Sum rate 52
7.2 CDF of Sum Rate 57
7.3 Execution Time in Matlab 58
7.4 Bit Error Rate (BER) 58

8 Conclusions **60**

A Appendix **64**
A.1 Linear Detection Derivations [1] 64

List of figures

Figure 2.1 QPSK diagram constellation and symbols [2] 17
Figure 2.2 The first stage of the MIMO system model considered 19
Figure 2.3 Base band system model of multi-user MIMO with 1-bit ADCs and comparator network [1] 20
Figure 2.4 Circuit representation of comparator network [1] 20

Figure 4.1 In the left side is shown the effective noise covariance $\mathbf{C}_{n'_R}$, in the right side show as scaled identity matrix 29
Figure 4.2 Comparison of Sum Rate lower bound $N_t = 2$ and $N_r = 2$ 29
Figure 4.3 Comparison of Quantized Comparator Network Capacity for 3x8 and 3x16 MIMO 31

Figure 7.1 Sum Rate of the proposed optimized connected comparator networks (CN) and benchmarks versus SNR for $N_t = 4$, $N_r = 12$, and different alpha values $\alpha = 2N_r, 4N_r, 8N_r$, considering 1-bit ADCs in MIMO-UL. 54
Figure 7.2 Sum Rate of the proposed optimized comparator networks (CN) and benchmarks versus number of receive antennas, SNR=15dB, $N_t = 4$,$\alpha = 4N_r$, considering 1-bit ADCs. 54
Figure 7.3 Comparison between the proposed optimized Comparator Network and the optimal exhaustive search sum rate for $N_t = 2$, $N_r = 3$ and $\alpha = 2N_r$. 55
Figure 7.4 Sum Rate of the proposed optimized connected comparator networks (CN) and benchmarks versus SNR for $N_t = 4$, $N_r = 12$, and $\alpha = 2N_r$, considering 1-bit ADCs in MIMO-UL. 55
Figure 7.5 Sum Rate of the proposed optimized connected comparator networks (C.N.) and benchmarks versus SNR for $N_t = 4$, $N_r = 4$, and alpha $\alpha = 2N_r$, considering 1-bit ADCs in MIMO-UL. 56
Figure 7.6 Cumulative Distribution Function (CDF) of the sum rate for the proposed optimized Comparator Networks (CN) and benchmarks configurations, with $N_t = 4$, $N_r = 16$, and $\alpha = 4N_r$, SNR = 20dB, The scenario considers 1-bit ADCs in MIMO-UL 2000 channel realizations. 57
Figure 7.7 BER performance of MMSE, Matched Filter, and Zero Forcing detectors with 1-bit quantization across different SNR values. 59
Figure 7.8 Performance of BER with comparator network $N_r = 4$ for $\alpha = 2N_r$ 59

List of tables

Table 5.1 Comparison of Linear Detection Techniques 35
Table 5.2 Computational complexity for a typical linear detector 41
Table 5.3 Computational complexity with LRA-MMSE [3] 42
Table 5.4 Computational cost with LRA-MMSE hardware cost $N_t = 2$ $N_r = 10$ [3] 43

Table 6.1 Comparison between optimization strategies comparator selection methods 51

Table 7.1 Comparison of Algorithm Execution Times 58

List of Abbreviations

ADCs Analog-to-Digital Converter

AGC Automatic Gain Control

AWGN Additive White Gaussian Noise

BER Bit Error Rate

BS Base Station

CAPES Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

CN Comparator Network

CSI Channel State Information

DEE Departamento de Engenharia Elétrica

IDD Iterative Detection and Decoding

LRA Low-Resolution Aware

LTE Long-Term Evolution

MF Matched Filter

MIMO Multiple-Input Multiple-Output

MMSE Minimum Mean Square Error

$\Sigma\Delta$ sigma delta

QAM Quadrature Amplitude Modulation

QPSK Quadrature Phase Shift Keying Modulation

SINR Signal Interference Noise to Radio

ZF Zero Forcing

1
Introduction

Multiple Input Multiple Output (MIMO) technology allows the simultaneous transmission of multiple data streams over time, frequency, and space [4]. MIMO applications include 5G, 6G, the Internet of Things, data augmentation among others [4]. However, one of the main issues in MIMO systems is the increased cost and power consumption of high-speed and wide-bandwidth analog-to-digital converters (ADCs). The power consumption of the ADCs scales with the sampling frequency and exponentially with the number of bits in the resolution [5]. This introduces critical design trade-offs, especially for wireless systems where energy efficiency is crucial, such as battery-powered devices, or large-scale infrastructure such as base stations [6]. Despite these efforts, finding an optimal balance between the system cost and performance remains challenging. One approach to reduce complexity and power consumption is the adoption of low-resolution 1-bit quantization, particularly in large-scale MIMO systems. However, 1-bit quantization is limited to only two binary values (+1 and -1), which restricts the amount of information that can be transmitted through the network. This limitation arises information loss in 1-bit ADCs based systems, as discussed in [6] and [7]. In particular, in MIMO systems with low resolution, the performance degradation in terms of sum rate is a known issue. However, the reduction is not drastic in some cases [7]. To mitigate this effect and improve performance, several techniques have been proposed, such as temporal oversampling [8] and spatial sigma-delta $\Sigma\Delta$ [9], [10].

An alternative approach leverages the use of an additional comparator network, as presented in recent studies [11], [12], and [13]. This work considers a virtual extension of the channel representation by incorporating switches and comparators into the Multi User-MIMO system. In this context, this study aims to develop a convex optimization-based comparator network configuration for improving the sum rate under low-resolution constraints by addressing approximations for modeling the performance in MIMO communication systems, including the effect of comparator networks. We develop a similar strategy as proposed in [14], which is an optimization technique for antenna selection, with low computational complexity. The key objective is to approach the performance of a fully connected comparator network with a reduced number of

comparators. Building on this foundation, we formulate a convex optimization technique within the extended MIMO system model for comparator selection.

This work relies on the utilization of Bussgang theorem [15], [16] to the system of the comparator networks with 1-bit ADCs, to obtain a sum rate lower bound expression considering Gaussian input. Furthermore, we develop an approximation assuming uncorrelated noise samples. The sum lower bound expression is the objective function in the formulation of a convex optimization strategy for the comparator network configuration. The numerical results compare the sum rate of the proposed convex optimization method with the exhaustive search solution that achieves optimal performance and practical comparator network configurations, namely, randomly and fully connected comparator networks, which confirms the efficiency of the proposed method. The applications of MIMO with comparator network and 1-bit quantization are in 5G, 6G in Internet of Things (IoT), low cost MIMO radars.

1.1 Motivation and Prior works

One of the main issues in MIMO systems is the increased cost and energy consumption associated with high-speed and wide-bandwidth Analog-to-Digital Converters (ADCs). However, energy consumption increases exponentially as the sampling rate and resolution of ADCs increase to meet the demands of higher data rates and bandwidth. The power consumption of the ADCs scales with the number of bits in the resolution and sampling frequency. This introduces critical design trade-offs, especially for wireless systems where energy efficiency is crucial, such as battery-powered devices or large-scale infrastructure like base stations. Therefore, there is growing interest in reducing the resolution of ADCs as a potential solution to mitigate these challenges. Low-resolution ADCs, including a 1-bit quantization approach, significantly reduce power consumption and cost while maintaining acceptable system performance. These techniques aim to balance the potential loss of information owing to coarse quantization while applying the advantages of reduced energy consumption and hardware complexity. Related works have applied Iterative Detection and Decoding (IDD) [17]. In [18], an analysis of the application of 1-bit quantization ADCs/DCAs in a MIMO system at the transmitter and receiver sides [19]. The study [20] focuses on the challenge of multi-user data detection in massive MIMO systems equipped with 1-bit ADCs. The paper [21] provides a numerical evaluation of the expected value of the soft-estimated symbols with zero-forcing (ZF) and minimum mean squared error (MMSE) receivers for a multi-U-E setting with Rayleigh fading [22]. In [20], a MIMO

uplink with a Rayleigh fading channel distribution and a single user antenna was proposed, and 16-QAM and 1-bit quantization were considered. The 1-bit quantization exhibits good performance at high frequencies.

1.2 Contributions

The main contribution of this work is to understand and perform uplink channel simulations with low-resolution quantization using a linear strategy. This work also develops a sum-rate bound for uplink systems for low-resolution MIMO reception using real-valued notation adding the comparator network. Another contribution of this study is the extension of the sum-rate dimension for uplink systems with a comparator-assisted low-resolution MIMO receiver. This thesis presents the optimization of the sum rate of 1-bit quantization MIMO. The use of the comparator network in a low-quantization system helps improve the sum rate by obtaining additional information through the creation of virtual channels, which increases the complexity when using a fully connected network. Options are sought to reduce complexity and achieve results similar to those of fully connected networks; among these options are random and partially connected networks. Additionally, to improve the sum rate, a convex optimization problem is formulated with the sum rate as the objective function. A key contribution of this thesis is the improvement in the sum rate, achieved through the Bussgang theorem and the proposed convex optimization algorithm via relaxation. The other contribution of this study is the formulation of an optimization problem for comparator selection. This study also develops a strategy for computing a feasible solution based on the relaxed solution, and obtains the numerical results in terms of the sum rate lower bound and BER respectively.

1.3 Thesis Outline

This work is organized as follows. Chapter 2 describes the MIMO system model with 1-bit quantization and a comparator network, and describes different types of configurations to connect the comparator with the signal. Chapter 3 presents 1-bit quantization ADCs. Chapter 4 introduces the sum rate approximation through Bussgang decomposition. Chapter 5 presents the techniques for detection on the reception side with and without a 1-bit coarse quantization, including after the comparator network. Chapter 6 introduces the different types of optimization strategies with comparator network and low resolution as Greedy MMSE algorithm, SINR algorithm and present the

proposed convex optimized strategy to sum rate maximization. Chapter 7 shows the numerical results for the detection and sum rate comparison for the different comparator network techniques, Cumulative Distribution Function (CDF) of sum rate comparison of benchmarks and BER. Finally, in Chapter 8 conclusions are presented.

1.4 Notation

The following notations are considered throughout the dissertation. $\mathbb{E}\{\cdot\}$ denotes the expected value. Scalar values are represented by lowercase letters in regular font, for example, a. Complex vectors and matrices are denoted by bold lowercase and uppercase letters, respectively. $\mathbf{I}$ represents the identity matrix. The quantization operator is represented as $\mathcal{Q}(\cdot)$, $\|\cdot\|_2^2$ represents the squared Euclidean norm.

2
MIMO System Model

This chapter presents the MIMO system model considered. This study is based on the MIMO system model, considered in [11] with N_r received antennas and N_t transmit antennas, as shown in Fig. 2.3. MIMO antenna selection offers greater flexibility than single-input Single-Output (SISO) systems in Up Link transmission.

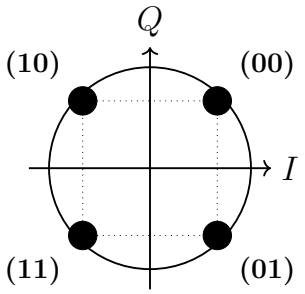
MIMO systems are versatile and can use many modulation techniques, depending on the application and environment. To choose the modulation scheme, the following characteristics can be considered

- In Quadrature Phase Shift Keying (QPSK) and Quadrature Amplitude Modulation (QAM) schemes, it can achieve high spectral efficiency values.
- The requirements of the application determine the priorities. For example, video streaming might focus on achieving high data rates and quality, whereas IoT applications might emphasize low power consumption and reliability.

QPSK and QAM (16-QAM and 64-QAM) are common choice variants owing to their balance between efficiency and robustness. The QPSK technique has applications in communications that transmit with phase deviation. QPSK duplicates the data rate compared with BPSK. Each symbol is represented by a quadrature deviation. The real part (I) and the imaginary part (Q) are both discrete values and token-specific value according to the phase angle of the symbol. In QPSK, the Gray Code is used to minimize errors in multiple symbols. It simplifies the decoding process in digital systems [2]. The objective of the Hamming distance is to minimize the distance between the symbols in the constellation map. The Gray code is used to ensure that the Hamming distance between adjacent symbols is 1, which reduces various transmission errors.

- Gray Code: QPSK modulation employs a Gray code map, a representation to differentiate two consecutive symbols in one bit.

- Hamming distance is the comparison between two consecutive symbols with the same length and measures the difference of bits: It is used in QPSK to reduce the error probability and improve the data transmission.



Symbols:
(00): $1 + j$
(01): $1 - j$
(11): $-1 - j$
(10): $-1 + j$

Figure 2.1: QPSK diagram constellation and symbols [2]

The wireless channel response is typically modeled as a random process due to multipath propagation caused by reflection, diffraction scattering, and refraction, as well as signal attenuation from absorption. These effects lead to variations in amplitude, phase, and delay, resulting in fading. Stochastic processes characterize the behavior of a random MIMO channel by considering temporal variations, such as in the case of a Rayleigh fading channel.

The compact vector representation is given by (2-1)

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}. \tag{2-1}$$

The channel matrix is denoted by $\mathbf{H}$ represents the communication channel between the transmit and receive antennas in the MIMO system. $\mathbf{H}$ operates on $\mathbf{x}$ to generate $\mathbf{y}$, and $\mathbf{n}$ adds the variation in the output.
The channel $\mathbf{H}$ has dimensions $\mathbb{C}^{N_r \times N_t}$. The received vector is denoted by $\mathbf{y}$, and has dimensions $\mathbb{C}^{N_r \times 1}$.
The transmitted symbol vector $\mathbf{x}$ is complex and has dimensions $\mathbb{C}^{N_t \times 1}$.
The complex noise vector is denoted by $\mathbf{n}$, and has dimensions $\mathbb{C}^{N_r \times 1}$ with a complex Gaussian distribution for each antenna element $\mathcal{CN}(0, \sigma_n^2)$.

The following representation is used to explain how each component of

the different channel coefficients contributes to the received signal. A simple visualization vector expansion of the MIMO system is represented in (2-2)

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{N_r} \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1N_t} \\ h_{21} & h_{22} & \cdots & h_{2N_t} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_r1} & h_{N_r2} & \cdots & h_{N_rN_t} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{N_t} \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_{N_r} \end{bmatrix}. \tag{2-2}$$

The input to the channel matrix consists of independent and identically distributed complex Gaussian symbols with zero mean and unit variance. In this system model, the transmission power was divided between both the real part $\Re\{\mathbf{x}_i\}$ and the imaginary part $\Im\{\mathbf{x}_i\}$ of each symbol $\mathbf{x}_i$. These parts of the transmitted symbols are normalized as

$$\mathbb{E}\left\{\Re\{\mathbf{x}_i\}^2\right\} = \mathbb{E}\left\{\Im\{\mathbf{x}_i\}^2\right\} = \tfrac{1}{2}\sigma_{\mathbf{x}}^2, \tag{2-3}$$

where $\Re\{\mathbf{x}_i\}$ and $\Im\{\mathbf{x_i}\}$ are the real and imaginary parts of the complex symbol $\mathbf{x}_i$, $\sigma_{\mathbf{x}}^2$ is the variance of the transmitted symbols. The factor $\frac{1}{2}$ arises because the total power of the complex symbol $\mathbf{x}_I$, which are statistically independent [11]. To express each component in real values, each element in (2-1) can be expanded as follows

$$\begin{bmatrix} \Re\{\mathbf{y}\} \\ \Im\{\mathbf{y}\} \end{bmatrix} = \begin{bmatrix} \Re\{\mathbf{H}\} & -\Im\{\mathbf{H}\} \\ \Im\{\mathbf{H}\} & \Re\{\mathbf{H}\} \end{bmatrix} \begin{bmatrix} \Re\{\mathbf{x}\} \\ \Im\{\mathbf{x}\} \end{bmatrix} + \begin{bmatrix} \Re\{\mathbf{n}\} \\ \Im\{\mathbf{n}\} \end{bmatrix}, \tag{2-4}$$

where $\Re\{.\}$ indicates real values and $\Im\{.\}$ indicates imaginary representation. This expanded representation doubles the dimension due to the inclusion of both real and imaginary parts. It is important to note that the last equation offers advantages and reduces the complexity of the process. The following equation represents the equivalent real value representation of the equation (2-4)

$$\mathbf{y}_{\mathrm{R}} = \mathbf{H}_{\mathrm{R}}\mathbf{x}_{\mathrm{R}} + \mathbf{n}_{\mathrm{R}}, \tag{2-5}$$

where the equivalent representation for the received vector $\mathbf{y}_{\mathrm{R}}$ has dimensions $\mathbb{R}^{2N_r \times 1}$. The real compact valued channel matrix $\mathbf{H}_{\mathrm{R}}$ has dimensions $\mathbb{R}^{2N_r \times 2N_t}$. The real compact valued transmitted vector $\mathbf{x}_{\mathrm{R}}$ has dimensions $\mathbb{R}^{2N_t \times 1}$. The real compact-valued noise vector $\mathbf{n}_{\mathrm{R}}$ has dimensions $\mathbb{R}^{2N_r \times 1}$ with Gaussian distribution with mean 0 and variance $\frac{\sigma_n^2}{2}$.

In the first stage of the system model, an MIMO system with N_t transmitters

is considered. At the receiver, an expanded real-value representation is used, as shown in Figure 2.2.

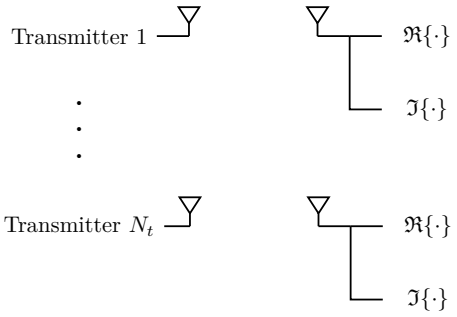


Figure 2.2: The first stage of the MIMO system model considered

In the second stage, the MIMO system model incorporates 1-bit analog to digital converters (ADCs), and a comparator network is added to provide additional information about the received signal. In the third stage, the output of the 1-bit quantization and comparator networks is the input of the Low Resolution Aware (LRA)-MMSE detector; the whole system model is shown in Figure 2.3. The overall system model for the proposed multi-input multi-output (MIMO) receiver architecture system model with 1-bit quantization reduces power consumption is simultaneous with the comparator network 2.3. The channel was considered a statistical and perfect Channel State Information (CSI) at the receiver. The MIMO system model 1-bit quantization and comparator network makes the subtraction from two different antennas, and then compares it with a umbral in the output, which has binary values considering the real value representation [20]. In [23], the different states that can be taken in the network of comparators are shown. The graph 2.4 represents the circuit representation of the comparator network.

2.0.1 Comparator network design

A comparator network system compares pairs of input signals and produces binary symbol outputs, typically representing whether one input is greater or less than the other. Consequently, these networks are commonly applied in signal processing, and decision-making applications. In communication systems, the comparator network in low-resolution MIMO receivers has benefits mainly in terms of power consumption and low computational cost. The matrix $\mathbf{B}$ represents the relationship between the received signals and extended versions of those signals. However, combining these elements with appropriate detection techniques makes it possible to recover the transmitted

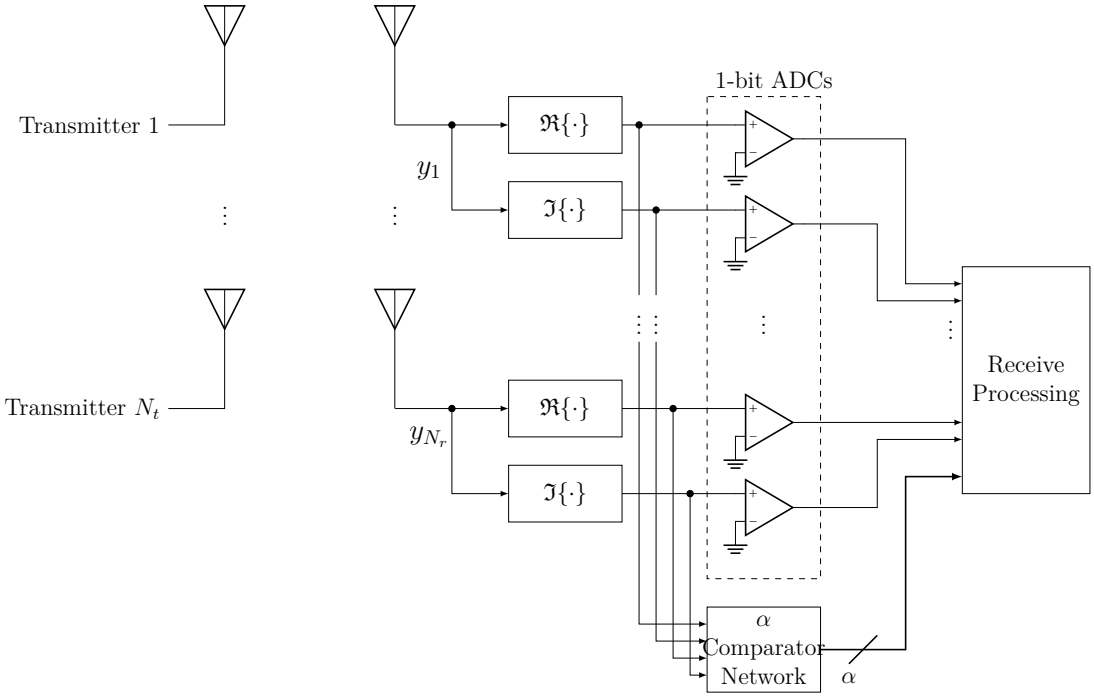


Figure 2.3: Base band system model of multi-user MIMO with 1-bit ADCs and comparator network [1]

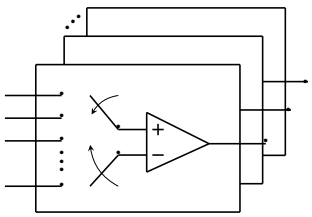


Figure 2.4: Circuit representation of comparator network [1]

signal and optimize the system. The matrix $\mathbf{B}$ is presented in the following equation

$$\mathbf{B} = \left(\begin{bmatrix} \mathbf{I}_{2N_r} \\ \mathbf{B'} \end{bmatrix} \right), \tag{2-6}$$

where $\mathbf{I}_{2N_r}$ is an identity matrix with dimension in $\mathbb{R}^{2N_r \times 2N_r}$. $\mathbf{B'} \in \mathbb{R}^{\alpha \times 2 \cdot Nr}$ is the matrix related to the processing or filtering of the signal after passing through the quantization network or comparator network.

2.1 Fully Comparator Network

This configuration implies the connection between all antennas [1]. The fully connected comparator matrix $\mathbf{B'}$ is constructed to include all possible pairwise connections

$$\mathbf{B}' = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 & 0 & 0 & \cdots & 0 \\ -1 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & -1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & 0 & 0 & \cdots & -1 \\ 0 & 0 & 0 & \cdots & 1 & -1 \end{bmatrix}. \tag{2-7}$$

Each comparator is represented for each row, and each row contains 1 and -1, and the other components are zero. The main drawback of the fully connected network is the significant number of comparators α_{full}, which grows approximately proportional to the square of the number of receive antennas N_r [1].

For example, if a system with $N_r = 2$ receive antennas is considered, it requires $\alpha_{\text{full}} = \binom{4}{2} = 6$ comparator in this network

$$\mathbf{B}' = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{bmatrix}, \tag{2-8}$$

where $\mathbf{B}' \in \mathbb{R}^{\alpha \times 2N_r}$ represents the normalized comparator network [1].
In this network, every pair of received signals is compared, which means that the number of comparators is given as follows

$$\alpha_{\text{full}} = \binom{2N_r}{2} = N_r \cdot (2N_r - 1), \tag{2-9}$$

where α_{full} number of comparators networks.

2.2 Partially Connected Network

The partially connected network configuration involves choosing a specific subset of elements based on predetermined criteria or performance metrics. This method maximizes a particular objective. One option as example for a partially connected network is represented in the following matrix [1]

$$\mathbf{B}'_{\alpha_{\text{partial}}} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{bmatrix}. \tag{2-10}$$

The random comparator network α_{partial} selects some rows from α_{fully}. The random comparator network is a combinatorial problem and exists in different combinations $\alpha_{\text{partial}} = \binom{N_r \cdot (2 \cdot N_r - 1)}{\alpha_{\text{partial}}}$.

3
Quantization at the receiver side

In the UL-MIMO system, the quantization process at the received side is performed with different resolutions. Typically low resolution ADCs are represented as comparators.

3.1
Quantization with 1-bit ADCs at the receiver side

In UL-MIMO systems received signal at the receiver side was quantized by one bit, significantly affecting the detection process. Let $\mathbf{y}_R \in \mathbb{R}^{N_r}$ be the actual received signal, and $\mathcal{Q}(\mathbf{y})$ represent the nonlinear quantization function applied element-by-element to $\mathbf{y}_R$. In the case of one-bit quantization, the function $\mathcal{Q}(\cdot)$ produces values constrained to $\{-1, +1\}$, depending on whether the input is below or above a predefined threshold is typically considered zero. This process introduces significant challenges in transmission due to information loss. The quantization of the received signal $\mathbf{y}$ is represented by

$$\mathbf{z}_{\mathcal{Q}} = \mathcal{Q}(\mathbf{y}) = \mathcal{Q}(\mathbf{H}\mathbf{x} + \mathbf{n}), \tag{3-1}$$

where the transmitted vector is denoted by $\mathbf{x}$, and has dimensions $\mathbb{C}^{N_t \times 1}$. $\mathcal{Q}(\cdot)$ represents the non linear 1-bit quantization constrained to $\{-1, +1\}$.

Quantization is a non-linear process applied to real and imaginary parts separated such as shown in the following equation

$$\mathbf{z}_{\mathcal{Q}} = \mathcal{Q}(\mathbf{y}) = \mathcal{Q}(\Re\{\mathbf{y}\}) + j\mathcal{Q}(\Im\{\mathbf{y}\}). \tag{3-2}$$

where $\mathbf{z}_{\mathcal{Q}}$ is the quantized signal at the output of 1-bit ADCs. Additional quantization noise was introduced in the quantization process introduces a non linearity.

The Bussgang decomposition allows an approach to the noise quantization of such a linear function; this process enables management and mathematical analysis [24]. This linear approximation of the noise quantization is represented by $\mathbf{n}_q$. The following equation represents this approach to the linear noise quantization employing Bussgang decomposition

$$\mathbf{z}_{\mathcal{Q}} = \mathbf{A}(\mathbf{H}\mathbf{x} + \mathbf{n}) + \mathbf{n}_q, \tag{3-3}$$

The matrix $\mathbf{A}$ is a linear operator. The transmitted vector is denoted by $\mathbf{x}$, and has dimensions $\mathbb{C}^{N_t \times 1}$. And, $\mathbf{n}_q$ is an equivalent linear approach to quantization noise.

3.2 Quantization at the receiver side with 1-bit ADCs and Comparator Networks

Quantization at the receiver side, a 1-bit ADC and comparator network in MIMO systems, enables a simple, low-cost solution. The quantization process involves both the received signal $\mathbf{y}_{\mathrm{R}}$ and the comparator network. Each row of the matrix B represents a comparator network. Compares two received signals and quantizes the differences. The network compares pairs of components from the received signal and outputs quantized values ± 1 [3]. The following equation (3-7) represents the quantization at the output of the ADCs with the comparator network. The subtraction in each comparator is through the multiplication of $\mathbf{B}'$ with $\mathbf{y}_{\mathrm{R}}$ [11], which is described in the following equation

$$\mathbf{z}_{\mathcal{Q}}^{\mathrm{R}} = \mathcal{Q}\left(\begin{bmatrix} \mathbf{y}_{\mathrm{R}} \\ \mathbf{B}' \cdot \mathbf{y}_{\mathrm{R}} \end{bmatrix}\right), \tag{3-4}$$

where $\mathbf{B}'$ with corresponds to a modified matrix related to the processing or filtering of the signal after it has been passed through the quantization network or comparator network. Factorizing the received real-valued signal, the equivalent form of representation is

$$\mathbf{z}_{\mathcal{Q}}^{\mathrm{R}} = \mathcal{Q}\left(\left[\mathbf{I}_{2N_r}\mathbf{B}'\right]^T \cdot \mathbf{y}_{\mathrm{R}}\right). \tag{3-5}$$

The 1-bit samples and comparator output are represented in the following equation (3-6); both are considered in the detection process. Taking into account that matrix $\mathbf{B}$ [11] is the representation of the concatenation of the identity matrix with $\mathbf{B}$, the final equation to represent quantization at the output of the ADCs and the comparator network reads as

$$\mathbf{z}_{\mathcal{Q}}^{\mathrm{R}} = \mathcal{Q}\left(\mathbf{B} \cdot \mathbf{y}_{\mathrm{R}}\right), \qquad \text{where} \tag{3-6}$$

$$\mathbf{B} = \begin{bmatrix} \mathbf{I}_{2N_r} \\ \mathbf{B}' \end{bmatrix}. \tag{3-7}$$

4
Sum Rate

This section presents the sum rate for the proposed receiver with 1-bit ADCs and a comparator network, as the sum rate strongly depends on the comparator network configuration. Employing 1-bit quantization renders the system inherently non-linear. To be described as a linear process, was employed the Bussgang decomposition.

4.1
Sum Rate Lower Bound

In this section, we develop a sum rate expression for the proposed receiver with 1-bit ADCs and a comparator network, as the sum rate strongly depends on the comparator network configuration. Based on the Bussgang theorem, this approach helps preserve the correlation of the original signal despite the non-linearity. The transmission capacity decreases with 1-bit quantization due to the low resolution in signal detection [11]. A virtual channel is created because each comparator with a different threshold generates an observation. This is mathematically expressed as an antenna. These virtual channels allow the MIMO system to exploit its spatial diversity. Even if no additional antennas are physically added, the comparator network creates a virtual extension in the number of signals that can be processed and optimized. These are additional lines of digital information that improve the representation of the original MIMO channel. The addition of a few comparators improves performance without the need for a more complex structure of ADCs. Each comparator generates a new digital output based on the difference between two signals. These new outputs can be considered information channels added to the original system. Although no new transmitted antennas are used, they provide additional information to the receivers [3].

The quantization and comparator operation step is a nonlinear process. However, this process can be described as a linear process through the Bussgang decomposition [7]. With this, the sum rate expression, considering 1-bit ADCs and a comparator network, can be given as:

$$I(\mathbf{x};\mathbf{z}_Q) = \frac{1}{2}\log_2\left|\left(\mathbf{I}_{2Nr+\alpha} + \mathbf{C}_{\eta'_{\mathrm{R}}}^{-1}\mathbf{H'}_{\mathrm{R}}\mathbf{C}_{\mathbf{x}_{\mathrm{R}}}\mathbf{H'}_{\mathrm{R}}^{T}\right)\right|, \tag{4-1}$$

where $I(\mathbf{x};\mathbf{z}_Q)$ is the mutual information between the input $\mathbf{x}$ and the output $\mathbf{z}_Q$, considering Gaussian input, and is used as a lower bound of the sum rate. $\mathbf{C}_{\eta'_{\mathrm{R}}}$ is the covariance matrix of the effective noise including the quantization effect in terms of quantization noise. $\mathbf{H'}_{\mathrm{R}}$ is the effective channel taking into account the 1-bit ADCs and the comparator network. $\mathbf{C}_{\mathbf{x}_{\mathrm{R}}}$ is the covariance matrix of the transmitted signal. In the following, it is described in detail how to compute the expression for noise covariance and effective channel.

4.2 Proposed Approximation of the Sum Rate with comparator aided low-resolution MIMO Receiver

To calculate $\mathbf{H'}_{\mathrm{R}}$, we start with the expression of the autocorrelation of the quantized signal wich depend of the autocorrelation at the input, which is presented in [25] as

$$\mathbf{C}_{\mathbf{z}_{\mathrm{R}}} = \mathbf{B}\mathbf{H}_{\mathrm{R}}\mathbf{C}_{\mathbf{x}_{\mathrm{R}}}\mathbf{H}_{\mathrm{R}}^{T}\mathbf{B}^{T} + \mathbf{B}\mathbf{C}_{\mathbf{n}_{\mathrm{R}}}\mathbf{B}^{T}. \tag{4-2}$$

where $\mathbf{C}_{\mathbf{z}_{\mathrm{R}}}$ is the autocorrelation matrix of the received signal before the quantization step, given by

$$\mathbf{C}_{\mathbf{z}_{\mathcal{Q}}} = \frac{2}{\pi}\sin^{-1}\left(\mathrm{diag}(\mathbf{C}_{\mathbf{z}_{\mathrm{R}}})^{-\frac{1}{2}}\mathbf{C}_{\mathbf{z}_{\mathrm{R}}}\,\mathrm{diag}(\mathbf{C}_{\mathbf{z}_{\mathrm{R}}})^{-\frac{1}{2}}\right), \tag{4-3}$$

Considered $\mathbf{K}_{\mathrm{R}} = \mathrm{diag}(\mathbf{C}_{\mathbf{z}_{\mathrm{R}}})^{-\frac{1}{2}}$ the autocorrelation matrix of the quantized received signal from (4-3) reads as follows

$$\mathbf{C}_{\mathbf{z}_{\mathcal{Q}}} = \frac{2}{\pi}\sin^{-1}\left(\mathbf{K}_{\mathrm{R}}\mathbf{C}_{\mathbf{z}_{\mathrm{R}}}\mathbf{K}_{\mathrm{R}}\right). \tag{4-4}$$

The effective channel with the 1-bit ADCs and comparator network system, can be expressed based on [7], [11] as

$$\mathbf{H'}_{\mathrm{R}} = \mathbf{C}_{\mathbf{z}_{\mathcal{Q}}\mathbf{z}_{\mathrm{R}}}(\mathbf{C}_{\mathbf{z}_{\mathrm{R}}})^{-1}\mathbf{B}\mathbf{H}_{\mathrm{R}}, \tag{4-5}$$

Based on Bussgang theorem the cross correlation similarly, the equation obtained in the previous section is calculated based on [25] express as $\mathbf{C}_{\mathbf{z}_{\mathcal{Q}}\mathbf{z}_{\mathrm{R}}} = \sqrt{\frac{2}{\pi}}\mathbf{K}_{\mathrm{R}}\mathbf{C}_{\mathbf{z}_{\mathrm{R}}}$, is the cross correlation between the quantized and unquantized versions of the received signal. Hence, the effective channel can be written as

$$\mathbf{H}'_{\mathrm{R}} = \sqrt{\frac{2}{\pi}}\mathbf{K}_{\mathrm{R}}\mathbf{B}\mathbf{H}_{\mathrm{R}} \in \mathbb{R}^{2N_r+\alpha\times 2N_t}. \tag{4-6}$$

Moreover, the effective noise covariance matrix based on [7], [11] can be

expressed as

$$\mathbf{C}_{\eta'_{\mathrm{R}}} = \frac{2}{\pi}\Big[\, \sin^{-1}(\mathbf{K}_{\mathrm{R}}\mathbf{C}_{\mathbf{z}_{\mathrm{R}}}\mathbf{K}_{\mathrm{R}}) - \mathbf{K}_{\mathrm{R}}\mathbf{C}_{\mathbf{z}_{\mathrm{R}}}\mathbf{K}_{\mathrm{R}} + \mathbf{K}_{\mathrm{R}}\mathbf{B}\mathbf{C}_{\mathbf{n}_{\mathrm{R}}}\mathbf{B}^{T}\mathbf{K}_{\mathrm{R}}\Big], \tag{4-7}$$

where the matrix $\mathbf{B}\mathbf{B}^{T}$ is approximately a scaled identity matrix $\in \mathbb{R}^{2N_r+\alpha\times 2N_r+\alpha}$.

Substituting (4-5) and (4-7) into (4-1), we have the sum rate for the system with 1-bit ADCs and a comparator network. To formulate the objective function of the convex optimization problem, an approximation of (4-7) is useful. First, the approximation of the first term can be obtained through the first-order Taylor series, and can be written as

$$\sin^{-1}(\mathbf{K}_{\mathrm{R}}\mathbf{C}_{\mathbf{z}_{\mathrm{R}}}\mathbf{K}_{\mathrm{R}}) \approx \mathbf{K}_{\mathrm{R}}\mathbf{C}_{\mathbf{z}_{\mathrm{R}}}\mathbf{K}_{\mathrm{R}} + \left(\frac{\pi}{2} - 1\right)\mathbf{I}. \tag{4-8}$$

Hence, (4-7) can be re-expressed as the effective noise covariance

$$\mathbf{C}_{\eta'_{R}} \approx \frac{2}{\pi}\Big[\left(\frac{\pi}{2} - 1\right)\mathbf{I} + \mathbf{K}_{\mathrm{R}}\mathbf{B}\mathbf{C}_{\mathbf{n}_{\mathrm{R}}}\mathbf{B}^{T}\mathbf{K}_{\mathrm{R}}\Big]. \tag{4-9}$$

To approximate $\mathbf{K}_{\mathrm{R}}$, first we rewrite $\mathbf{C}_{\mathbf{z}_{\mathrm{R}}}$, considering $\mathbf{C}_{\mathbf{x}_{\mathrm{R}}} = \frac{\sigma_x^2}{2}\mathbf{I}$ and $\mathbf{C}_{\mathbf{n}_{\mathrm{R}}} = \frac{\sigma_n^2}{2}\mathbf{I}$, both are considered in especial case that is mention after in Chapter 6 when is used the perfect channel state information. With this, we rewrite (4-2) as

$$\mathbf{C}_{\mathbf{z}_{\mathrm{R}}} = \mathbf{B}\left(\frac{\sigma_x^2}{2}\mathbf{H}_{\mathrm{R}}\mathbf{H}_{\mathrm{R}}^{T} + \frac{\sigma_n^2}{2}\mathbf{I}\right)\mathbf{B}^{T}. \tag{4-10}$$

In the following, we consider a Rayleigh fading channel with $\mathbb{E}\{\mathbf{H}_{\mathrm{R}}\mathbf{H}_{\mathrm{R}}^{T}\} = N_t\mathbf{I}$. With this, (4-10) can be expressed as

$$\mathbf{C}_{\mathbf{z}_{\mathrm{R}}} = \left(N_t\frac{\sigma_x^2}{2} + \frac{\sigma_n^2}{2}\right)\mathbf{B}\mathbf{B}^{T}. \tag{4-11}$$

Consequently, the approximated autocorrelation matrix of signal $\mathbf{z}_{\mathrm{R}}$ reads as

$$\mathbf{C}_{\mathbf{z}_{\mathrm{R}}} \approx \left(N_t\frac{\sigma_x^2}{2} + \frac{\sigma_n^2}{2}\right)\mathbf{I}. \tag{4-12}$$

With this, matrix $\mathbf{K}_{\mathrm{R}}$ can be approximated as

$$\mathbf{K}_{\mathrm{R}} \approx \left[\left(N_t\frac{\sigma_x^2}{2} + \frac{\sigma_n^2}{2}\right)\right]^{-\frac{1}{2}}\mathbf{I}. \tag{4-13}$$

Thus, we can present the identity-scaled approximation for the third term of (4-7) as follows

$$\mathbf{K}_{\mathrm{R}}\mathbf{B}\mathbf{C}_{\mathbf{n}_{\mathrm{R}}}\mathbf{B}^{T}\mathbf{K}_{\mathrm{R}} \approx \frac{\sigma_n^2}{2}\left[\left(N_t\frac{\sigma_x^2}{2}+\frac{\sigma_n^2}{2}\right)\right]^{-1}\mathbf{I}. \tag{4-14}$$

Substituting (4-14) in (4-9), the effective noise covariance matrix can be represented as

$$\mathbf{C}_{\eta'_R} \approx \frac{2}{\pi}\Big[\left(\frac{\pi}{2}-1\right)\mathbf{I}+\frac{\sigma_n^2}{2}\left(N_t\frac{\sigma_x^2}{2}+\frac{\sigma_n^2}{2}\right)^{-1}\mathbf{I}\Big]. \tag{4-15}$$

Considering that $\mathbf{C}_{\eta'_R} \approx \lambda\cdot\mathbf{I}$, the scaled identity approximation factor λ of the effective noise covariance (4-7), reads as

$$\lambda = \frac{2}{\pi}\left[\frac{\pi}{2}-1+\frac{\sigma_n^2}{2}\left(N_t\frac{\sigma_x^2}{2}+\frac{\sigma_n^2}{2}\right)^{-1}\right]. \tag{4-16}$$

Substituting (4-16) in (4-1), we get the sum rate approximation for 1-bit ADCs and comparator network systems as a function of a scaled identity matrix. This expression is used later as an objective function for the optimization problem and simplified version reads as

$$\tilde{I}(\mathbf{x};\mathbf{z}_{\mathcal{Q}}) = \frac{1}{2}\log_2\left(\left|\mathbf{I}_{2Nr+\alpha}+\frac{1}{\lambda}\frac{\sigma_x^2}{2}\mathbf{H'}_{\mathrm{R}}\mathbf{H'}_{\mathrm{R}}^{T}\right|\right). \tag{4-17}$$

The Fig. 4.1 show the approximation the effective noise covariance as an scaled identity matrix. The figure of the left side is the original effective noise covariance matrix, and the right side represents the approximation as scaled identity matrix. This approximation is useful for the optimization process. This approximation allows convex optimization to be applied (as described in Chapter 6 , the original problem is non convex and combinatorial). The performance is maintained close to the optimal, as verified in the simulations. The Fig. 4.2 shows the sum rate simulation comparison between quantized real and unquantized real. The graph shows that as SNR increases, the sum rate also increases, but the different techniques behave differently. Unquantized techniques (in both the complex and real equivalent representation) achieve higher sum rates than quantized techniques. Below is a description of the techniques used to obtain the curves, considering 1-bit of quantization and a network of comparators.

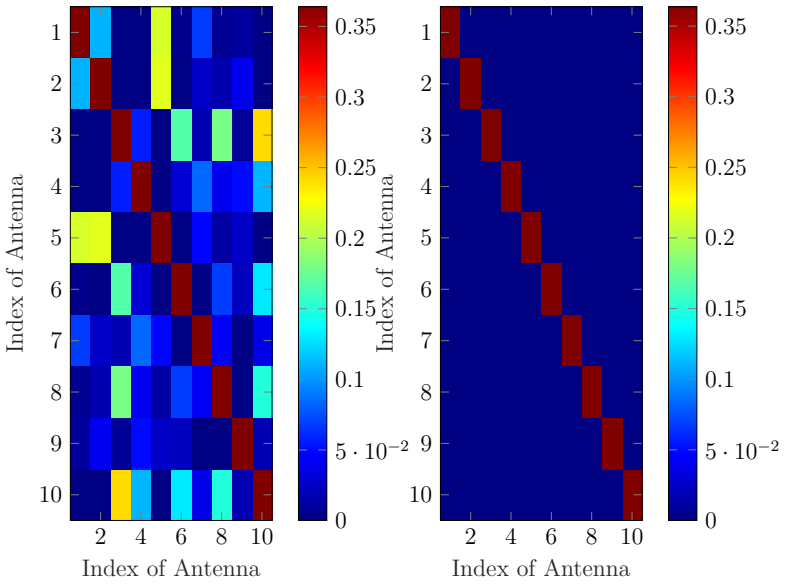


Figure 4.1: In the left side is shown the effective noise covariance $\mathbf{C}_{n'_R}$, in the right side show as scaled identity matrix

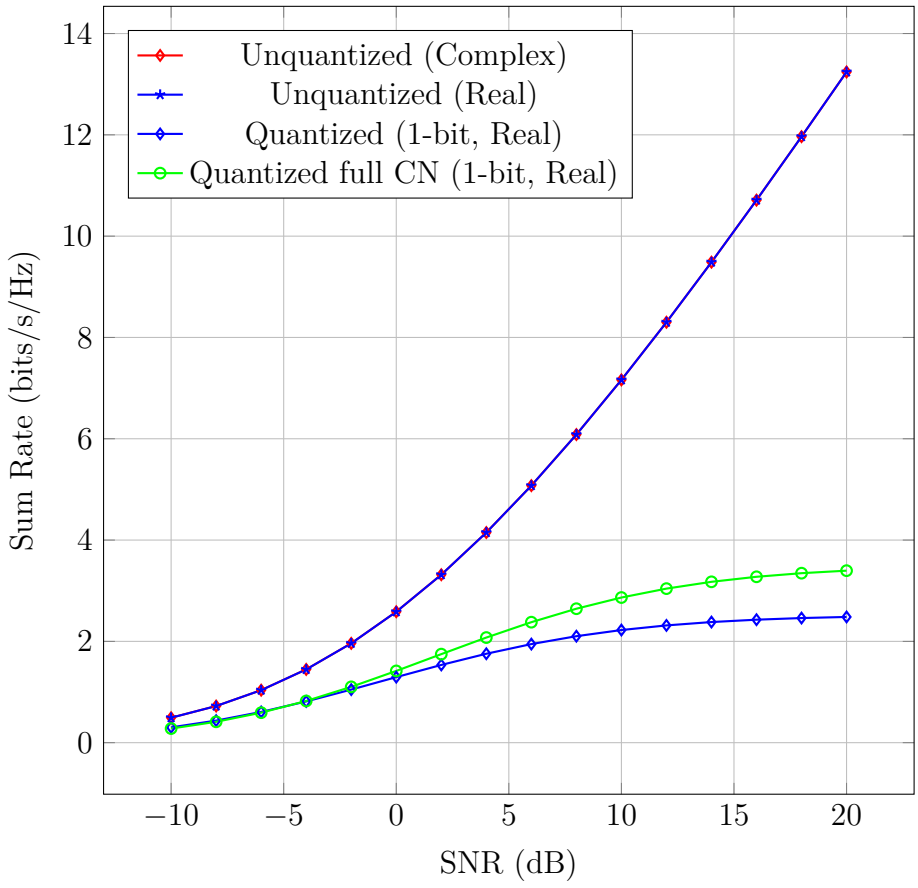


Figure 4.2: Comparison of Sum Rate lower bound $N_t = 2$ and $N_r = 2$

- **Unquantized Complex** Indicates the sum rate without quantization in the complex domain.
- **Unquantized Real** Represented the sum rate without quantization in the real equivalent value.
- **Quantized 1-bit real** Shows the blue rhombus wich represents the sum rate with 1-bit quantization in the real equivalent value.
- **Quantized full connected CN** Indicates the 1-bit quantization included the comparator network in the real equivalent value. In the figure is the green one and represent and improvement compared with only 1-bit quantization in real equivalent representation (blue rhombus curve).

The Fig. 4.3 compares the network capacity of a 1-bit quantized MIMO and comparator network system for two configurations considering the approximation of the sum rate expression, $3N_t$ number of transmitted antennas and $8N_r$ (the blue curve) and $3N_t$ and $16N_r$ (the red curve). At low SNR values ($\geq$ 5dB), both systems show very low capacity. As SNR increases, the capacity rises sharply, but the $3N_t$ $16N_r$ system consistently outperforms the $3N_t$, $8N_r$ system. At high SNR values($\geq$ 10dB), both curves saturate around 10–11 bits/s/Hz, with the $3N_t$ and $16N_r$ configuration always achieving a higher rate. In simple terms, having more receiving antennas $3N_t$ and $16N_r$ increases the transmission capacity in quantized MIMO systems, especially as signal quality improves.

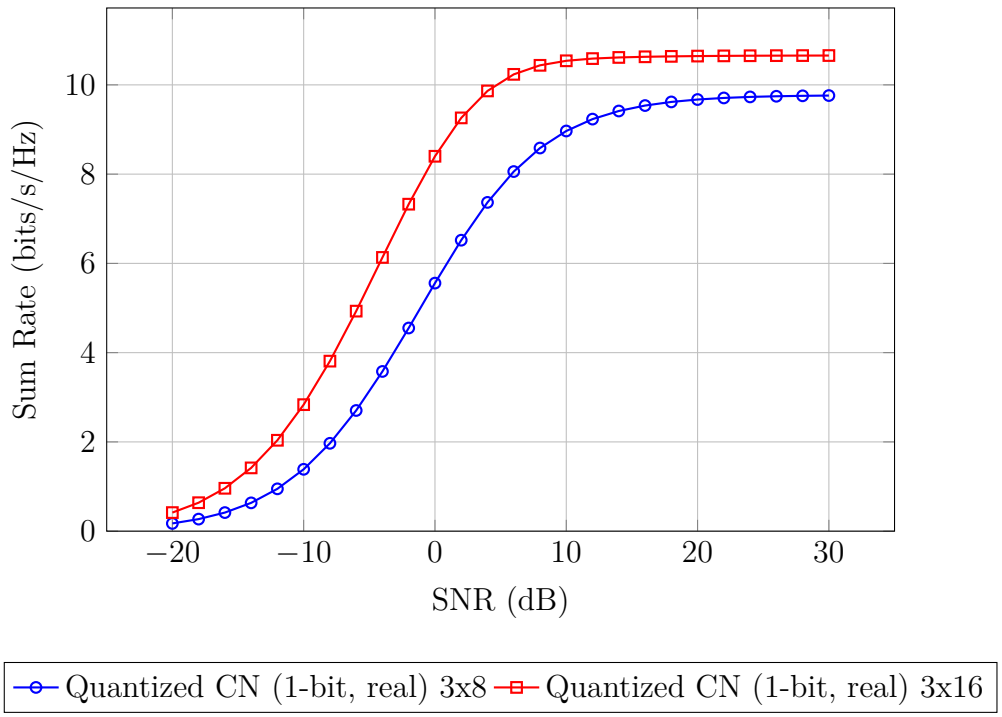


Figure 4.3: Comparison of Quantized Comparator Network Capacity for 3x8 and 3x16 MIMO

5
Linear Detection

The main goal and challenge of linear detection is to recover the transmitted signal from the received signal. For this purpose, both estimation and decision-making were taken into account. For this reason, an estimation method is used. The input signals $\mathbf{x}_{N_t}$ belong to a discrete set denoted as $\mathcal{A}$ [26]. At the output of the linear detection, $\tilde{\mathbf{x}}$ is received. The closest symbol maps $\tilde{\mathbf{x}}$ to the nearest symbol in the set $\mathcal{A}$, which consequently produces $\hat{\mathbf{x}}$ [26]. This approximation is a decision-making operation based on the proximity of the transmitted signal in the constellation. This is mapped using the Euclidean distance. The linear detector is also known as a filter, it reduces the effects of interference. For the estimated signal approach is given in the following approximation

$$\tilde{\mathbf{x}} \rightarrow \hat{\mathbf{x}}, \tag{5-1}$$

where $\tilde{\mathbf{x}}$ is the estimated signal after applying of linear filter, such as shown in the (5-2). $\hat{\mathbf{x}}$ is the detected signal, and is the result of applying a quantization process. The main objective is to obtain the transmitted signal $\tilde{\mathbf{x}} \in \mathbb{R}^{N_t \times 1}$. From equation (4-1), the estimated transmitted signal vector is obtained as follows

$$\tilde{\mathbf{x}} = \mathbf{W} \cdot \mathbf{y}, \tag{5-2}$$

where $\mathbf{W}$ acts like a filter.
Considering the equivalent real-value representation, we have

$$\tilde{\mathbf{x}}_R = \mathbf{W}_R \cdot \mathbf{y}_R. \tag{5-3}$$

The following equation defines the properties for the transmitted signal and noise vectors. They express the expectation of the cross product between the transmitted signal vector $\mathbf{x}$ and the noise vector $\mathbf{n}$.

$$\mathbb{E}[\mathbf{n} \cdot \mathbf{x}^H] = 0, \tag{5-4}$$

where the transmitted vector is denoted by $\mathbf{x}$, and has dimensions $\mathbb{C}^{N_t \times 1}$. This relationship indicates that the noise $\mathbf{n}$ is uncorrelated with the transmitted signal $\mathbf{x}$.

5.1 Infinite Resolution

Infinite resolution is a theoretical scenario in which an unlimited number of bits represent all possible values in the continuous domain, allowing the digital representation to perfectly match the original analog signal. Thus, no quantization error arises. In this context, infinite resolution can achieve a significantly lower BER, resulting in higher system performance.

Zero-Forcing Linear Detector

The Zero-Forcing (ZF) linear detector is based on the Moore-Penrose pseudo-inverse technique. Zero-Forcing detection aims to cancel interference from the different transmitted signals by using the inverse or pseudo-inverse of the channel [27]. Although it effectively reduces interference, it amplifies noise, especially when the channel matrix is ill-conditioned. Importantly, the ZF approach does not account for the effects of noise during detection, which can lead to suboptimal performance in practical scenarios. Among the linear detector schemes with lower computational complexity are ZF, Matched Filter (MF), and Minimum Mean Square Error (MMSE) [27]. The ZF filter $\mathbf{W}_{ZF}$ is represented based on [27] by the following equation

$$\mathbf{W}_{ZF} = \left(\mathbf{H}^H \mathbf{H}\right)^{-1} \mathbf{H}^H, \tag{4-6}$$

where the channel matrix is denoted by $\mathbf{H}$, and has dimensions $\mathbb{C}^{N_r \times N_t}$.

MMSE Linear Detector

Compared to ZF, the MMSE detector balances interference cancellation and noise amplification, avoiding excessive noise amplification [27]. The MMSE provides better performance in terms of BER, especially in low-SNR channels. The MMSE linear detector is expressed as

$$\mathbf{W}_{\text{MMSE}} = \arg\min_{\mathbf{W}} \mathbb{E}[\|\mathbf{x} - \tilde{\mathbf{x}}\|_2^2], \tag{5-5}$$

where $\mathbf{W}_{\text{MMSE}}$ is the MMSE filter. Similarly, $\mathbf{x}$ is the original signal vector with dimensions $\mathbb{C}^{N_t \times 1}$.

The expected value represents the squared error, which is defined as the mean of the squared Euclidean norm. The Euclidean norm measures the squared distance between a transmitted signal and its estimate. In other words, the squared error was computed for all realizations. The expected value is then used to calculate the mean squared error between the original transmitted signal and the estimated signal after filtering.

Expanding this expression, one obtains the following equation

$$\mathbb{E}[\|\mathbf{x}-\tilde{\mathbf{x}}\|_2^2] = \mathbb{E}\left[\|(\mathbf{I}_{\mathbf{N_t}}-\mathbf{WH})\mathbf{x}\|_2^2\right] + \mathbb{E}\left[\|\mathbf{Wn}\|_2^2\right]. \tag{5-6}$$

Differentiating with respect to $\mathbf{W}$ and setting it to zero gives the MMSE filter. Applying the matrix inversion lemma to the MMSE, the expression for the MMSE filter is given by

$$\mathbf{W}_{\text{MMSE}} = \left(\mathbf{C}_{\mathbf{x}}^{-1} + \mathbf{H}^H\mathbf{C}_{\mathbf{n}}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^H\mathbf{C}_{\mathbf{n}}^{-1}, \tag{5-7}$$

where $\mathbf{C_x}$ is the correlation matrix of the transmitted signal, $\mathbf{C_n}$ is the noise correlation matrix, and $\mathbf{H}$ is the channel matrix. Under low SNR conditions, the MMSE detector approaches the MF solution, maximizing the received signal relative to the noise. At a high SNR, the MMSE detector approaches the ZF filter solution, eliminating signal interference, as shown in the following equation

$$\lim_{\text{SNR}\to 0}\mathbf{W}_{\text{MMSE}} \approx \mathbf{C_x}\mathbf{H}^H\mathbf{C}_{\mathbf{n}}^{-1}. \tag{5-8}$$

When the SNR approaches infinity, the filter equivalent is obtained as follows

$$\lim_{\text{SNR}\to\infty}\mathbf{W}_{\text{MMSE}} \approx \left(\mathbf{H}^H\mathbf{C}_{\mathbf{n}}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^H\mathbf{C}_{\mathbf{n}}^{-1}. \tag{5-9}$$

A special case is considered when specific values for the signal and noise covariance matrices are assigned, as shown in the following equation

$$\mathbf{C_n} = \sigma_n^2\mathbf{I}_{N_r}, \tag{5-10}$$

where σ_n^2 is the noise variance, and N_r is the number of receive antennas. $\mathbf{I}_{N_r}$ is the identity matrix of dimension $N_r \times N_r$.

The signal covariance, considering the real value representation, is expressed as follows

$$\mathbf{C}_{\mathbf{x}_\text{R}} = \frac{\sigma_{\mathbf{x}}^2}{2}\mathbf{I}_{N_{tR}}, \tag{5-11}$$

where $\mathbf{C}_{\mathbf{x}_\text{R}}$ is a real value representation of the covariance matrix of the transmitted signal vector $\mathbf{x}$. Normalization facilitates the comparison of algorithms

without relying on power scales. The total transmission power is controlled by an explicit parameter $\sigma_{\mathbf{x}}^2$ rather than being hidden in the constellation. With the last consideration, the MMSE filter is characterized as follows

$$\mathbf{W}_{\text{MMSE}} = \left(\frac{N_t \sigma_{\mathbf{n}}^2}{E_{tx}} \mathbf{I}_{N_t} + \mathbf{H}^H \mathbf{H} \right)^{-1} \mathbf{H}^H. \tag{5-12}$$

The noise covariance matrix is denoted as $\mathbf{C_n} = \sigma_{\mathbf{n}}^2 \mathbf{I}_{N_R}$, with $\sigma_{\mathbf{n}}^2$ being the variance of the noise. The identity matrix for the receiver antennas is $\mathbf{I}_{N_R}$. In the MMSE framework, the noise covariance minimizes the expected squared error in estimating the transmitted signal. For high SNR values, the equations for Zero Forcing and Minimum Mean Squared Error are equivalent.

Matched Filter Detector

The Matched Filter is a linear detector that maximizes the SNR at its output and requires less computation. The Matched Filter does not account for interference effects. It is represented by the following equation:

$$\mathbf{W_{MF}} = \mathbf{H}^H \mathbf{y}, \tag{5-13}$$

where $\mathbf{W_{MF}}$ denotes the weight vector obtained through the Matched Filter technique, which maximizes the SNR by aligning the received signal with the transmitted signal. The following table 5.1 based on [26] presents a comparison and summary of basic techniques of detection

Table 5.1: Comparison of Linear Detection Techniques

Zero Forcing	**MMSE**	**Matched Filter**
Requires CSI	Requires CSI	Requires CSI
Minimize the interference	Minimize the MSE	Best performance at low SNR
Matrix Inversion	Matrix Inversion	No Matrix Inversion

5.2 Detection MMSE with 1-bit Quantization ADCs

The linear MMSE filter aims to minimize the Mean Squared Error between the transmitted signal $\mathbf{x}_{\text{R}}$ and the estimate $\tilde{\mathbf{x}}_{\text{R}}$, which is given by

$$\mathbf{W}_{\text{MMSE,1-bit}} = \arg\min_{\mathbf{W}} E\left[\|\mathbf{x} - \mathbf{W}\mathbf{z}\|_2^2 \right], \tag{5-14}$$

where $\mathbf{x}$ is the transmitted signal, and $\mathbf{z}$ is the quantized received signal. $\|\cdot\|_2^2$ represents the squared Euclidean norm, which measures the squared distance between the transmitted signal $\mathbf{x}$ and the filtered quantized signal $\mathbf{Wz}$.

The objective function can be expressed as

$$\mathbb{E}\left[\|\mathbf{x}-\mathbf{Wz}\|_2^2\right]=\mathbb{E}\left[\mathbf{x}^H\mathbf{x}\right]-\mathbb{E}\left[\mathbf{x}^H\mathbf{Wz}\right]-\mathbb{E}\left[\mathbf{z}^H\mathbf{W}^H\mathbf{x}\right]+\mathbb{E}\left[\mathbf{z}^H\mathbf{W}^H\mathbf{Wz}\right]. \tag{5-15}$$

Only the last two terms on the right-hand side depend on $\mathbf{W}^*$, which can be rewritten as

$$\mathbb{E}\left[\mathbf{z}^H\mathbf{G}^H\mathbf{x}\right]+\mathbb{E}\left[\mathbf{z}^H\mathbf{W}^H\mathbf{Wz}\right]=-\mathrm{tr}\left(\mathbb{E}\left[\mathbf{zx}^H\right]\mathbf{W}^H\right)+\mathrm{tr}\left(\mathbb{E}\left[\mathbf{zz}^H\right]\mathbf{W}^H\mathbf{W}\right). \tag{5-16}$$

Taking the derivative with respect to $\mathbf{W}^*$ and setting it to zero yields the optimal filter

$$\mathbf{W}_{MMSE}=\mathbf{C}_{\mathbf{x},\mathbf{z}}\mathbf{C}_{\mathbf{z}}^{-1}, \tag{5-17}$$

where $\mathbf{C}_{\mathbf{x},\mathbf{z}}$ is the cross-correlation matrix between the transmitted signal $\mathbf{x}$ and the covariance of quantized received signal $\mathbf{z}$. $\mathbf{C}_{\mathbf{z}}$ is the autocorrelation matrix of the received signal $\mathbf{z}$. The cross-correlation matrix $\mathbf{C}_{\mathbf{x},\mathbf{z}}$ can be expressed as

$$\mathbf{C}_{\mathbf{x},\mathbf{z}}=\mathbb{E}[\mathbf{x}\mathbf{z}^H], \tag{5-18}$$

where $\mathbf{x}$ is the transmitted signal. $\mathbf{z}^H$ is the Hermitian of the received signal $\mathbf{z}$. $\mathbf{C}_{\mathbf{x},\mathbf{z}}$ is the cross-correlation matrix between the transmitted signal $\mathbf{x}$ and the received signal $\mathbf{z}$. Take into account the [24] [16] Bussgang decomposition obtain the following equation

$$\mathbf{C}_{\mathbf{x},\mathbf{z}}=\mathbb{E}[\mathbf{x}(\mathbf{x}^H\mathbf{H}^H\mathbf{A}^H+\mathbf{n}^H\mathbf{A}^H+\mathbf{n}_q^H)], \tag{5-19}$$

$\mathbf{n}_q$ is the quantization linear approach noise vector.
Consequently, the cross-correlation is expressed as follows

$$\mathbf{C}_{\mathbf{x},\mathbf{z}}=\mathbf{C}_{\mathbf{x}}\mathbf{H}^H\mathbf{A}^H. \tag{5-20}$$

And the cross-correlation between the input $\mathbf{y}$ and the quantized output based on [11] $\mathbf{z}$ is given by

$$\mathbf{C}_{\mathbf{z},\mathbf{y}} = \sqrt{\frac{2}{\pi}}\mathbf{C}_{\mathbf{z}}\,\mathrm{diag}(\mathbf{C}_{\mathbf{z}})^{-\frac{1}{2}}. \tag{5-21}$$

From the two expressions for the cross-correlation function obtained

$$\sqrt{\frac{2}{\pi}}\mathbf{C}_{\mathbf{z}}\,\mathrm{diag}(\mathbf{C}_{\mathbf{z}})^{-\frac{1}{2}} = \mathbf{C}_{\mathbf{z}}\mathbf{A}^{H}. \tag{5-22}$$

Finally, the following value is obtained for $\mathbf{A}$

$$\mathbf{A} = \sqrt{\frac{2}{\pi}}\,\mathrm{diag}(\mathbf{C}_{\mathbf{z}})^{-\frac{1}{2}}, \tag{5-23}$$

Where $\mathbf{A}$ represents a scaling matrix that is applied to the autocorrelation matrix of the quantized received signal$\mathbf{C}_{\mathbf{z}}$ is computing as [25].

$$\mathbf{C}_{\mathbf{z}} = \frac{2}{\pi}\Bigg(\sin^{-1}\left(\mathrm{diag}(\mathbf{C}_{\mathbf{y}})^{-\frac{1}{2}}\,\mathrm{Re}(\mathbf{C}_{\mathbf{y}})\,\mathrm{diag}(\mathbf{C}_{\mathbf{y}})^{-\frac{1}{2}}\right) \tag{5-24}$$

$$+ j\cdot\sin^{-1}\left(\mathrm{diag}(\mathbf{C}_{\mathbf{y}})^{-\frac{1}{2}}\,\mathrm{Im}(\mathbf{C}_{\mathbf{y}})\,\mathrm{diag}(\mathbf{C}_{\mathbf{y}})^{-\frac{1}{2}}\right)\Bigg), \tag{5-25}$$

This matrix describes how each element of the quantized signal is related to the others.

Then, the covariance matrix of the unquantized signal is given

$$\mathbf{C}_{\mathbf{y}} = \mathbb{E}[\mathbf{y}\mathbf{y}^{H}] = \mathbf{H}\mathbf{C}_{\mathbf{x}}\mathbf{H}^{H} + \mathbf{C}_{\mathbf{n}}. \tag{5-26}$$

Thus, the MMSE detection at the receiver expressed in function of $\mathbf{H}, \mathbf{C}_{\mathbf{x}}$ and

$\mathbf{C_n}$, becomes as follows

$$\begin{aligned}\mathbf{W}_{\text{MMSE, 1-bit}} = \sqrt{\frac{\pi}{2}}\,\mathbf{C_x}\mathbf{H}^H \operatorname{diag}\left(\mathbf{H}\mathbf{C_x}\mathbf{H}^H + \mathbf{C_n}\right)^{-\frac{1}{2}} \\ \times \Bigg[\Bigg(\sin^{-1}\Big(\operatorname{diag}\left(\mathbf{H}\mathbf{C_x}\mathbf{H}^H + \mathbf{C_n}\right)^{-\frac{1}{2}} \\ \times \operatorname{Re}\left(\mathbf{H}\mathbf{C_x}\mathbf{H}^H + \mathbf{C_n}\right)\operatorname{diag}\left(\mathbf{H}\mathbf{C_x}\mathbf{H}^H + \mathbf{C_n}\right)^{-\frac{1}{2}}\Big) \\ + j\cdot\sin^{-1}\Big(\operatorname{diag}\left(\mathbf{H}\mathbf{C_x}\mathbf{H}^H + \mathbf{C_n}\right)^{-\frac{1}{2}} \\ \times \operatorname{Im}\left(\mathbf{H}\mathbf{C_x}\mathbf{H}^H + \mathbf{C_n}\right)\operatorname{diag}\left(\mathbf{H}\mathbf{C_x}\mathbf{H}^H + \mathbf{C_n}\right)^{-\frac{1}{2}}\Big)\Bigg)^{-1}\Bigg].\end{aligned} \tag{5-27}$$

We can make the following consideration

$$\mathbf{K}_{\mathrm{R}} = \operatorname{diag}(\mathbf{C}_{\mathbf{z}_{\mathrm{R}}})^{-\frac{1}{2}}, \tag{5-28}$$

where $\mathbf{K}_{\mathrm{R}} \in \mathbb{R}^{(2N_r+\alpha)\times(2N_r+\alpha)}$ related to the covariance. The equation (5-27) can be written with the real equivalent representation. With this, MMSE receiver matrix with 1-bit quantization at the receiver for equivalent real value representation becomes

$$\mathbf{W}_{\text{MMSE,1-bit}_{\mathrm{R}}} = \sqrt{\frac{\pi}{2}}\,\mathbf{C}_{\mathbf{x}_{\mathrm{R}}}\mathbf{H}^T\mathbf{K}_{\mathrm{R}}\left(\sin^{-1}\left(\mathbf{K}_{\mathrm{R}}\operatorname{Re}\{\mathbf{C}_{\mathbf{z}_{\mathrm{R}}}\}\mathbf{K}_{\mathrm{R}}\right)\right). \tag{5-29}$$

5.3 Detection Low Resolution Aware Minimum Mean Square Error (LRA-MMSE) with 1-bit ADCs and Comparator Networks

The LRA-MMSE with 1-bit ADCs and Comparator Network approach allows for modeling the non-linearity of 1-bit quantization. Based on the Bussgang theorem, this approach helps preserve the correlation of the original signal despite the non-linearity [26]. The Bussgang theorem enables a model of the autocorrelation between the quantized signal $\mathbf{z}_R$ as a function of the unquantized signal $\mathbf{y}_R$, simplifying the analysis and design of the LRA-MMSE detector.

The channel capacity decreases with 1-bit quantization owing to the low resolution at the receiver [11]. Virtual channels were introduced through a comparator network to compensate for this limitation. Despite binary quantization, these virtual channels represent a novel method for extracting useful information from the received signals. These virtual channels enable the

MIMO system to leverage its spatial diversity

The received signal is $\mathbf{y_R}$ and is quantized by 1-bit ADCs. The matrix $\mathbf{B}$ represents the comparator network and quantized differences, such as +1 or -1, with a threshold. The received signal is sent to the 1-bit ADCs and the comparator network. The main objective in the detection is to recover the original transmitted signal, for these reasons the proposed strategy detection based on [12] was considered. The process for defining the filter is presented in the following equation

$$\mathbf{W}_{\text{R,LRA-MMSE}} = \arg\min_{\mathbf{W}_\text{R}} \ \mathbb{E}\left[\left|\left|\mathbf{x}_\text{R} - \mathbf{W}_\text{R}^T \cdot \mathbf{z}_{\mathcal{Q}}^\text{R}\right|\right|_2^2\right], \tag{5-30}$$

where $\mathbf{z}_{\mathcal{Q}}^\text{R}$ is the quantized signal after passing through the ADCs and comparator based on Bussgang theorem, and $\in \mathbb{R}^{(2N_r+\alpha)\times(2N_r+\alpha)}$. The solution to the formulated problem considered the same steps of the previous section for linear detection (5-14) is

$$\mathbf{W}_{\text{R,LRA-MMSE}} = \mathbf{C}_{\mathbf{z}_{\mathcal{Q}}^\text{R}}^{-1}\mathbf{C}_{\mathbf{z}_{\mathcal{Q}}^\text{R}\mathbf{x}_\text{R}}, \tag{5-31}$$

where $\mathbf{C}_{\mathbf{z}_{\mathcal{Q}}^\text{R}}$ represents the covariance matrix of the quantization at the output of the ADCs and the comparator network. $\mathbf{C}_{\mathbf{z}_{\mathcal{Q}}^\text{R}\mathbf{x}_\text{R}}$ is the covariance matrix describing the relationship between the quantized signal vector $\mathbf{z}_{\mathcal{Q}}^\text{R}$ and the transmitted signal $\mathbf{x}$. This matrix configuration involves computations based on Bussgang theorem [24] such as

$$\mathbf{C}_{\mathbf{z}_{\mathcal{Q}}^\text{R}} = \frac{2}{\pi}\sin^{-1}\left(\mathbf{K}_\text{R}\,\text{Re}\left\{\mathbf{C}_{\mathbf{z}_\text{R}}\right\}\mathbf{K}_\text{R}\right), \tag{5-32}$$

where $\mathbf{C}_{\mathbf{z}_\text{R}}$ is an autocorrelation matrix and matrix $\mathbf{K}_\text{R}$ with dimensions $\mathbf{K}_\text{R} \in \mathbb{R}^{(2N_r+\alpha)\times(2N_r+\alpha)}$ is given by

$$\mathbf{K}_\text{R} = \text{diag}(\mathbf{C}_{\mathbf{z}_\text{R}})^{-\frac{1}{2}}. \tag{5-33}$$

The covariance matrix of received unquantized signal (the derivation is shown in the appendice A1) can be expressed as

$$\mathbf{C}_{\mathbf{z}_\text{R}} = \mathbf{B}\mathbf{H}_\text{R}\mathbf{C}_{\mathbf{x}_R}(\mathbf{H}_\text{R})^T\mathbf{B}^T + \mathbf{B}\mathbf{C}_{\mathbf{n}_\text{R}}\mathbf{B}^T, \tag{5-34}$$

where the covariance of the noise is real-valued. It is represented in the

following equation:

$$\mathbf{C}_{\mathbf{n}_\mathrm{R}} = \mathbb{E}[\mathbf{n}_\mathrm{R}\mathbf{n}_\mathrm{R}^T], \tag{5-35}$$

where $\mathbf{C}_{\mathbf{n}_\mathrm{R}}$ is the real-valued covariance matrix of the noise, and $\mathbf{n}_\mathrm{R}$ represents the real-valued noise vector.

The following equation (5-36) is a simplified and rearranged version of the last equation, where common terms are grouped and the filter is expressed by the following equation

$$\begin{aligned}\mathbf{W}_{\mathrm{R,LRA\text{-}MMSE}} = \sqrt{\frac{\pi}{2}}\Big[\sin^{-1}\Big(\mathbf{K}_\mathrm{R}\,\mathrm{Re}\Big\{\mathbf{B}\mathbf{H}_R^T\mathbf{C}_{\mathbf{x}_R}\mathbf{H}_R\mathbf{B}^T \\ + \mathbf{B}\mathbf{C}_{\mathbf{n}_R}\mathbf{B}^T\Big\}\mathbf{K}_R\Big)\Big]^{-1}\mathbf{K}_\mathbf{R}\mathbf{B}\mathbf{H}_\mathbf{x}\mathbf{C}_{\mathbf{x_R}}.\end{aligned} \tag{5-36}$$

The LRA-MMSE matrix with the comparator network is computed by derivation considered the same steps of the previous section for linear detection (5-14) is as follows.

$$\mathbf{W}_{\mathrm{R,LRA,full}} = \mathbf{C}^{-1}_{\mathbf{z}^\mathrm{R}_\mathcal{Q},\mathrm{full}}\mathbf{C}_{\mathbf{z}^\mathrm{R}_\mathcal{Q}\mathbf{x}_\mathrm{R},\mathrm{full}}, \tag{5-37}$$

where $\mathbf{C}_{\mathbf{z}^\mathrm{R}_\mathcal{Q}\mathbf{x}_\mathrm{R},\mathrm{full}}$ is the covariance matrix that describes the relationship between the quantized signal vector and the transmitted signal $\mathbf{x}_\mathrm{R}$.

Taking into the account the (5-36). For the fully connected comparator network the receiver matrix can be expressed as follows

$$\begin{aligned}\mathbf{W}_{\mathrm{R,LRA\text{-}MMSE}_\mathrm{full}} = \sqrt{\frac{\pi}{2}}\mathbf{C}_{\mathbf{x}_R}\mathbf{H}_\mathbf{x}^T\mathbf{B}^T\mathbf{K}_{\mathrm{full}\,R}\Big[\sin^{-1}\Big(\mathbf{K}_{\mathrm{full}\,R}\,\mathrm{Re}\Big\{\mathbf{B}_\mathrm{full}\mathbf{H}_R^T\mathbf{C}_{\mathbf{x}_R}\mathbf{H}_R\mathbf{B}_\mathrm{full}^T \\ + \mathbf{B}_\mathrm{full}\mathbf{C}_{\mathbf{n}_R}\mathbf{B}_\mathrm{full}^T\Big\}\mathbf{K}_{\mathrm{full}\,R}\Big)\Big]^{-1}.\end{aligned} \tag{5-38}$$

The equivalent expression for random configuration due to derivation is given by

$$\mathbf{W}_{\mathrm{R,LRA,random}} = \mathbf{C}^{-1}_{\mathbf{z}^\mathrm{R}_\mathcal{Q},\mathrm{random}}\mathbf{C}_{\mathbf{z}^\mathrm{R}_\mathcal{Q}\mathbf{x}_\mathrm{R},\mathrm{random}}, \tag{5-39}$$

where $\mathbf{C}_{\mathbf{z}^\mathrm{R}_\mathcal{Q}\mathbf{x}_\mathrm{R},\mathrm{random}}$ is the covariance matrix describing the relationship between the quantized $\mathbf{z}^\mathrm{R}_\mathcal{Q}$ and the representation of real equivalent representation of transmitted signal $\mathbf{x}_\mathrm{R}$. The following equation (5-40) presents the general compact representation of the filter and the estimated signal

$$\mathbf{W}_\mathrm{R} = \mathbf{C}^{-1}_{\mathbf{z}^\mathcal{Q}_R}\mathbf{C}_{\mathbf{z}^\mathcal{Q}_R\mathbf{x}_R}, \qquad \tilde{\mathbf{x}}_\mathrm{R} = \mathbf{W}_\mathrm{R}^T\mathbf{z}^\mathcal{Q}_\mathrm{R} \tag{5-40}$$

The representation of the random connected comparator network expressed by the following equation

$$\mathbf{W}_{\mathrm{R,LRA\text{-}MMSE}_{\mathrm{random}}} = \left(\frac{2}{\pi}\sin^{-1}\left(\mathbf{K}_{\mathrm{random}_{\mathrm{R}}}\,\mathrm{Re}\left\{\mathbf{B}\mathbf{H}_{\mathrm{R}}\mathbf{C}_{\mathbf{x}_{\mathrm{R}}}\mathbf{H}_{R}^{T}\mathbf{B}^{T}+\mathbf{B}\mathbf{C}_{\mathbf{n}_{\mathrm{R}}}\mathbf{B}^{T}\right\}\right)\right)^{-1} \times\left(\sqrt{\frac{2}{\pi}}\mathbf{K}_{\mathrm{random}_{\mathrm{R}}}\mathbf{B}\mathbf{H}_{\mathrm{R}}\mathbf{C}_{\mathbf{x}_{\mathrm{R}}}\right). \tag{5-41}$$

5.4 Computational complexity of typical Linear Detector

This section presents the computational cost associated with typical linear detectors, the Zero Forcing (ZF) and Minimum Mean Square Error (MMSE) detection techniques. In both cases, the main computational load arises from inverting the $N_t \times N_t$ matrix. Despite this, both algorithms have the same order of complexity, predominated by $\mathcal{O}(N_t^3)$ for inversion and $\mathcal{O}(N_r N_t^2)$ for matrix multiplications. The computational complexity for ZF and MMSE is equal to $\mathcal{O}(N_t^3 + N_r N_t^2)$ [28]. The following Table 5.2 shows a summary

Table 5.2: Computational complexity for a typical linear detector

Linear Detector	**Computational Complexity**
Zero Forcing	$\mathcal{O}(N_t^3 + N_t^2 \cdot N_r)$
Matched Filter	$\mathcal{O}(N_t \cdot N_r)$
MMSE	$\mathcal{O}(N_t^3 + N_t^2 \cdot N_r)$

comparison of the typical linear detection strategies. The Matched Filter has less computational cost. In Fig. 7.7 was simulated the BER performance of the basic strategies of detection.

The ZF detector is computationally expensive because it cancels interference effectively, while the MMSE detector provides the best trade-off between complexity and performance. The computational complexity values shown in Table 5.3 for $N_t = 2$ and $N_r = 2$ reflect the number of basic mathematical operations required by each linear detection algorithm. The (ZF) detector uses the expression

$$\hat{\mathbf{x}}_{\mathrm{ZF}} = \left(\mathbf{H}^{H}\mathbf{H}\right)^{-1}\mathbf{H}^{H}\mathbf{y}. \tag{5-42}$$

The calculation involves the Gram matrix $\mathbf{H}^H\mathbf{H}$, which involves eight multiplications for matrices of size $N_r = 2 \times N_t = 2$. The inversion process itself consists of four operations for a matrix of the same size and includes two additional matrix multiplications, each comprising four operations. This result

excluding the inversion in a complexity of $\mathcal{O}(16)$ [28]. Linear detection (MF) involves a simple matrix-vector multiplication

$$\hat{\mathbf{x}}_{\mathrm{MF}} = \mathbf{H}^H \mathbf{y}, \tag{5-43}$$

where $\mathbf{H}^H$ is the Hermitian transpose of the channel matrix $\mathbf{H}$, and $\mathbf{y}$ is the received-signal vector. In the process, there is a multiplication between a matrix $N_t = 2 \times N_r = 2$ and a reception vector 2×1 $\mathbf{y}$. Finally, we obtain that four multiplications are required in total, with a computational complexity of $\mathcal{O}(4)$ [20].

These complexity estimates are useful approximations based on the number of multiplications involved and are particularly relevant when evaluating the performance and implementation cost of linear detectors in small MIMO systems. Computing the computational cost of detectors with 1-bit quantization as shown in Table 5.3 [3] .

In the table 5.3 [3] shows the computational cost of the LRA MMSE detector for all configurations of the comparator network. Moreover, includes the other strategies of optimization as SINR and Greedy search MMSE. In the next section was made the comparison in term of BER and sum rate. When using representation equivalent real values, representation, including

Table 5.3: Computational complexity with LRA-MMSE [3]

Linear Detector	**Computational Cost LRA-MMSE detector**
Without network	$\mathcal{O}((2 \cdot N_r)^3 + 2 \cdot N_t \cdot (2 \cdot N_r)^2 + 4 \cdot N_r \cdot N_t)$
Random selection	$\mathcal{O}((2 \cdot N_r + \alpha_p)^3 + 2 \cdot N_t \cdot (2 \cdot N_r + \alpha_p)^2 + 2N_t \cdot (2N_r + \alpha_p))$
Fully Selection	$\mathcal{O}((2N_r + \alpha_f)^3 + 2 \cdot N_t \cdot (2 \cdot N_r + \alpha_f)^2 + 2N_t \cdot (2N_r + \alpha_f))$
SINR	$\mathcal{O}((2N_r + \alpha_p)^3 + 2 \cdot N_t \cdot (2 \cdot N_r + \alpha_p)^2 + 2N_t \cdot (2N_r + \alpha_p))$
Greedy search	$\mathcal{O}((2N_r + \alpha_p)^3 + 2 \cdot N_t \cdot (2 \cdot N_r + \alpha_p)^2 + 2N_t \cdot (2N_r + \alpha_p))$

comparator networks is represented in Table 5.3. The Table 5.3 shows the numerical values for computational cost and hardware cost considering LRA-MMSE Detection [3].

The next Table 5.4 [3] shows the compromise between computational cost and hardware cost is computed as $\alpha_{\mathrm{full}} = N_r(2N_r - 1)$ for the fully selection when selecting an optimization strategy in the context of MIMO with LRA-MMSE detection with a comparator network [11]. The Greedy MMSE search and the SINR search optimization strategies will be presented in the next chapter for comparison with the proposed algorithm with convex optimization.

Table 5.4: Computational cost with LRA-MMSE hardware cost $N_t = 2$ $N_r = 10$ [3]

Linear Detector	**Computational Cost**	**Hardware Cost**
Without network	$\mathcal{O}(9.68 \cdot 10^3)$	–
Random selection	$\mathcal{O}(7.05 \cdot 10^4)$	20
Fully Selection	$\mathcal{O}(9.44 \cdot 10^6)$	190
Greedy Search	$\mathcal{O}(2.18 \cdot 10^8)$	20
SINR	$\mathcal{O}(7.06 \cdot 10^5)$	20

6
Comparator Selection Optimization Algorithms

This section discusses three optimization strategies to approximate the full comparator network with a reduced number of comparators. The random comparator network α_{partial} selects some rows from α_{fully}. The main objective in this section is reduces the size of $\mathbf{B}$. The random comparator network is a combinatorial problem and exists in different combinations, where $\alpha_{\text{fully}} = \binom{N_r \cdot (2 \cdot N_r - 1)}{\alpha_{\text{partial}}}$. For this reason, three optimization techniques are used to improve performance with fewer comparators the Greedy search algorithm, the SINR, and the proposed convex optimization strategy for comparator selection.

6.1
MMSE Greedy Search Network

The Greedy search algorithm is an optimized strategy based on an iterative process. The main objective is to choose the option with the minimum value of MMSE. According to [3], this algorithm uses a greedy search to find the best subset of rows α_{partial} for the matrix $\mathbf{B}_{\alpha_{\text{partial}}}$ from the full matrix $\mathbf{B}$, with the aim of minimizing the MSE. The process starts by calculating an initial $l_{\min}$. The algorithm then compares each row (i) in the current subset with the candidate rows (j) from $\mathbf{B}'$. If swapping row i with row j leads to a new MSE (l) that is less than $l_{\min}$, the swap is made and $l_{\min}$ is updated. This step-by-step process continues until no further reduction in the MMSE is possible, resulting in an optimized matrix $\mathbf{B}$. This approach is more efficient than random comparator network selection because it ensures the largest possible reduction in the MSE at each step. The following outlines the steps in the algorithm for optimizing the MMSE

Algorithm 1 MMSE based Greedy Search [3]

Find the fully connected network $\mathbf{B}'$ in (2-8) and get the number of rows, defined by α_f ▷ Define all possible comparators
Extract the first α_{partial} rows of $\mathbf{B}'$, defined by $\mathbf{B}'_{\alpha_{\text{partial}}}$ ▷ Considering the initial selection of α_p comparators
Constitute $\mathbf{B}$ in (2-7) and calculate $\mathbf{W}_{\text{R}}$ in (5-30) ▷ The initial system model and MMSE receiver were set up.
Compute the MSE with $E\left[\|\mathbf{x}_{\text{R}} - \mathbf{W}_{\text{R}}\mathbf{z}_{\mathcal{Q}}\|_2^2\right]$, defined by $l_{\min}$ ▷ Initial minimum MSE
for $i = 1 : \alpha_{\text{partial}}$ **do** ▷ Iterate over each row currently in the selected set $\mathbf{B}'_{\alpha_{\text{partial}}}$
    Take the ith row of $\mathbf{B}'_{\alpha_{\text{partial}}}$ and freeze the other $\alpha_{\text{partial}} - 1$ rows
    for $j = 1 : \alpha_f$ **do** ▷ Each row of possible options from the entire set is examined $\mathbf{B}'$
        if the jth row of $\mathbf{B}'$ is already in $\mathbf{B}'_{\alpha_{\text{partial}}}$ **then**
            $j = j + 1$ ▷ Jump to the next row if the row is already selected
        else
            Replace the ith row of $\mathbf{B}'_{\alpha_{\text{partial}}}$ with the jth row of $\mathbf{B}'$
            Constitute $\mathbf{B}$ and calculate $\mathbf{W}_{\text{R}}$ ▷ Recalculate system parameters
            Compute the MSE value, defined by l ▷ Calculate MSE after the temporary exchange
            if $l < l_{\min}$ **then** ▷ Verify if the exchange improves performance
                $l_{\min} = l$ ▷ Update the minimum MSE
                Update $\mathbf{B}'_{\alpha_{\text{partial}}}$ ▷ Evaluate each candidate row for replacement within the entire set $\mathbf{B}'$
            end if
        end if
    end for
end for

The Greedy search algorithm computes the local MMSE for each row and the global MMSE after obtaining the optimized matrix to ensure the optimal matrix.

The MSE according to [3] for a system involving quantized observations is expressed in the following equation

$$\text{MSE} = \text{tr}\left(\mathbf{W}_{\text{R}}^{T}\mathbf{C}_{\mathbf{z}_{\text{Q}}^{R}}\mathbf{W}_{\text{R}}\right) - 2\,\text{tr}\left(\mathbf{C}_{\mathbf{z}_{\text{Q}}^{R}\mathbf{x}_{\text{R}}}\mathbf{W}_{\text{R}}^{T}\right) + \text{tr}(\mathbf{C}_{\mathbf{x}_{\text{R}}}), \tag{6-1}$$

where $\text{tr}(\mathbf{C}_{\mathbf{x}_R})$ is the constant term.

6.2 Sequential SINR Algorithm Search

The Signal-to-Interference-plus-Noise Ratio (SINR) reflects the relationship between the signal transmission power and interference noise for a given signal flow. According to [11] the main objective is performance-maximizing

Algorithm 2 Sequential Signal Interference Noise to Ratio (SINR) [11]

1: Define the matrix $\mathbf{B}'$ considered fully connected and compute the number of rows α_{full}.
2: Define the set of unused comparators $L = \{1, \ldots, \alpha_{\text{full}}\}$.
3: Compute $\text{SINR}_{l,k}$ for all l (row index in $\mathbf{B}'$) and k (signal stream) using equation (6-2).
4: Initialize the comparator network $\mathbf{B} = \mathbf{I}_{2N_r}$. ▷ Start without the added comparator network
5: **for** $i = 1 : \alpha_{\text{partial}}$ **do** ▷ Iterate until α_{partial} comparators are selected
6: Compute the receiver weighting matrix $\mathbf{W}_{\text{R}}$ using equation (5-7) based on $\mathbf{B}$.
7: Compute the MSE_k for the $2N_t$ signal streams using equation (6-3).
8: Identify the weakest signal index: $k_{\max} = \arg\max_k \text{MSE}_k$.
9: Identify the unused comparator with maximum SINR $k_{\max}$:
10: $l_{\max} = \arg\max_{l \in L} \text{SINR}_{k_{\max},l}$. ▷ Selection based on weakest stream
11: Extend the comparator network matrix with the $l_{\max}$-th row of $\mathbf{B}'$:
12: $\mathbf{B} = \begin{bmatrix} \mathbf{B} \\ \mathbf{B}'_{l_{\max}} \end{bmatrix}$.
13: Update the set $L = \{1, \ldots, l_{\max} - 1, l_{\max} + 1, \ldots, \alpha_{\text{full}}\}$.
14: **end for**

comparators were selected as the worst-performing comparators with a high MSE. Initially, $\mathbf{B}'$ with was considered. The next step is defined as L, which contains the indices of all available comparators in $\mathbf{B}'$. The set of comparators is represented as $L = \{1, \ldots, \alpha_{\text{full}}\}$. From this set, one comparator is selected in $\mathbf{B}'_{\alpha_{\text{partial}}}$. This decreases as the comparators are selected for the network. The SINR of each comparator was evaluated with respect to each signal flow k.

The SINR for a specific signal k-th real-valued channel input before quantization at the output of the comparator l. The numerator of the SINR corresponds to the desired signal contribution for the weakest stream, represented as $k_{\max}$. In the denominator, the expression is the sum of all interferences plus noise contributions. The SINR expression is shown as [3] follows

$$\mathrm{SINR}_{l,k} = \frac{\frac{1}{2}\sigma_x^2 |[\mathbf{B}'\mathbf{H}_{\mathrm{R}}]_{l,k}|^2}{\frac{1}{2}\sigma_x^2 \sum_{j=1, j\neq k}^{2N_t} |[\mathbf{B}'\mathbf{H}_{\mathrm{R}}]_{l,j}|^2 + \sigma_n^2}. \tag{6-2}$$

This indicates that the comparator improves the signal of interest against interference and noise. The system is initialized without a comparator network, and only the actual received signals are considered without additional operations. The process is then repeated from $i = 1$ to $i = \alpha_{\text{full}}$. The next step is to calculate W as a function of B, including B'. The next step is to calculate MSEk for N_t. The stream with the highest MSE ($k_{\max}$) is selected for improvement. Subsequently, the one that maximizes the SINR is selected. This selection helps improve the BER. The higher the SINR value, the higher the data transmission rate [3].

The MSE based on [3] the $k-th$ real-valued after passing through the filter $\mathbf{W}_{\mathrm{R}}$ is given by

$$\mathrm{MSE}_k = \left[\mathbf{W}_{\mathrm{R}}^T \mathbf{C}_{\mathbf{z}_Q^R} \mathbf{W}_R - \mathbf{W}_{\mathrm{R}}^T \mathbf{C}_{\mathbf{z}_{\mathbf{Q}}^R \mathbf{x}_{\mathrm{R}}} + \mathbf{C}_{\mathbf{x}_R}\right]_{k,k}. \tag{6-3}$$

This is a sequential process. In each search cycle, the comparator configuration with the highest SINR is selected [11]. The following are the steps to be followed in the SINR Search Algorithm [3]

6.3 Proposed Convex Optimization Strategy

This section presents the proposed convex optimization strategy based on [14]. In the first stage, the binary decision vector Δ is defined as $\Delta = [\Delta_1, \Delta_2, \ldots, \Delta_{2N_r+\alpha_{\text{full}}}]$, where

$$\Delta_i = \begin{cases} 1, & \text{if the } i^{\text{th}} \text{ row comparator is selected,} \\ 0, & \text{otherwise.} \end{cases} \tag{6-4}$$

Where index i represents the row number of the selected comparator in the network. From the delta vector Δ, diag(Δ) is constructed, which can be

represented, based on [14], as

$$\mathrm{diag}(\Delta) = \begin{bmatrix} \Delta_1 & 0 & \cdots & 0 \\ 0 & \Delta_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Delta_{2N_r+\alpha_{\mathrm{full}}} \end{bmatrix}. \tag{6-5}$$

Thus, the channel matrix can be modified considering the diagonal matrix Δ. The sum rate expression developed in (4-17) reads as

$$\tilde{I}(\Delta) = \tfrac{1}{2}\log\left|\mathbf{I}_{2Nr+\alpha_{\mathrm{full}}} + \frac{\sigma_x^2}{2\lambda}\mathrm{diag}(\Delta)\mathbf{H}'_{\mathrm{R}}\mathbf{H}'^{T}_{\mathrm{R}}\mathrm{diag}(\Delta)^T\right|. \tag{6-6}$$

With the determinant property as used in $\mathrm{diag}(\Delta)^T\mathrm{diag}(\Delta) = \mathrm{diag}(\Delta)$, (6-6) becomes

$$\tilde{I}(\Delta) = \tfrac{1}{2}\log\left|\mathbf{I}_{2N_t} + \frac{\sigma_x^2}{2}\frac{1}{\lambda}\cdot \mathbf{H}'^{\,T}_{\mathrm{R}}\cdot\mathrm{diag}(\Delta)\cdot\mathbf{H}'_{\mathrm{R}}\right|. \tag{6-7}$$

The corresponding optimization problem for selecting α comparators can be expressed as

$$\begin{aligned} &\max_{\Delta}\ \tfrac{1}{2}\log\left|\mathbf{I}_{2N_t} + \frac{\sigma_x^2}{2\lambda}\mathbf{H}'^{T}_{\mathrm{R}}\mathrm{diag}(\Delta)\mathbf{H}'_{\mathrm{R}}\right| \\ &\text{s.t.}\quad \Delta_{1:2N_r} = 1, \quad \Delta_i \in \{0,1\}, \quad \sum_{i=1}^{2N_r+\alpha_{\mathrm{full}}} \Delta_i = 2N_r + \alpha, \end{aligned} \tag{6-8}$$

which is non-convex, due to the binary constraint on Δ_i. Relaxing (6-8) by allowing for continuous values for Δ_i makes (6-7) a convex problem that can be solved by applying convex optimization. The relaxed problem reads as

$$\begin{aligned} &\max_{\Delta}\ \tfrac{1}{2}\log\left|\mathbf{I}_{2N_t} + \frac{\sigma_x^2}{2\lambda}\mathbf{H}'^{T}_{\mathrm{R}}\mathrm{diag}(\Delta)\mathbf{H}'_{\mathrm{R}}\right| \\ &\text{s.t.}\quad \Delta_{1:2N_r} = 1, \quad 0 \le \Delta_i \le 1, \quad \sum_{i=1}^{2N_r+\alpha_{\mathrm{full}}} \Delta_i = 2N_r + \alpha. \end{aligned} \tag{6-9}$$

Note that, if only $\alpha < \alpha_{\mathrm{full}}$ are used, the sum rate becomes a function of the selected comparators. For selecting the comparator configuration, the indices of the $(2N_r + \alpha)$ largest entries in Δ are identified. These indices are used to select the rows in matrix $\mathbf{B}$ of the fully connected comparator network. The matrix of the corresponding partially connected network is termed $\mathbf{B}_{\mathrm{select}}$, and the sum rate can be computed via (4-1).

The comparator selection problem can be solved through numerical methods, such as semi-definite programming. This section presents the formulation of the optimization problem, where we apply a similar strategy to the one used in the receive antenna selection problem [14].

Proposed Convex optimization Strategy CVX

1. **Base channel matrices**

$$\mathbf{H} \in \mathbb{C}^{N_r \times N_t}, \quad \mathbf{H}_{\mathrm{R}} = \begin{bmatrix} \Re\mathbf{H} & -\Im\mathbf{H} \\ \Im\mathbf{H} & \Re\mathbf{H} \end{bmatrix} \in \mathbb{R}^{2N_r \times 2N_t},$$

$$\sigma_{\mathbf{n}} = \sqrt{\sigma_x^2/\mathrm{SNR}}, \quad \mathbf{C}_{\mathbf{x}_R} = \tfrac{1}{2}\sigma_x^2 \mathbf{I}_{2N_t}, \quad \mathbf{C}_{\mathbf{n}_{\mathrm{R}}} = \tfrac{\sigma_n^2}{2}\,\mathbf{I}_{2N_r}$$

2. **Fully comparator network and selection vector**

$$\mathbf{B}_{\mathrm{full}} = \tfrac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I}_{2N_r} \\ \mathbf{B}'(N_r, \texttt{full}) \end{bmatrix}, \qquad \Delta \in \mathbb{R}^{M'_{\mathrm{full}}}.$$

3. **Convex optimization objective function**

$$\max_{\Delta} \ \frac{1}{2} \log\det\Big(\mathbf{I}_{2N_t} + \tfrac{\sigma_x^2}{2\lambda}(\mathbf{H'_{full}}^{T} \operatorname{diag}(\Delta)\, \mathbf{H'}_{\mathrm{full}}\Big)$$

subject to

$$\Delta_i = 1 \ (i \le 2N_r), \quad 0 \le \Delta_i \le 1, \quad \sum_i \Delta_i = 2N_r + \alpha.$$

4. **Computational of sum rate with** (4-1)

$$\mathbf{H}'_{\mathrm{full_R}} \ = \ \sqrt{\tfrac{2}{\pi}}\, \mathbf{K}_{\mathrm{full_R}}\, \mathbf{B}_{\mathrm{full}}\, \mathbf{H}_{\mathrm{R}}.$$

   (a) Select the $2N_r + \alpha$ largest indices from Δ to form the binary vector $\boldsymbol{\Delta_0}$.

   (b) Build $\mathbf{B}_{\mathrm{select},\alpha}$ by taking the corresponding rows from $\mathbf{B}_{\mathrm{full}}$.

   (c) Computation of the quantization covariance

$$\mathbf{C}_{\mathbf{z}_{\mathrm{R}}} = \mathbf{B}_{\mathrm{select},\alpha}(\mathbf{H_r}\mathbf{C}_{\mathbf{x}_{\mathrm{R}}}\mathbf{H}_{\mathrm{R}}^{T})\mathbf{B}_{\mathrm{select},\alpha}^{T} + \mathbf{B}_{\mathrm{select},\alpha}\mathbf{C}_{\mathbf{n}_{\mathrm{R}}}\mathbf{B}_{\mathrm{select},\alpha}^{T}.$$

   (d) The diagonal normalization matrix was obtained and expressed in

the following equation

$$\mathbf{K}_{\mathrm{R}} = \mathrm{diag}\Big(\frac{1}{\sqrt{(\mathbf{C}_{\mathbf{z}_{\mathrm{R}}})}}\Big).$$

(e) Define the effective channel trough Bussgang decomposition

$$\mathbf{H'}_{\mathrm{R}} = \sqrt{\tfrac{2}{\pi}}\,\mathbf{K}_{\mathrm{R}}\mathbf{B}_{\mathrm{select},\alpha}\mathbf{H}_{\mathrm{R}}.$$

(f) The next step is to compute the effective noise covariance

$$\mathbf{C}_{\eta'_{\mathrm{R}}} = \tfrac{2}{\pi}\Big(\arcsin(\mathbf{K}_{\mathrm{R}}\mathbf{C}_{z_{\mathrm{R}}}\mathbf{K}_{\mathrm{R}}) - \mathbf{K}_{\mathrm{R}}\mathbf{C}_{z_{\mathrm{R}}}\mathbf{K}_{\mathrm{R}}\Big) + \mathbf{K}_{\mathrm{R}}\mathbf{B}_{\mathrm{select},\alpha}\mathbf{C}_{n_{\mathrm{R}}}\mathbf{B}^{T}_{\mathrm{select},\alpha}\mathbf{K}_{\mathrm{R}}.$$

(g) Finally, by substituting, the expression for evaluating the actual sum rate was obtained.

$$\mathbf{I}_{\mathrm{select}} = \tfrac{1}{2}\log_2\Big(\Big|\mathbf{I}_{(2N_r+\alpha)} + \mathbf{C}^{-1}_{\eta'_{\mathrm{R}}}\tfrac{\sigma_x^2}{2}\mathbf{H'}_{\mathrm{R}}\mathbf{H'}^{T}_{\mathrm{R}}\Big|\Big).$$

Table 6.1 below, presents the description and characteristics of the three optimization strategies; Greedy Search MMSE, Sequential SINR [14], and Proposed Convex Optimization.

Table 6.1: Comparison between optimization strategies comparator selection methods

Method	**Description**	**Characteristics**
Greedy Search MMSE [3]	This strategy includes the iterative selection of comparators by evaluating all possibilities at each step.	- Iterative process to select the comparators. - Evaluates all remaining rows at each step. - Chooses and replace the one with the minimum value of MSE.
Sequential SINR [11]	The SINR optimization focuses on maximizing the weakest signal.	- Focuses on the weakest stream worst SINR and high value of MMSE. - The main objective is the comparator maximizing SINR gain.
Proposed Convex Optimization	Address the non-convex optimization challenge by employing convex optimization to relax the binary-selection issue.	- By relaxation binary selection to continuous values. - Solves a convex optimization problem. - Selects the α comparators with the highest contribution.

7
Numerical results

This section presents the simulation results for the sum rate (b/s/Hz) of different comparator network configurations and the BER of the detectors. The analysis covers performance and computational complexity using the Quadrature Phase Shift Keying (QPSK) modulation scheme. The SNR is a measure to relate the signal power and the noise power and we it is considered as (σ_x^2/σ_n^2). In the following, we consider $\sigma_x^2 = 1$. The transmission power is considered normalized. Then, the inverse of noise variance represents the radio Signal Noise to Radio represented by the following equation

$$\mathrm{SNR}_{\mathrm{dB}} = 10 \log_{10} \left(\frac{1}{\sigma_n^2} \right), \tag{7-1}$$

where σ_n^2 represents the noise covariance.

The quantization process introduces additional noise independent of channel noise. The intensity of this noise is influenced by the SNR and the configuration of the channel matrix. However, it introduces nonlinear effects and performance degradation.

7.1
Sum rate

To improve the sum rate, simulations were considered $N_t = 4$ transmitted antennas and $N_r = 12$ receive antennas, MU-MIMO UL. In the Fig. 7.1 we show the simulation for full, random, proposed-optimized, with the comparators network selection, and without the comparators network. The simulations were tested with 2000 channel realizations and different values, for $\alpha = 2N_r, 4N_r, 8N_r$. We can observe the improvement of the sum rate by employing the comparators' network selection. The fully connected network shows the best performance among others due to the largest number of antennas, which increases the computational cost. In

this case, for $N_r = 12$ the $\alpha_{full} = 276$. On the other hand, the proposed optimized comparator network achieves better performance than both the configuration without a comparator network and randomly connected comparator network. The performance of proposed comparators increases with the number of receive antennas and the α factor. In Fig. 7.1 with $\alpha = 8N_r$, it can be observed that the results of the proposed optimized CN can nearly reach the full comparator one, but using a smaller number of received comparators in the selection with low computational cost. The second Fig. 7.2 shows the comparison of the sum rate of the proposed optimized CN and benchmarks considering fixed SNR=15dB lower bound versus different number of receive antennas N_r (4,6,8,10,12,14,16), $N_t = 4$ and $\alpha = 4N_r$, and we can observe that the proposed optimized CN show a nearly performance to Full CN. In Fig. 7.3 it is shown the comparison between the proposed optimized CN and the optimal exhaustive search sum rate considering $N_t = 2$, $N_r = 3$, and for $\alpha = 2N_r$, both curves show an increasing behavior of the sum rate as the SNR increases. The Fig. 7.4 presents the sum rate of the proposed optimized connected comparator networks (CN) and benchmarks versus SNR for $N_t = 4$, $N_r = 12$, including the optimization techniques of partially connected network and alpha $\alpha = 2N_r$, considering 1-bit ADCs in MIMO-UL. The last figure Fig. 7.5 in this section shows the sum rate of the proposed optimized connected CN and benchmarks versus SNR for $N_t = 4$, $N_r = 4$, and alpha $\alpha = 2N_r$, considering 1-bit ADCs in MIMO-UL including the optimal exhaustive search and compared with the other optimization techniques.

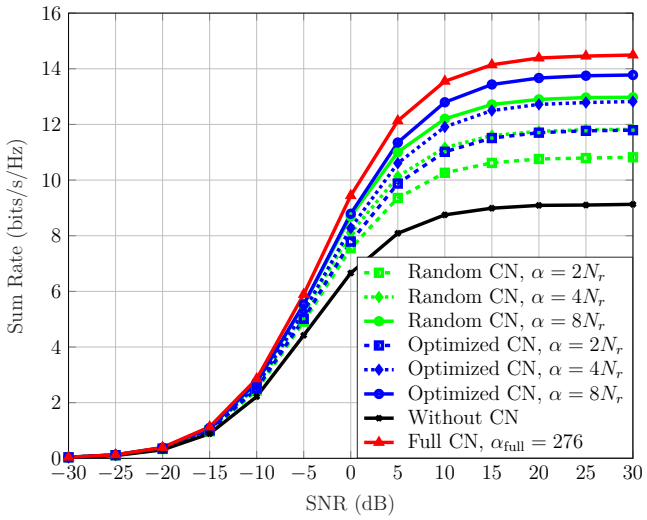


Figure 7.1: Sum Rate of the proposed optimized connected comparator networks (CN) and benchmarks versus SNR for $N_t = 4$, $N_r = 12$, and different alpha values $\alpha = 2N_r, 4N_r, 8N_r$, considering 1-bit ADCs in MIMO-UL.

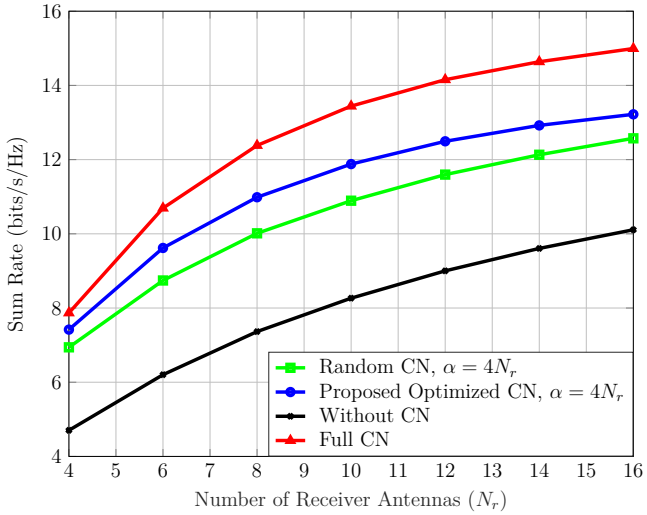


Figure 7.2: Sum Rate of the proposed optimized comparator networks (CN) and benchmarks versus number of receive antennas, SNR=15dB, $N_t = 4$,$\alpha = 4N_r$, considering 1-bit ADCs.

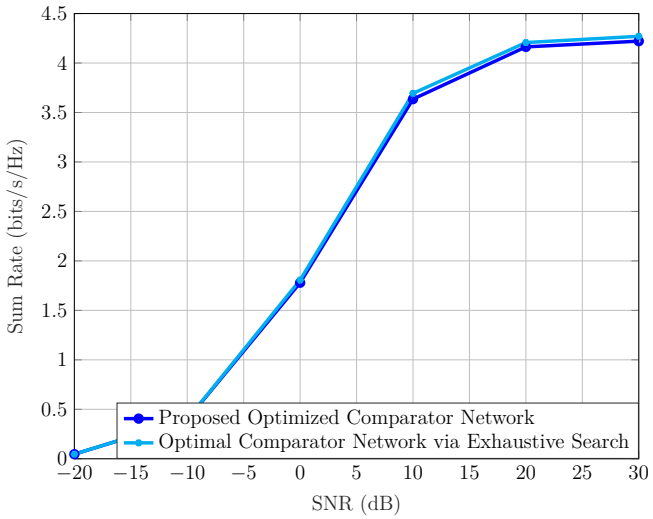


Figure 7.3: Comparison between the proposed optimized Comparator Network and the optimal exhaustive search sum rate for $N_t = 2$, $N_r = 3$ and $\alpha = 2N_r$.

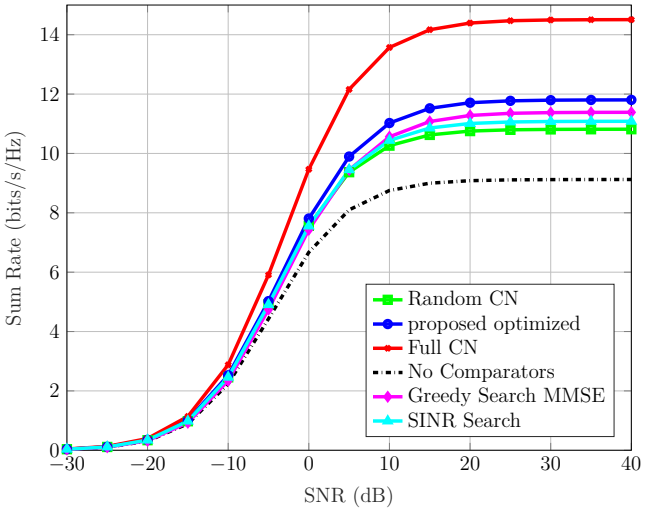


Figure 7.4: Sum Rate of the proposed optimized connected comparator networks (CN) and benchmarks versus SNR for $N_t = 4$, $N_r = 12$, and $\alpha = 2N_r$, considering 1-bit ADCs in MIMO-UL.

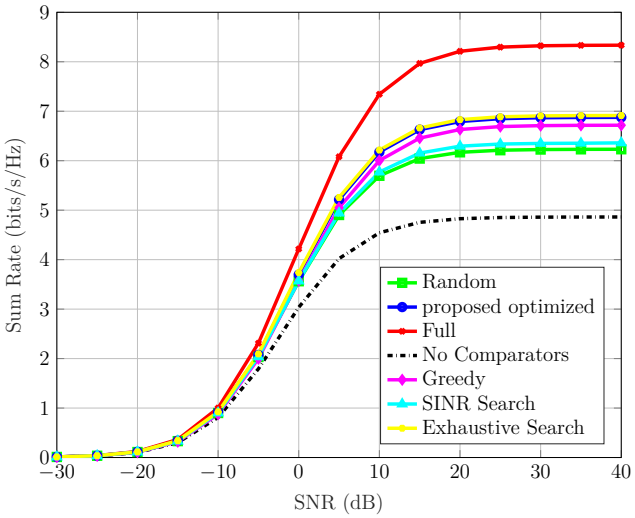


Figure 7.5: Sum Rate of the proposed optimized connected comparator networks (C.N.) and benchmarks versus SNR for $N_t = 4$, $N_r = 4$, and alpha $\alpha = 2N_r$, considering 1-bit ADCs in MIMO-UL.

7.2 CDF of Sum Rate

The Cumulative Distribution Function (CDF) evaluates a wide range of performance metrics in communication systems, such as SNR, BER, and latency. In this study, we employed the CDF to provide insights into the behavior and performance of the system in terms of sum rate across different comparator network configurations, including fully connected, optimized proposed method, randomly connected, and without a comparator network. The simulation results presented in Fig. 7.6, were obtained using $N_r = 16$, $N_t = 4$, and $\alpha = 4N_r$, with 2000 channel realizations. In this case, a CDF curve further to the right indicates that the system achieves higher transmission rates with greater probability. The results demonstrate that the fully connected comparator network yields the best performance, followed by the proposed optimized configuration, the randomly connected network, and the setup without a comparator network. As shown in Fig. 7.6, the median sum-rate for the configuration without a comparator network is approximately 10.5 bits/s/Hz, for the randomly connected network is around 12.6 bits/s/Hz, for the optimized proposed configuration is about 13.5 bits/s/Hz, and for the fully connected comparator network is approximately 15.5 bits/s/Hz.

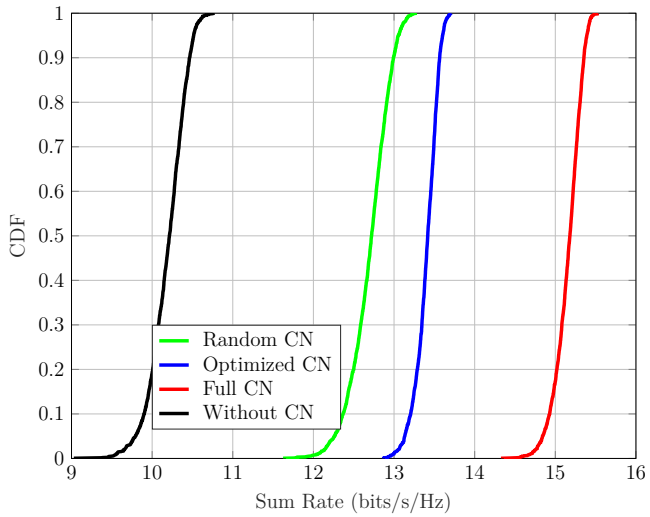


Figure 7.6: Cumulative Distribution Function (CDF) of the sum rate for the proposed optimized Comparator Networks (CN) and benchmarks configurations, with $N_t = 4$, $N_r = 16$, and $\alpha = 4N_r$, SNR = 20dB, The scenario considers 1-bit ADCs in MIMO-UL 2000 channel realizations.

7.3
Execution Time in Matlab

The table 7.1 shows a comparison of the execution times of different algorithms used in an antenna or sub network selection problem in a MIMO system wit 1-bit quantization, with fixed SNR in 15dB. The PC, equipped with an Intel(R) Core(TM) i-7 processor and 16.0 GB of RAM, was used for 1000 channel realizations with N_t=4 and N_r=4, and $\alpha = 2N_r$. We can observe in Table 7.1 shows that the execution time for the proposed optimized convex is longer than that of the other optimization strategies despite performing well. The exhaustive search technique demands the most time due to its exceptionally high computational complexity.

Table 7.1: Comparison of Algorithm Execution Times

Algorithm	**Time (seconds)**
Proposed (CVX)	1.0576
Greedy Search	0.8372
SINR Selection	0.0005
Exhaustive Search	254.64

7.4
Bit Error Rate (BER)

The BER is defined as the fractional ratio between the bit error and the total number of bits transmitted. The Fig. 7.7 shows the BER performance between typical techniques detection such as Minimum Mean Square Error (MMSE), Matched Filter (MF), Zero Forcing (ZF), included 1-bit quantization, where it can be observed that the matched filter does not have the best performance and that the linear MMSE detector has better performance. The Fig. 7.8 presents the simulation of the BER comparison benchmarks using QPSK modulation and LRA MMSE detection for all techniques Random, Full, Greedy search, proposed convex optimization, SINR with comparator network. It can be seen that the Greedy Algorithm has better performance than the others. The Greedy algorithm has the better performance because the main objective is to reduced the square error, the second best performance is proposed convex optimization and the last one is the SINR strategy but is nearly to the proposed convex optimization.

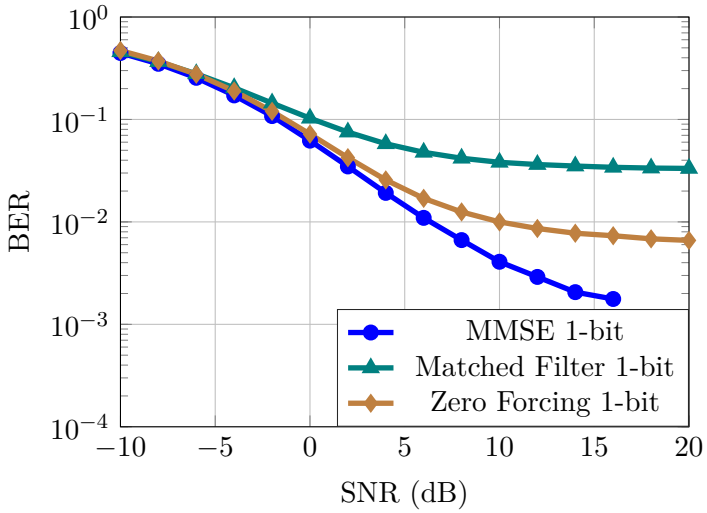


Figure 7.7: BER performance of MMSE, Matched Filter, and Zero Forcing detectors with 1-bit quantization across different SNR values.

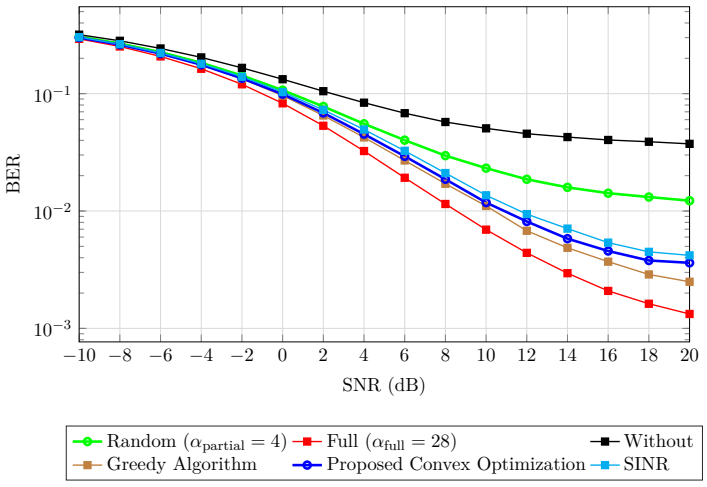


Figure 7.8: Performance of BER with comparator network $N_r = 4$ for $\alpha = 2N_r$

8
Conclusions

This thesis addresses the problem of designing a comparator network that provides additional information for multiple antenna receivers with 1-bit ADCs. This problem is a combinatorial problem that is difficult to solve optimally for medium and large-scale antenna systems. The proposed method focuses on the maximization of the sum-rate approximation. By relaxing the corresponding binary problem formulation, the optimization problem can be solved using convex optimization methods with polynomial complexity. The numerical results confirm that the optimized comparator networks nearly achieve the performance of the exhaustive search solution in simulation in example for $N_r = 3$ and $\alpha = 2N_r$. By using the optimized proposed comparator network achieves nearly to 12 (bits/s/Hz), the sum rate can be significantly improved in comparison to random comparator networks and systems with only 1-bit ADCs. The other test is about CDF of sum rate obtained the following values without a comparator network is approximately 10.5 bits/s/Hz, for the randomly connected network is around 12.6 bits/s/Hz, for the optimized proposed configuration is about 13.5 bits/s/Hz, and for the fully connected comparator network is approximately 15.5 bits/s/Hz. Another numerical result is about execution time of all optimization strategies in the partially comparator network despite the proposed optimization strategy achieves better results in the performance the time execution is more than the others. With this consideration as a future work would be consider development of a practical algorithm as an alternative to CVX. In addition to that, it could be interesting to extend the proposed system model to scenarios involving spatial or temporal oversampling. It is also interesting to incorporate coding schemes to create a more practical system employing iterative detection and decoding (IDD) strategies adapted to the comparator network.

Bibliography

[1] FERNANDES, A. B. L. B., **On mimo communications systems with 1-bit quantization and comparator networks at the receiver**, 2021.

[2] WIRASAPTA, A. H.; NUGROHO, P. ; WIBOWO, S. B., **Designing qpsk modulator using ltspice-based discrete components**, In: 2022 IEEE INTERNATIONAL CONFERENCE ON COMMUNICATION, NETWORKS AND SATELLITE (COMNETSAT), pp. 283–288, 2022.

[3] FERNANDES, A. B. L. B.; LANDAU, L. T. N., **Comparator Network Aided Channel Estimation and Achievable rates for MIMO Receivers with 1-bit Quantization**, In: 2021 IEEE STATISTICAL SIGNAL PROCESSING WORKSHOP (SSP), pp. 381–385, 2021.

[4] PENNANEN, H.; HÄNNINEN, T.; TERVO, O.; TÖLLI, A. ; LATVA-AHO, M., **6g: The intelligent network of everything**, IEEE Access, 2024.

[5] WALDEN, R., **Analog-to-digital converter survey and analysis**, IEEE J. Sel. Areas Commun., vol. 17, no. 4, pp. 539–550, 1999.

[6] MO, J.; HEATH, R. W., **Capacity Analysis of One-Bit Quantized MIMO Systems With Transmitter Channel State Information**, IEEE Trans. Signal Process., vol. 63, no. 20, pp. 5498–5512, 2015.

[7] LI, Y.; TAO, C.; SECO-GRANADOS, G.; MEZGHANI, A.; SWINDLEHURST, A. L. ; LIU, L., **Channel Estimation and Performance Analysis of One-Bit Massive MIMO Systems**, IEEE Trans. Signal Process., vol. 65, no. 15, pp. 4075–4089, Aug. 2017.

[8] SHAO, Z.; LANDAU, L. T. N. ; DE LAMARE, R. C., **Dynamic oversampling for 1-bit ADCs in large-scale multiple-antenna systems**, IEEE Trans. Commun., vol. 69, no. 5, pp. 3423–3435, 2021.

[9] RAO, S.; SECO-GRANADOS, G.; PIRZADEH, H.; NOSSEK, J. A. ; SWINDLEHURST, A. L., **Massive MIMO channel estimation**

with low-resolution spatial sigma-delta ADCs, IEEE Access, vol. 9, pp. 109320–109334, 2021.

[10] PIRZADEH, H.; SECO-GRANADOS, G.; RAO, S. ; SWINDLEHURST, A. L., **Spectral efficiency of one-bit sigma-delta massive MIMO**, IEEE J. Sel. Areas Commun., vol. 38, no. 9, pp. 2215–2226, 2020.

[11] FERNANDES, A. B. L. B.; SHAO, Z.; LANDAU, L. T. N. ; DE LAMARE, R. C., **Multiuser-MIMO Systems Using Comparator Network-Aided Receivers With 1-Bit Quantization**, IEEE Trans. Commun., vol. 71, no. 2, pp. 908–922, 2023.

[12] CHEN, Z.; LONG, W.-X.; CHEN, R. ; MORETTI, M., **Low-cost OAM spatial oversampling receiver with 1-bit quantized comparators and DL-based detector**, IEEE Commun. Lett., vol. 28, no. 6, pp. 1407–1411, 2024.

[13] SAMPAIO, L.; LANDAU, L., **Comparator network-aided receivers for spatially and temporally correlated MIMO systems with 1-bit quantization**, In: 2025 IEEE WIRELESS COMMUNICATIONS AND NETWORKING CONFERENCE (WCNC), pp. 1–6, Milan, Italy, 2025.

[14] DUA, A.; MEDEPALLI, K. ; PAULRAJ, A., **Receive antenna selection in MIMO systems using convex optimization**, IEEE Trans. Wireless Commun., vol. 5, no. 9, pp. 2353–2357, 2006.

[15] BUSSGANG, J. J., **Crosscorrelation functions of amplitude-distorted gaussian signals**, 1952.

[16] MEZGHANI, A.; NOSSEK, J., **Capacity Lower Bound of MIMO Channels with Output Quantization and Correlated Noise**, In: 2012 INTERNATIONAL SYMPOSIUM ON INFORMATION THEORY, Cambridge, MA, Jul. 2012.

[17] SHAO, Z.; DE LAMARE, R. C. ; LANDAU, L. T., **Iterative detection and decoding for large-scale multiple-antenna systems with 1-bit ADCs**, IEEE Wireless Communications Letters, vol. 7, no. 3, pp. 476–479, 2017.

[18] JEON, Y.-S.; LEE, N.; HONG, S.-N. ; HEATH, R. W., **One-bit sphere decoding for uplink massive MIMO systems with one-bit ADCs**, IEEE Trans. Wireless Commun., vol. 17, no. 7, pp. 4509–4521, 2018.

[19] BAZRAFKAN, A.; ZLATANOV, N., **Asymptotic capacity of massive MIMO with 1-bit ADCs and 1-bit DACs at the receiver and at the transmitter**, IEEE Access, vol. 8, pp. 152837–152850, 2020.

[20] RADBORD, A.; ATZENI, I. ; TÖLLI, A., **Multi-user Data Detection in Massive MIMO with 1-bit ADCs**, In: ICASSP 2023 - 2023 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING (ICASSP), pp. 1–5, 2023.

[21] RADBORD, A.; ATZENI, I. ; TÖLLI, A., **Enhanced data detection for massive MIMO with 1-bit ADCs**, In: 2023 57TH ASILOMAR CONFERENCE ON SIGNALS, SYSTEMS, AND COMPUTERS, pp. 427–432, 2023.

[22] ABDELHAMEED, D.; HUSSIEN, H. S. ; MOHAMED, E. M., **Low complexity MIMO detection technique for 1-bit ADC MIMO-CEM using adaptive sphere decoding**, In: 2017 34TH NATIONAL RADIO SCIENCE CONFERENCE (NRSC), pp. 335–342, 2017.

[23] SAMPAIO, L.; LANDAU, L., **Spatially and temporally correlated channel estimation and detection for comparator network-aided MIMO receivers with 1-bit ADCs**, EURASIP Journal on Advances in Signal Processing, vol. 2025, no. 34, 2025.

[24] BUSSGANG, J. J., **Crosscorrelation functions of amplitude-distorted Gaussian signals**, Res. Lab. Elec., MIT, vol. Tech. Rep. 216, Mar. 1952.

[25] JACOVITTI, G.; NERI, A., **Estimation of the autocorrelation function of complex Gaussian stationary processes by amplitude clipped signals**, IEEE Trans. Inf. Theory, vol. 40, no. 1, pp. 239–245, 1994.

[26] YANG, S.; HANZO, L., **Fifty years of MIMO detection: The road to large-scale mimos**, IEEE Communications Surveys Tutorials, vol. 17, no. 4, pp. 1941–1988, 2015.

[27] DAHA, M. Y.; KHURSHID, K.; ASHRAF, M. I. ; HADI, M. U., **Optimizing signal detection in MIMO systems: Ai vs approximate and linear detectors**, IEEE Communications Letters, vol. 28, no. 10, pp. 2387–2391, 2024.

[28] PAULRAJ, A.; NABAR, R. ; GORE, D., *Introduction to space-time wireless communications*, Cambridge university press, 2003.

A
Appendix

A.1
Linear Detection Derivations [1]

A.1 Derivation of the Cross-Correlation Matrix between Received and Transmitted Data Signals

The cross-correlation matrix between $\mathbf{z}_R$ and $\mathbf{x}_R$ from (5-5) is calculated as follows

$$\begin{aligned}\mathbf{C}_{z_R x_R} &= \mathbb{E}\left[\mathbf{z}_R \mathbf{x}_R^H\right] = \mathbb{E}\left[(\mathbf{B}\mathbf{y}_R)\,\mathbf{x}_R^H\right] = \mathbb{E}\left[(\mathbf{B}(\mathbf{H}_R\mathbf{x}_R + \mathbf{n}_R))\,\mathbf{x}_R^H\right] \\ &= \mathbb{E}\left[(\mathbf{B}\mathbf{H}_R\mathbf{x}_R + \mathbf{B}\mathbf{n}_R)\,\mathbf{x}_R^H\right] \\ &= \mathbb{E}\left[\mathbf{B}\mathbf{H}_R\mathbf{x}_R\mathbf{x}_R^H + \mathbf{B}\mathbf{n}_R\mathbf{x}_R^H\right] \\ &= \mathbf{B}\mathbf{H}_R\mathbf{C}_{x_R},\end{aligned} \tag{B-1}$$

where it is considered that $\mathbf{x}_R$ is uncorrelated with $\mathbf{n}_R$. Thus, $\mathbb{E}\left[\mathbf{n}_R\mathbf{x}_R^H\right] = 0$.

A.2 Derivation of the Auto-Correlation of the Received Data Signal

The auto-correlation of $\mathbf{z}_R$ from (5-6) is calculated as follows

$$\begin{aligned}\mathbf{C}_{z_R} &= \mathbb{E}\left[\mathbf{z}_R\mathbf{z}_R^H\right] = \mathbb{E}\left[(\mathbf{B}\mathbf{y}_R)(\mathbf{B}\mathbf{y}_R)^H\right] = \mathbb{E}\left[\mathbf{B}\mathbf{y}_R\mathbf{y}_R^H\mathbf{B}^H\right] \\ &= \mathbb{E}\left[\mathbf{B}(\mathbf{H}_R\mathbf{x}_R + \mathbf{n}_R)(\mathbf{H}_R\mathbf{x}_R + \mathbf{n}_R)^H\mathbf{B}^H\right] \\ &= \mathbb{E}\left[(\mathbf{B}\mathbf{H}_R\mathbf{x}_R + \mathbf{B}\mathbf{n}_R)(\mathbf{x}_R^H\mathbf{H}_R^H + \mathbf{n}_R^H)\mathbf{B}^H\right] \\ &= \mathbb{E}\Big[\mathbf{B}\mathbf{H}_R\mathbf{x}_R\mathbf{x}_R^H\mathbf{H}_R^H\mathbf{B}^H + \mathbf{B}\mathbf{H}_R\mathbf{x}_R\mathbf{n}_R^H\mathbf{B}^H \\ &\quad + \mathbf{B}\mathbf{n}_R\mathbf{x}_R^H\mathbf{H}_R^H\mathbf{B}^H + \mathbf{B}\mathbf{n}_R\mathbf{n}_R^H\mathbf{B}^H\Big] \\ &= \mathbf{B}\mathbf{H}_R\mathbf{C}_{x_R}\mathbf{H}_R^H\mathbf{B}^H + \mathbf{B}\mathbf{C}_{n_R}\mathbf{B}^H,\end{aligned} \tag{B-2}$$

where it is considered that $\mathbf{x}_R$ is uncorrelated with $\mathbf{n}_R$. Thus, $\mathbb{E}\left[\mathbf{x}_R\mathbf{n}_R^H\right] = \mathbb{E}\left[\mathbf{n}_R\mathbf{x}_R^H\right] = 0$.