Addis Ababa University

Addis Ababa Institute of Technology (AAiT)

School of Electrical and Computer Engineering

Performance Evaluation of a Low Complexity Hybrid Beamforming for 5G mmWave MIMO Systems

By: Mengistu Shumet

Advisor: Dr. Murad Ridwan

A Thesis Submitted to Addis Ababa University in Partial Fulfilment of the Requirements for the Degree of Master of Science in Electrical and Computer Engineering (Communication and Networking)

November 2025

Addis Ababa, Ethiopia

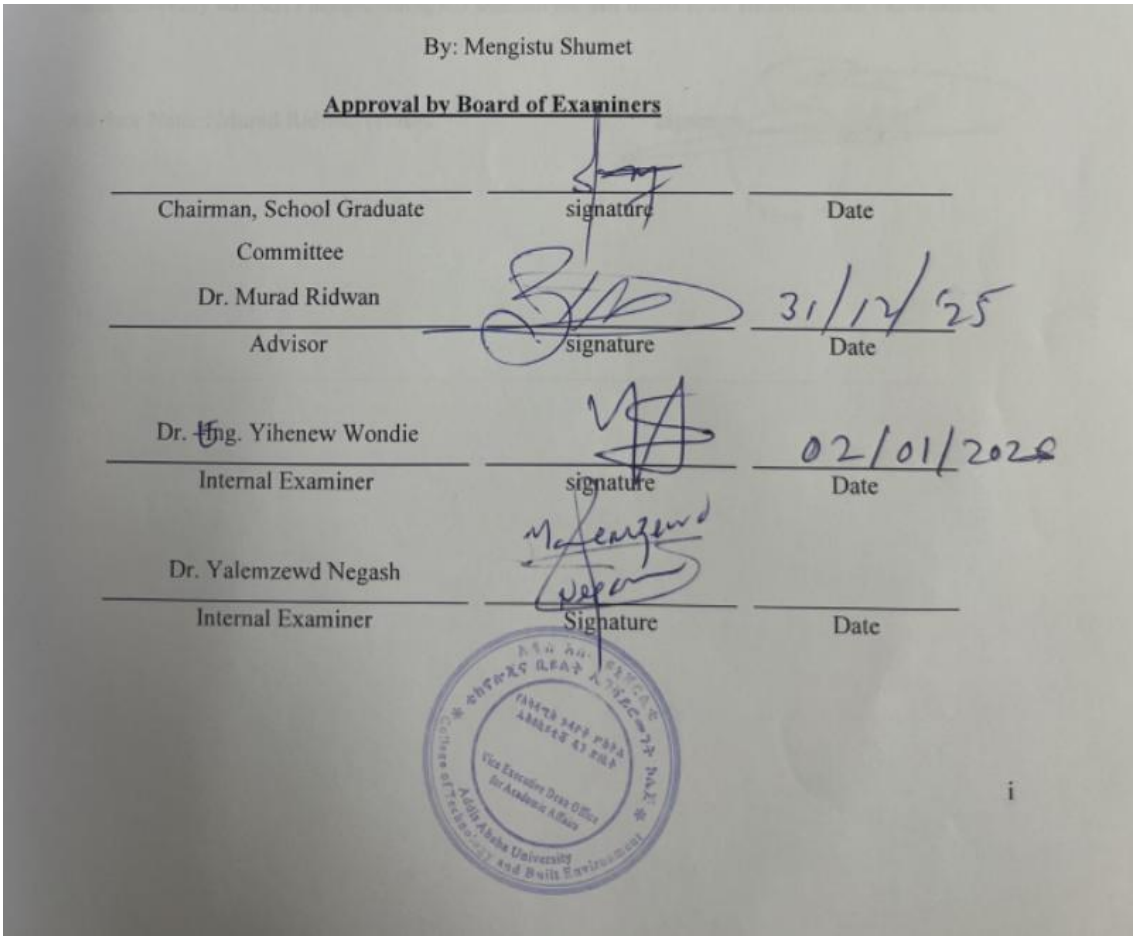
Addis Ababa University

Addis Ababa Institute of Technology (AAiT)

School of Electrical and Computer Engineering

A Thesis Submitted to Addis Ababa University in Partial Fulfilment of the Requirements for the Degree of Master of Science in Electrical and Computer Engineering (Communication and Networking)

Performance Evaluation of a Low Complexity Hybrid Beamforming for 5G mmWave MIMO Systems



By: Mengistu Shumet

Approval by Board of Examiners

Chairman, School Graduate Committee	signature	Date
Dr. Murad Ridwan Advisor	signature	31/12/25 Date
Dr. Ing. Yihenew Wondie Internal Examiner	signature	02/01/2026 Date
Dr. Yalemzewd Negash Internal Examiner	Signature	Date

i

Declaration

The undersigned thereby declares that the thesis is entirely my original work with complete citations to all sources and has not been submitted for credit toward a degree at this or any other university.

Name: Mengistu Shumet

Signature:_____.

Date of Submission:_____.

Place: Addis Ababa

As a university adviser, I am providing my consent for this thesis to be submitted for examination.

Advisor Name: Murad Ridwan (PHD). Signature:_____.

Acknowledgement

My heartfelt thanks go to my advisor Dr. Murad Ridwan for the support, coaching, guidance and continued support as of the thesis idea, throughout the whole process.

My utmost gratitude goes to my beautiful wife, Nardos Taddesse, as her encouragement, support in all the way she can please me so that to maintain my momentum, especially at times I felt exhausted and impatient.

Finally, my sincere acknowledgment goes to Dr.Fekadu Mihret, my undergraduate instructor at Bahirdar university, for his unwavering support and dedicated coaching from the start to the end.

Abstract

The demand for higher network capacity is drastically increasing and the effort to exploit new capacity boosting techniques are also growing simultaneously. Fifth-Generations networks have found millimeter wave signals, from 10mm to 1 mm wavelengths, to be the ideal frequency ranges to support the growing higher data rates requirement which reaches up to 10 Gbit/s with a vast bandwidth.

In massive MIMO systems, dependency on traditional beamforming methodologies is no longer valid due to issues such as increased computational burden, scalability issues, cost, and power requirements. For large-scale systems that are limited by power and cost the use of fully digital beamforming is not practical anymore because the baseband can only accommodate a certain number of analog to digital converters (ADCs) and signal mixers. The analog beamforming technique, however, is limited by flexibility and scalability issues, and unsuitable for MIMO systems.

To mitigate the mmWave signal attenuation and achieve higher data speeds as well as improved spectral efficiency, deployment of massive MIMO and hybrid analog-digital beamformers is now viewed as a feasible solution. In this thesis, a performance evaluation and complexity analysis has been carried out for the low complexity hybrid beamforming technique called the PE-AltMin algorithm. In particular, the spectral efficiency and computational complexity evaluation of this algorithm is studied in contrast with the conventional fully digital optimal beamformer and the MO-AltMin algorithm for the fully connected hybrid beamformer mapping. This thesis paper has shown that PE-AltMin algorithm is computationally less complex than MO-AltMin and fully-digital beamforming techniques. The PE-AltMin algorithm provides favorable tradeoff for better computational resources utilization where hardware simplicity matters, and real-time deployment are prioritized. MO-AltMin algorithm is preferable when computational resources are not the challenge and performance margins matter the most.

Keywords: mmWave, Massive MIMO, Hybrid Beamforming, Alternating Minimization, Phase Extraction, Manifold optimization, Conjugated Gradient Descent (CGC), Riemannian Conjugate Gradient (RCG), Low complexity.

Table of Contents

List of figures

List of tables

List of Acronyms

mmWave	Millimeter Wave	Tr(A)	Trace of matrix A
5G	Fifth Generation	$[\,]^H$	Hermitian transpose
HBF	Hybrid Beamforming	$\mathbb{E}[A]$	Mean of matrix A
ABF	Analog Beamforming	SVD	Singular value decomposition
DBF	Digital Beamforming	H	Channel matrix
AltMin	Alternating minimization	Y(t)	Received signal at the user
PE-AltMin	Phase Extraction Alternating Minimization	$\mathbf{a}_r$	Received array vector
MO-AltMin	Manifold Optimization	σ_n^2	Noise variance
SDR- AltMin	Semi Definite Relaxation	n	Noise
SIC	Successive Interference Cancellation	K	Number of users
FS-AltMin	Fast Search Alternative Minimization	F	Number of subcarriers
CSI	Carrier Symbol Interference	I	Identity matrix
MIMO	Multiple Input Multiple Output	Arg(A)	returns a matrix with the phases of the entries of **A** [1]
USPA	U-Slot Patch Antenna	$\mathcal{M}_{cc}^m$	Riemannian submanifold of $\mathbb{C}^m$
OMP	Orthogonal Matching Pursuit	$\nabla f(\boldsymbol{x})$	Euclidean gradient of $f(\boldsymbol{x})$
FLOPs	Floating Point Operations	s	vector of transmitted data symbols
MFLOPs	Mega Floating-Point Operations	$\mathbb{C}$	Complex circle
RCG	Riemannian Conjugate Gradient	$\text{Retr}_{\mathbf{x}}$	Retraction of vector x
SE	Spectral Efficiency	EE	Energy efficiency
N_t	Number of transmit antennas	N_s	Number of transmitted data streams
N_r	Number of receiving antennas	N_{RF}	Number of RF chains

N_{RF}^{t}	Numer of RF chain at the transmitter	N_{RF}^{r}	Number of RF chain at the user
F_{RF}	Analog beamformer matrix	F_{opt}	Fully digital optimal beamformer
F_{BB}	Digital baseband beamformer	F_{DD}	Unitary digital beamformer
W_{RF}	Analog combiner	W_{RF}	Digital combiner
$\|\| \quad \|\|_F$	Frobenius norm		

1 Chapter One: Introduction

1.1 Background and Motivation

Fifth generation communication networks are aiming to increase the network capacity by the factor of one thousand (1000x) as the result of drastically growing demands of internet users [2]. The available wireless networks must be expanded proportionally so that to support the rapidly increasing interest of multimedia usage. In other words, 5G systems aim to offer gigabits per second (Gbits/s) data rates specifically to the high demanding wireless users. The bottleneck of existing wireless systems is the limited capacity of available frequency spectrums. Available frequency spectrums, the likes of microwave spectrum have limitations to accommodate the gigabit per second communication requirements. By this fact, it is intuitive to utilize the underutilized frequency bands as well as the yet to utilize frequency spectrums for cellular communication.

The existing frequency spectrums, such as the microwave spectrum, cannot satisfy the requirements of Gbits/s rate transmission and as a result, the scarcity of the frequency spectrum in the current wireless system is a bottleneck/fundamental challenge to further capacity increase. Therefore, to meet increasing demand for communications, use of underutilized frequency/spectrum bands (i.e., the frequencies that are yet exploited for wireless communications) is necessary. By this motivation, the research institutions and researchers found and attracted their attention towards the millimeter wave communication technology and it is being implemented by it in the 5G networks.

Among the methods to maximize network capacity, improving the spectral efficiency by using techniques like massive multiple input multiple output (mMIMO) and advanced channel coding are dominant these times [1]. To mention other options available and used for improving spectral efficiency available and used for improving spectral efficiency are the likes of enabling advanced operations such as cloud radio access network (cloud-RANs) and network densifications such as deploying small cells.

Many analog subarrays are combined to create massive hybrid arrays, and each of the subarrays has its own digital processing circuit. It is an appealing option for the upcoming millimeter-wave

(mmWave) mobile communications as it might offer the benefit of balancing cost and performance for large arrays. In addition, the hybrid setup can efficiently utilize signal energy in sparse mmWave channels by using analog beamforming subarrays such as phased arrays. However, the configuration of many digital chains provides the device additional beamforming flexibility and multiplexing capabilities [3].

Frequency ranges around 30 GHz to 300 GHz having a wavelength of 10 mm to 1 mm, respectively are commonly called millimeter waves and they were previously thought to be only useful for carrying outdoor point-to-point backhaul links and indoor high-resolution multimedia transmissions. However, they are nowadays identified as potential candidates for 5G wireless systems with a bandwidth capacity of 10 GHz [4].

First, the major limitations or problems faced by millimeter wave cellular systems are increased path loss and attenuation by structures (buildings, human bodies, and vehicles) and precipitation [5]. Since in mmWave systems, MIMO beamforming/precoding requires the use of large antenna sizes at both the transmitter and the receiver sides, this allows us to mitigate higher path loss and structurally produce highly focused antenna beams. Multiple academic research articles reveals that large-scale antenna array and beam formation strategy is the magic bullet to reduce the signal loss we incur in mm wave transmissions. The short wavelength of millimeter waves allows fitting additional antennas in the same physical space, leading to highly spatial multiplexing and robust directional beamforming. Notwithstanding these advantages, the major issue at present is on how to address the massive path loss in this segment of the spectrum.

Regardless of its benefits, millimeter wave communication is challenged with several factors [6] like.

1. Large Free-Space Attenuation: High-frequency signals suffer from greater free-space attenuation and short transmission reach.
2. High Sensitivity to Blockages: MmWave signals are sensitive to blockages (e.g buildings, human bodies or weather).
3. Beamforming condition: Due to large path-loss, a directive beamforming is required to concentrate the signal power to the desired receiver.

4. Hardware Complexity and Power Consumption: To support fully digital beamforming, each RF chain supports one antenna element, which results in a high-power consumption and hardware costs at mmWave frequencies.

Beamforming is a key technology to improve signal quality and the drawbacks of mmWave channel. It allows the transmitter and receiver to direct the signal toward a desired direction to utilize the signal-to-noise ratio (SNR) and reduce interference.

We have three main categories of beamforming structure:

1. Fully Digital Beamforming: Fully digital precoding requires radio frequency (RF) chains in the same number as antenna elements, which are composed of attenuators, signal mixers, analog-to-digital/digital-to-analog converters (ADCs/DACs) and power amplifiers [1] equivalent to the antenna elements. Conventional beamforming also referred to as traditional is performed in a digital domain and involves each antenna to have its own individual RF link. Nevertheless, because the RF link is expensive and high-power-consuming, the full digital beamforming is not applicable to wide scale system, except for the small-scale antenna system.

   Furthermore, as the millimeter wave system needs to integrate with a large-scale antenna array, the conventional method is no longer suitable for applying MIMO to the mmWave system. In addition, the usage of millimeter wave MIMO could be very costly and power intensive, especially for mobile devices making it impractical [7], even though the use of high number of antenna components is made possible by the tiny wavelengths of millimeter wave frequencies.
2. Analog beamforming: in this technique, a single radio frequency (RF) chain drives multiple antenna elements using phase shifters that reduce hardware complexity with a limited beamforming flexibility.
3. Hybrid Beamforming: A combination of digital and analog beamforming, where a reduced number of RF chains are connected to multiple antennas via phase shifters. This approach balances complexity and performance, making it suitable for mmWave MIMO systems [8]. There are two structures of hybrid beamforming, namely, the fully connected hybrid beamforming structure which enables each RF chain to connect to all antennas through

phase shifters, offering higher spectral efficiency and the partially connected hybrid beamforming structure which reduces the number of phase shifters by connecting each RF chain to a subset of antennas, leading to lower hardware costs and power consumption but a higher computational complexity compared to the fully connected hybrid beamforming structure [9].

The motivation behind studying low-complexity hybrid beamforming techniques stems from the fact that although hybrid beamforming reduces the number of RF chains, computational complexity remains a major concern, especially for real-time beamforming adaptation in dynamic environments. Traditional (fully-digital hybrid beamforming solutions rely on high-complexity optimization techniques like manifold optimization alternating minimization (MO-AltMin), which are computationally expensive for practical deployment [10] and due to the need to bridge the gap between fully digital beamforming and pure analog beamforming. Fully digital architecture achieves optimal performance but is impractical due to excessive power consumption and hardware costs. On the other hand, analog beamforming simplifies hardware but compromises spectral efficiency and beam flexibility.

Hybrid beamforming implementations through PE-AltMin, MO-AltMin and SDR-based approaches create a realistic middle ground that provides the following benefits:

- As the hybrid structures are demanding fewer radio frequency chains compared with the fully digital beamforming implementation, the hardware complexity and the power consumption requirement are significantly improving.

- The hybrid beamforming approaches can provide a closely approximate spectral efficiency performance of the optimal fully digital techniques by proper designing of the digital and analog beamformers.

- The capacity of phase extraction alternating minimization and semidefinite relaxation algorithms to leverage optimization techniques to improve the gain and to mitigate interference provides beamforming flexibility enhancements.

- Can mitigate the drawbacks of conventional beamforming techniques to support large scale MIMO system deployment without incurring additional computational costs.

Given the importance of hybrid beamforming in practical mmWave MIMO deployments, this thesis investigates and compares different hybrid precoding techniques to evaluate their performance, computational complexity, and feasibility for 5G mmWave systems.

Even though lots of hybrid beamforming algorithms are being formulated, there is still no optimum low complexity design yet. It's intuitive to assess the available beamforming strategies for 5G mmWave MIMO systems to know how complex, practical and to contribute on gaps. Therefore, this thesis aims to evaluate the complexity reduction of an alternating minimization-based algorithm called phase extraction alternating minimization algorithm compared with the manifold optimization alternating minimization counterpart and with that of the convention digital baseband beamformer.

1.2 Statement of the research problem

Traditional beamforming approaches are failing to meet the needs of 5G mmWave communication because of the fast wireless network expansion demand and rising need for high-capacity networks by users. Large-scale MIMO systems do not entertain fully digital beamforming because they require too many RF chains which incur computational complexity and hardware expenses are too high. Analog beamforming demonstrates an inability to handle multi-stream transmission therefore it fails to serve massive MIMO deployment needs.

The emerging solution of hybrid beamforming for the5G network is aiming to create a balanced trend by cutting down RF chain numbers while preserving multi-stream data transmission capabilities. Extensive field research is yet to establish a hybrid beamforming method that can be considered universally optimal. A hybrid beamforming system requires improvement to achieve near-optimal spectral efficiency at a significantly reduced computational complexity level which is necessary for practical implementation in extensive 5G MIMO systems utilizing millimeter-wave technology in the 30 GHz to 300 GHz frequency spectrum range.

Nevertheless, there are algorithms by researchers proposed as a low complexity beamforming technique for the 5G millimeter wave large scale MIMO systems to mitigate the constraints of the traditional techniques, efforts are still underway to get the optimum digital beamforming techniques. PE-AltMin, manifold Optimization Alternating Minimization (MO-AltMin) and

Semidefinite Relaxation (SDR) algorithms are among the proposed techniques. The later algorithm is used for a partially connected hybrid structure whereas the first two for the fully connected hybrid structure.

PE-AltMin is better low-complexity hybrid beamforming approach which serves as a practical choice instead of fully digital beamforming according to [1] & [11]. However, the evaluation of this algorithm's computational complexity and spectral efficiency should be done with comparative assessment of other alternative hybrid beamforming approaches. The investigation of this research focuses on evaluating the PE-AltMin algorithms to establish comparative performance against fully digital optimal, and MO-AltMin beamforming algorithms. The examination focuses on computational complexity and will overview spectral efficiency to establish understanding of measurement of their practical usability in mmWave systems implementing 5G technology.

The thesis addresses three primary questions through its investigation.

1. How PE-AltMin algorithm presents a less computational complexity beamforming technique compared with other approaches using the Big O analysis.
2. Can PE-AltMin achieve near-optimal spectral efficiency while maintaining low complexity? This is shown using simulation of spectral efficiency performance.
3. Tradeoffs between spectral efficiency and computational complexity.

The evaluation is based on complex metrics like time and number of iteration as well as spectral efficiency measurement for predefined data streams and RF chains. The PE-AltMin and manifold optimization alternating minimization (MO-AltMin) will be investigated using numerical formulations and spectral efficiency evaluation based on MATLAB simulations will be made for the two algorithms and for the conventional optimal digital precoder. For the spectral evaluation, two types of evaluations will be made: firstly, for an equal number of data streams and RF chains, and secondly, for the number of RF chains slightly greater than the number of data streams [1].

By formulating and analyzing these hybrid beamforming methods, this research seeks to identify an efficient algorithm that minimizes computational overhead while maintaining high spectral efficiency, making it suitable for next-generation wireless networks.

1.3 Objectives of the study

1.3.1 General Objectives

The main goal of this thesis is to evaluate the Phase Extraction Alternating Minimization (PE-AltMin) as a low-complexity hybrid beamforming approach for 5G mmWave MIMO systems through performance and complexity assessment. This thesis analyzes computational complexities and spectral efficiency of the PE-AltMin algorithm when comparing it to MO-AltMin, and Fully Digital Optimal beamforming methods.

1.3.2 Specific Objectives

a) Study the PE-AltMin algorithm for fully connected hybrid beamforming
b) Compare the computational complexity of PE-AltMin with MO-AltMin and Fully Digital Optimal beamformers to highlight its efficiency.
c) Study the impact of RF chains, number of antennas and data streams to complexity and spectral efficiency.

1.4 Scope and significance of the research

1.4.1 Scope of Research

This thesis evaluates the Phase Extraction Alternating Minimization (PE-AltMin) algorithm which serves as a minimal complexity hybrid beamforming solution for fully connected beamforming structure for the 5G mmWave MIMO networks. The evaluation examines PE-AltMin computational complexity and spectral efficiency together with other beamforming approaches namely,

- Manifold Optimization Alternating Minimization (MO-AltMin) algorithm
- Fully Digital Optimal Precoder

The research examines the performance of fully connected hybrid beamforming structure for the PE-AltMin and MO-AltMin. Numerical simulations implemented through MATLAB assess both spectral efficiency and computational complexity within the evaluation process. MATLAB is used to simulate spectral efficiency and computation complexity evaluation.

The investigation doesn't include algorithmic energy efficiency aspects because its main objective deals with the evaluation of complexity versus spectral efficiency trade-offs. This thesis does not contain analysis of hardware limitations or problems related to hybrid beamforming implementation that appear in actual hardware applications.

1.4.2 Significance of the Research

Hybrid beamforming techniques function as critical technology for 5G millimeter wave MIMO systems, although most current methods have high algorithmic computation or processing demands that make real-world deployment difficult. This thesis explores efficient yet low-complexity beamforming techniques that approach the near optimal performance known as PE-AltMin algorithm. Benchmarking of PE-AltMin algorithm with MO-AltMin and fully digital precoding enables researchers to understand performance levels to select complex and efficient hybrid beamforming approaches accordingly. The implementation of low-complexity algorithm, reduced-cost hybrid beam-forming solutions through algorithms like PE-AltMin becomes more viable compared to standard digital precoding because digital precoding remains impractical because of its high associated hardware expenses. The thesis outcomes will help develop hybrid beamforming technologies, which are essential for next-generation 5G and beyond networks that need efficient ultrahigh-frequency beamforming mechanisms. The established groundwork in this investigation provides a solid basis for both technical improvements and operational applications of low-complexity hybrid beamforming methods, which will maintain their suitability in wireless technology's progressing ecosystem.

1.5 Contributions

The PE-AltMin algorithm stands as the focus of this thesis as a low complexity hybrid beamforming technique while conducting a comparative performance evaluation with other hybrid beamforming techniques for 5G millimeter wave MIMO system.

a). Evaluation of the computational complexity of the PE-AltMin hybrid beamforming using the algorithm's upper bound referred to as Big O analysis in contrast to the MO-AltMin & optimal fully digital precoder, and

b). Comparison of algorithmic performance and spectral efficiency are the contributions of this thesis.

1.6 Literature Review

1.6.1 Overview of 5G mmWave Communication

The 5G network requires millimeter wave (mmWave) communication as its fundamental component because it enables large bandwidths along with high data rates. The bandwidth range of mmWave frequencies spans from 30 to 300 GHz (1mm to 10 mm wavelengths), enabling multi-Gbps data rates, while sub-6 GHz frequencies operate at lower intervals [4]. They support high antenna array density, enabling massive MIMO and beamforming benefits.

Advanced signal processing techniques are necessary to keep communication reliable because mmWave signals undergo high path loss along with atmospheric absorption and suffer from blocking caused by objects [2] and [12].

1.6.2 Challenges and Opportunities in mmWave Spectrum Utilization

Various barriers need to be resolved before mmWave technology can achieve successful implementation because it enables data speed beyond multi-gigabit-per-second rates.

- **High Path Loss and Signal Attenuation**: Signal attenuation and path loss significantly affect higher mmWave signals at free space when compared to lower-frequency bands therefore reducing their operational distance. The wide propagation losses of mmWave waves need dense small-cell deployment combined with advanced beamforming techniques to reduce the signal attenuation as described in [4]. The Friss transmission formula for free-space path loss (FSPL) in equation (2.1) demonstrates that path loss grows higher as antenna separation distance increases together with transmission frequency.
- **Blockage Sensitivity:** mmWave signals easily lose signal quality because of obstacles such as buildings and trees along with human bodies when passing through their path. Signal degradation becomes severe due to even minimal interference because these signals need robust beam-tracking systems and intelligent reflection methods [5]. This affects the indoor communication unless line of sight (LOS) is maintained.

- **Atmospheric Absorption and Weather Effects:** The reliability of mmWave communication links over long distances deteriorates because of rain and humidity as well as atmospheric gases which cause additional signal absorption. The situation becomes problematic especially when mmWave technologies are used for outdoor transmission and backhaul networks [4].
- **Hardware and Power Consumption Challenges:** The technical requirements of implementing mmWave communication technology include several hardware components combined with high power drawn by radio frequency chains using high resolution phase shifters and arranging extremely large antenna arrays. Technical limitations related to energy consumption and expenses block the progress of mmWave communication protocol implementation [7].
- **Beam Alignment and Mobility Management:** Due to their directional signals, mmWave frequencies require precise beam alignment methods for mobility management together with beam direction adjustment. The process of sustaining network stability remains complex for mobile devices moving in various locations since beam adjustment systems need adaptive methods [6].
- **Spectrum Availability and Regulation:** The availability of mmWave spectrum depends on the regulatory standards which differ from one region to another throughout the world. The maximum efficient spectrum utilization requires international regulatory standardization so the global networks can work together [9].

These challenges of millimeter wave technology do not hinder its adoption because it enables ultra-dense networks along with high-speed communications for vehicles and improved user experiences with augmented reality applications [9].

1.6.3 Review of Traditional Beamforming Approaches

The wireless communications process requires beamforming as a signal processing tool which both optimizes signal transmission towards the receivers while minimizing unwanted signal interference. A classification of beamforming solutions divides them into analog, digital and fully digital systems. The different approaches present different combinations of strengths and weaknesses specifically for millimeter-wave massive MIMO systems.

1.6.4 Analog Beamforming:

This technique applies to one RF signal chain together with phase shifters to guide antenna beams toward specified directions. Its power-efficient operation comes with the disadvantage of the reduced capability to handle multiple users at once [6]. At earlier stages, it was using single phase shifters to create beam at single frequency. Later, adjustable phase shifters used at each antenna elements so that to create adaptive and more directive beams. The general cons of analog beamformer are difficulties implementing advanced beamforming and beam steering technologies like creating nulls in specific directions, beam fine tuning limitations due to low resolution of quantized phase shifters, and difficulties supporting multiple streaming.

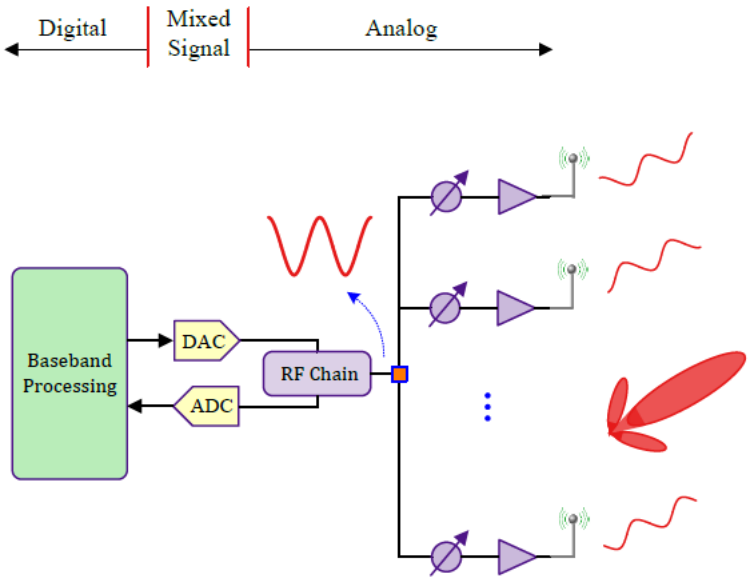


Figure 1.1:Simplified analog beamforming

1.6.5 Digital Beamforming:

Precise beam shaping and multi-stream transmission become possible because digital beamforming uses individual RF chains for every antenna. The implementation of this technology requires significant power usage combined with expensive hardware platforms [8] [10]. This technique supports multiple spatial streaming for spatial multiplexing. Digital beamforming provides frequency selective options to maximize bandwidth.

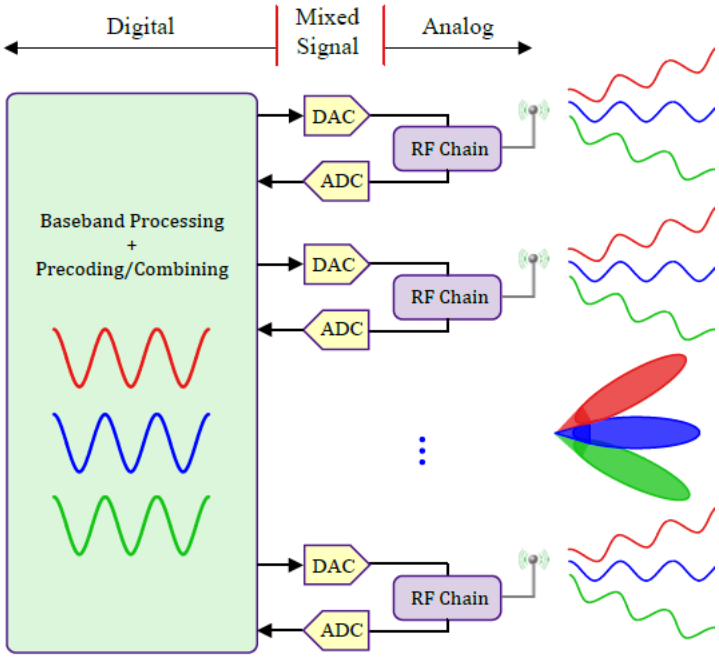


Figure 1.2: Simplified digital beamforming

1.6.5.1 Introduction to Hybrid Beamforming and Its Advantages in mmWave Systems

The implementation of hybrid beamforming allows networks to unite analog and digital beamforming methods to reach optimum network efficiency together with system performance levels. Its reduced RF chain implementation supports multi-stream transmission which makes it suitable for mmWave MIMO systems [1]. With this method, the power-hungry but highly adaptable digital beamforming solutions are balanced with the low-power-consuming but less flexible analog beamforming. Key advantages include:

- Power requirements decrease because of using fewer RF chains in the system.
- Spectral Efficiency reaches higher standards through hybrid beamforming because it enables the transmission of multiple users with sustained high data rates [11].
- The technology integrates effectively with large antenna arrays because it maintains scalability even when it avoids unnecessary cost increases.

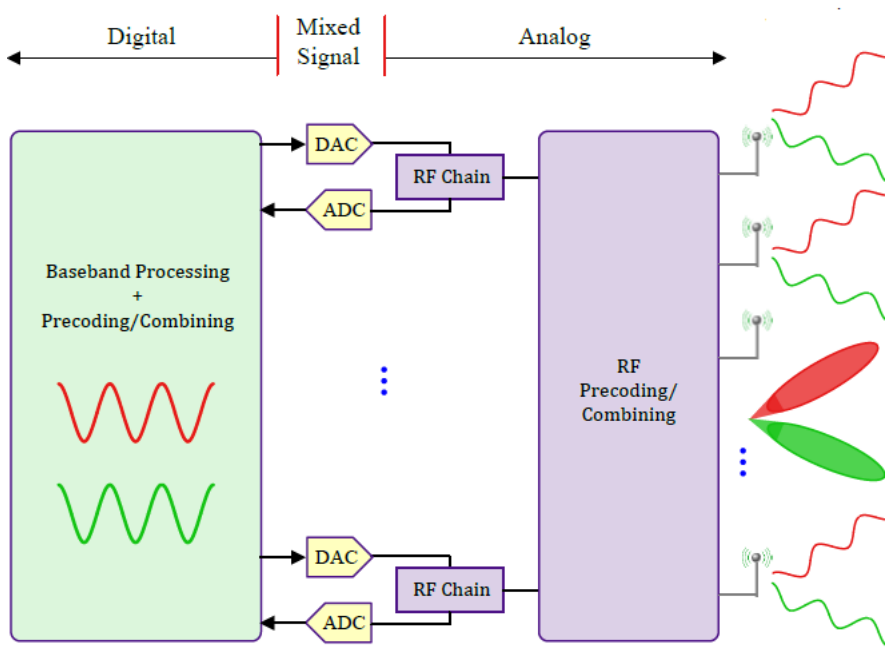


Figure 1.3: Simplified analog-digital hybrid beamforming.

1.6.6 Review of Existing Hybrid Beamforming Algorithms for 5G mmWave Systems

Several hybrid beamforming methods exist for 5G millimeter wave MIMO systems which enhance spectral efficiency by using less hardware and lower power consumption relative to digital beamforming techniques. Some notable approaches include:

- **The MO-AltMin (Manifold Optimization Alternating Minimization):** employs a conjugate gradient algorithm to upgrade both analog and digital precoders through an iterative process but demands elevated computational complexity. It achieves high spectral efficiency [1].
- **SDR-Based Beamforming**: as stated in [1] "the semidefinite algorithm is implemented using the alternating minimization concept for the partially connected where each RF chain is connected to a subarray of antennas". This method delivers suboptimal results through semidefinite relaxation technology to produce approximate optimal hybrid precoders which decrease processing requirements [13]. The hybrid structure undergoes conversion into a non-convex configuration during the process that transforms it into a convex semi-definite program. This method delivers effective results when operated on small system dimensions yet generates poor outcomes with large dimensions.

- **Greedy Algorithms**: Selects beamforming vectors iteratively based on predefined selection criteria, balancing performance and complexity [14].

- **Spatially Sparse precoding (SSP) algorithm**: The orthogonal matching pursuit (OMP) technique is used to transform digital beamformer into a sparse approximation of hybrid beamforming. Due to heavy computing requirements this beamforming technique delivers high spectral efficiency results [15].

- **Alternating Minimization (AltMin)**: This beamforming technique to optimize the analog and digital components sequentially through separate optimization processes. The paper introduced PE-AltMin to extract phases with simplified complexity as well as MO-AltMin which applies manifold optimization and conjugate gradient descent methods in the analog space for optimization [1] and [16]. Chapter three includes a complete description of the implemented algorithm.

- Machine Learning-Based hybrid beamforming: A new approach in beamforming technology employs supervised learning along with reinforcement learning methods to anticipate directional settings and optimize beamformer control parameters. The system applies deep learning approaches which include conventional neural networks (CNNs) and recurrent neural networks (RNNs) to identify beamformers through channel or beam space attributes [17].

1.6.7 Complexity, Performance, and Implementation Considerations

The maximum spectral efficiency of fully digital precoders remains out of reach but have limitations of applying them to large antenna array systems. On the other hand, the hybrid beamforming methods like PE-AltMin and MO-AltMin deliver nearly optimal performance outcomes while requiring less complexity in their implementation.

The process of SDR and manifold optimization-based techniques requires several iterative algorithm steps to increase the system's processing workload as clearly shown in [1]. Greedy methods together with phase extraction methods simplify system implementation but result in minor performance degradation.

Hybrid beamforming systems mandate precise implementation of hardware components, especially regarding phase shifter networking alongside RF chain arrangement [7]. These considerations are covered in chapter six.

1.6.8 Survey of Low-Complexity Approaches for Hybrid Beamforming

Hybrid beamforming represents a practical solution that decreases RF chains in mmWave MIMO systems without sacrificing performance optimality to get near optimal efficiency. Standard or traditional hybrid beamforming approaches demonstrate high computational requirements that make them impractical for real-life deployment situations, especially for MIMO systems. Research groups developed various low complexity hybrid beamforming approaches to achieve good performance efficiency with a lower computational complexity. This section explores fundamental low-complexity approaches.

1) Phase Extraction Alternating Minimization (PE-AltMin)

This hybrid beamforming technique uses a designated method which extracts phase elements from the fully digital precoder to produce iterative updates on analog and digital precoder sections with minimal hybrid-to-optimal digital error outcomes. The proposed approach conducts this operation with substantially lower computational complexity than standard hybrid beamforming optimization methods do. PE-AltMin operates through a procedure which alternates the updates of analog and digital precoders which results in quick convergence rates with reduced computational overhead. This approach provides practical mmWave MIMO solution because of its simplicity combined with efficiency [1].

2) Low-complexity Hybrid Precoding (LCHP) Algorithm

The algorithm splits the fully digital precoder into two distinct analog and digital parts. Firstly, SVD analysis of the channel matrix provides the solution for optimal fully digital precoder. Secondly, the analog precoder is constructed by extracting the phases of the optimal fully digital precoder and the phase values are quantized to match the hardware constraints of the phase shifters. “The digital precoder also calculates its values using the regularized Least Squares (RLS)

optimization that minimizes the error between the optimal fully digital precoder and the analog precoder operation on the optimal precoder [1] [16]".

3) Successive Interference Cancellation (SIC)-Based Hybrid Precoding

Successive Interference Cancellation (SIC)-based hybrid precoding provides a low-complexity approach to address interference challenges so that maintain less computational processing. In this technique hybrid beamforming executed successively with each iteration removing previously processed signal interferences. Signal recovery becomes more efficient together with a decrease in matrix operation complexity. Wireless communication systems choose SIC-based methods as their preferred method because they work well in situations where interference becomes dominant [11].

4) Eigenvector-Based Hybrid Precoding

Eigenvector-based hybrid precoding leverages the eigen-decomposition of the channel matrix to drive an efficient hybrid beamforming solution. This technique approximates the optimal fully digital precoder by selecting the dominant eigenvectors of the channel matrix. By doing this, it reduces computational complexity while maintaining good spectral efficiency. Its shown in [1] that "Eigenvector-based methods are particularly useful in scenarios where channel state information (CSI) is readily available, as they provide a structured way to optimize the beamforming process without extensive iterative computations" [6].

5) Spatially Sparse precoding

This algorithm introduces a hybrid analog-digital precoding method that exploits channel sparsity in mmWave systems. Instead of using the fully digital precoding, which requires considerable amount of radio frequency chains, the approach reduces the amount of radio frequency chains while maintaining reliable performance. It formulates the precoding problem as an optimization problem that approximates the fully digital precoder while considering hardware constraints of mmWave systems. The algorithm uses an iterative approach to find near-optimal hybrid precoder while keeping RF hardware complexity low to overcome the computational challenges of implementing hybrid precoding in large-scale massive MIMO systems [15].

2 Chapter Two: System Model and Problem Formulation

2.1 Millimeter wave MIMO system Model

The millimeter waves have wavelengths (λ) ranging between 1-10 millimeters. This frequency band extends from 30–300 GHz corresponds to this specific wavelength. The millimeter wave falls under the Extremely High Frequency (EHF) band [4], classified by the International Telecommunication Union (ITU). Millimeter wave signal propagation characterized by:

- Atmospheric attenuation
- High free space path loss
- Diffuse reflections
- Limited penetration depth

Free Space Path Loss: A major barrier to direct line-of-sight transmission of the millimeter wave radio frequency is the significant free space path loss (FSPL). The free space path loss (FSPL) changes inversely in terms of square of wavelength and can be given by:

$$FSPL = \left(\frac{4\pi d}{\lambda}\right)^2 \tag{2.1}$$

$$FSPL(dB) = 20\log_{10}(d) + 20\log_{10}(f) + 92.45$$

Where d is separation distance between the base station and the user, and f is the frequency signal [6]. As can clearly be seen from the equation (2.1), 10X decreases in the wavelengths and results in 100X increase in the free space path loss. Thus, the attenuation at millimeter wavelength is many orders of magnitude higher than the attenuation of frequencies deployed for of this day's wireless communications.

The main hurdle in mmWave communications is atmospheric signal loss because oxygen and water vapor molecules diminish transmission strength at frequency bands [10]. The rough texture of surfaces at mmWave frequencies produces diffuse reflection by scattering signal energy which diverts away from a single reflection path.

This reduces the amount of reflected energy reaching the receiver, making mmWave transmissions highly susceptible to shadowing and requiring near-LOS conditions [7].

Moreover, Limited Penetration is another major drawback of mmWave communications. The small wavelengths of mmWave signals cause them to fail at passing through regular construction materials therefore requiring straight LOS communication paths [9].

Despite these challenges, mmWave systems offer high data rates due to their large available bandwidth. However, they suffer from poor link budgets because obstacles comparable in size to the wavelength of the signal cause significant blockage [2]. To address these challenges, several researchers are presenting different techniques, like

1. Directional Antennas: These improve signal delivery when a LOS communication path exists between the transmitter and receiver.
2. Dynamic beam steering through Multi-Input Multi-Output (MIMO) Beamforming makes use of channel statistics to improve both multiplexing gain and beamforming gain [1].

Advanced hybrid beamforming methods enable mmWave MIMO systems to handle many challenges and improve the operational prospects of high-frequency wireless communication for 5G and subsequent standards [8].

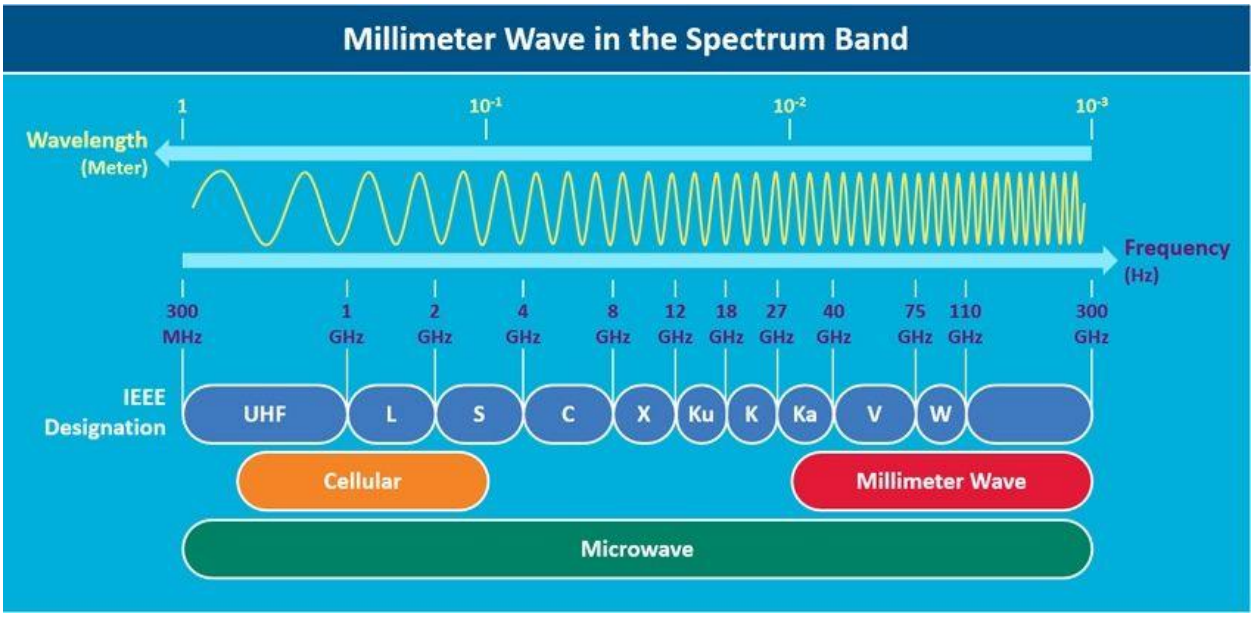


Figure 2.1: Millimeter wave frequency spectrum.

2.2 System Model

As it is shown in Figure *2.2* below, the figure illustrates the base station to user communication in a single cell multiple receiver/user millimeter wave massive multiple input multiple output OFDM architecture, a base station employee an N_t sized antenna configuration to provide K receivers with F narrow band subcarriers. For each subcarrier, BS offers N_s data streams for each user and the user have N_r antennas.

For the signal transmitter and each receiver, respectively, there are N_{RF}^{t} and N_{RF}^{r} possible radio frequency (RF) chains, which are constrained as $KN_s \leq N_{\mathrm{RF}}^{t} < N_t$ and $N_s \leq N_{\mathrm{RF}}^{r} < N_r$ [1].

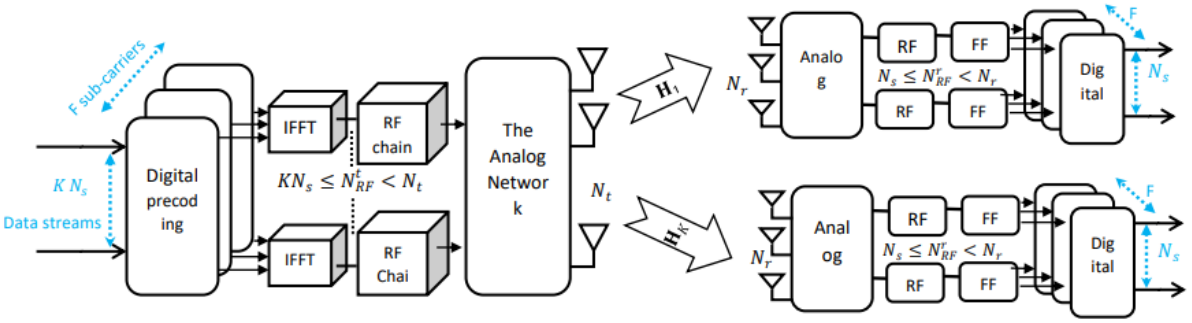


Figure 2.2: A multi-user millimeter wave MIMO system using OFDM, featuring a hybrid precoder based on a fully-connected phase shifter (FPS) design *[1]*.

Figure 2.2 represents that in the analog precoder model, dashed lines denote N_{RF}^{t} switches, and diagonal solid lines represent parallel signal transmissions.

The k-th user's received signal on the f-th subcarrier is provided as:

$$y_{(k,f)} = \mathbf{W}_{\mathrm{BB}_{(k,f)}}^{\mathrm{H}} \mathbf{W}_{\mathrm{RF}_k}^{\mathrm{H}} \left(\sqrt{\boldsymbol{\rho}}\mathbf{H}_{(k,f)} \sum_{j=1}^{K} \mathbf{F}_{\mathrm{RF}}\, \mathbf{F}_{\mathrm{BB}_{(j,f)}} s_{(j,f)} + \mathbf{n}_{(k,f)}\right) \tag{2.2}$$

$\sqrt{\rho}$ is the average received power; H stands for the channel matrix. The k-th user on the f-th subcarrier is represented using the subscript k, f, and the transmitted symbol vector $\mathrm{s}_{k,f} \in \mathbb{C}^{N_s}$ is such that $\mathbb{E}\left[s_{k,f} s_{k,f}^{H}\right] = \frac{1}{KN_sF}\mathbf{I}_{N_s}$. Digital baseband precoders are represented by $\mathbf{F}_{\mathrm{BB}_{k,f}} \in \mathbb{C}^{N_{\mathrm{RF}}^{t} \times N_s}$ and combiners by $\mathbf{W}_{\mathrm{BB}_{k,f}} \in \mathbb{C}^{N_{\mathrm{RF}}^{r} \times N_s}$.

The analog RF precoder, $\mathbf{F}_{\mathrm{RF}} \in \mathbb{C}^{N_t \times N_{\mathrm{RF}}^t}$ operates after the IFFT stage and is applied commonly across all subcarriers and users, since the digital baseband precoder already handles the mixing of signals intended for different users. In contrast, each user employs its own analog RF combiner, $\mathbf{W}_{\mathrm{RF}} \in \mathbb{C}^{N_r \times N_{\mathrm{RF}}^r}$, that functions independently on a per-subcarrier basis. The noise, $\mathbf{n}_{k,f} \in \mathbb{C}^{N_r}$ observed at the user side is represented by a matrix whose entries are modeled as independent complex Gaussian random variables with zero mean and unit variance $\mathcal{CN}(0, \sigma_n^2)$.

Assuming the transmitted symbols are Gaussian, the achievable sum rate on the f-th subcarrier can be expressed using the formula provided in in equation 2.3 as it is formulated using [1].

$$R_f = \sum_{k=1}^{K} \log\det\left(\mathbf{I}_{N_s} + \frac{\mathbf{1}}{KN_sF}\mathbf{W}_{k,f}^H \mathbf{H}_{k,f}\mathbf{F}_{k,f}\mathbf{F}_{k,f}^H\mathbf{H}_{k,f}^H\mathbf{W}_{k,f}\mathbf{\Omega}_{k,f}^{-\mathbf{1}}\right), \tag{2.3}$$

Where $\mathbf{F}_{k,f} = \mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}_{k,f}}$ and $\mathbf{W}_{k,f} = \mathbf{W}_{\mathrm{RF}_k}\mathbf{W}_{\mathrm{BB}_{k,f}}$ are the beamformer at the transmitter matrix and combiner matrix at the receiver, and $\mathbf{\Omega}_{k,f} = \mathbf{W}_{k,f}^H[\frac{\boldsymbol{\rho_k}}{KN_sF}\mathbf{H}_{k,f}(\sum_{j\neq k}\mathbf{F}_{j,f}\mathbf{F}_{j,f}^H)\mathbf{H}_{k,f}^H + \sigma_n^2\,\mathbf{I}_{N_r}]\mathbf{W}_{k,f}$ represents for the sum of the matrix of interference and noise.

2.3 Channel Model

Because of the significant free-space path loss at millimeter-wave frequencies, the MIMO channel, $\mathbf{H}_{k,f}$, between the base station and the k-th user on the f-th subcarrier is modeled using the Saleh-Valenzuela clustering approach, as described in references [1] [10] and [18].

The Saleh-Valenzuela model formulated the mm-wave channel matrix as

$$\begin{aligned}\mathbf{H}_{k,f} &= \sqrt{\frac{N_t N_r}{\rho_k N_{cl,k} N_{\mathrm{ray},k}}} \sum_{i=0}^{N_{cl,k}-1}\sum_{l=1}^{N_{\mathrm{ray},k}} \alpha_{il,k}\,\mathbf{a}_r(\phi_{il,k},\theta_{il,k})\mathbf{a}_t(\phi_{il,k},\theta_{il,k})^H \\ &= \gamma \sum_{i=0}^{N_{cl,k}-1}\sum_{l=1}^{N_{\mathrm{ray},k}} \alpha_{il,k}\,\mathbf{a}_r(\phi_{il,k},\theta_{il,k})\mathbf{a}_t(\phi_{il,k},\theta_{il,k})^H\end{aligned} \tag{2.4}$$

In this channel model, the factor γ acts as a normalization constant to ensure the total power is properly scaled. The parameters $N_{cl,k}$ and $N_{\mathrm{ray},k}$ simply tell us how many propagation clusters exist in the environment and how many individual rays (or multipath components) are present within each cluster, respectively.

For the k -th user, the symbol $\alpha_{il,k}$ captures the overall path loss — essentially, how much signal strength is lost as the wave travels from transmitter to receiver due to distance, obstacles, and other large-scale effects.

Within each cluster, l-th ray in the i-th cluster has its own complex gain, denoted by g_{n,r}. This gain accounts for small-scale fading effects, and it is modeled as a random variable that follows a specific statistical distribution (typically complex Gaussian in many mmWave or massive MIMO models).

On the antenna array side, , $\mathbf{a}_r(\theta_{il,k})$ and at $\mathbf{a}_t(\phi_{il,k})$ represents the receive array response vector, which describes how the receiving antenna array "sees" an incoming ray arriving from a particular direction. The angles $\theta_{il,k}$ and $\phi_{il,k}$ are the azimuth and elevation angles of arrival (AoA), respectively.

To make things concrete, we focus on a uniform square planar array (USPA), which is a common antenna configuration in modern wireless systems. This array consists of N antenna elements arranged in a square grid (for example, $\sqrt{N} \times \sqrt{N}$ elements along each side). Thanks to this regular geometry, we can explicitly write the array response vector for the l-th ray in the n-th cluster as it applies to the k-th user — the exact mathematical form is provided in [1].

$$\mathbf{a}(\phi_{il,k}, \theta_{il,k}) = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 \\ \vdots \\ e^{j\frac{2\pi}{\lambda} d(p\sin\phi_{il,k}\sin\theta_{il,k} + q\cos\theta_{il,k})} \\ \vdots \\ e^{j\frac{2\pi}{\lambda} d((\sqrt{N}-1)\sin\phi_{il,k}\sin\theta_{il,k} + (\sqrt{N}-1)\cos\theta_{il,k})} \end{bmatrix}^T \quad (2.5)$$

In this formulation, λ stands for the wavelength of the transmitted signal — a fundamental parameter that determines how the electromagnetic waves behave in space. The symbol d represents the uniform spacing between adjacent antennas in the array, typically chosen as half the wavelength ($d = \lambda/2$) to prevent grating lobes and ensure good performance, though the exact value can vary depending on the design.

The indices m and n correspond to the position of each antenna element within the two-dimensional grid of the square planar array. Specifically, m denotes the antenna's column index (along one horizontal direction), while n denotes its row index (along the vertical direction). Together, they uniquely identify every element in the $M = \sqrt{M} \times \sqrt{M}$ grid.

This particular channel model — which incorporates clustered multipath propagation, path loss, ray gains, and precise array steering through azimuth and elevation angles — will serve as the foundation for all our numerical simulations. By using the same realistic model across different algorithms, we can fairly compare their performance in terms of beamforming accuracy, user rates, or any other relevant metrics.

2.4 Hybrid beamforming

Beamforming and precoding are two words which have the same meaning and application and will be used interchangeably in this thesis paper. As shown Figure *2.2* above, the hybrid precoder is composed of two parts, which are the digital baseband beamformer and the analog RF beamformer. The comparison of hybrid beamformer to the traditional optimal baseband beamformer, the additional hardware in the hybrid RF-baseband beamformer counterpart is the analog elements in the RF domain, which is referred to as the analog architecture, which is the dominant factor to hardware complexities of the hybrid analog-digital beamformer for the millimeter wave massive MIMO systems.

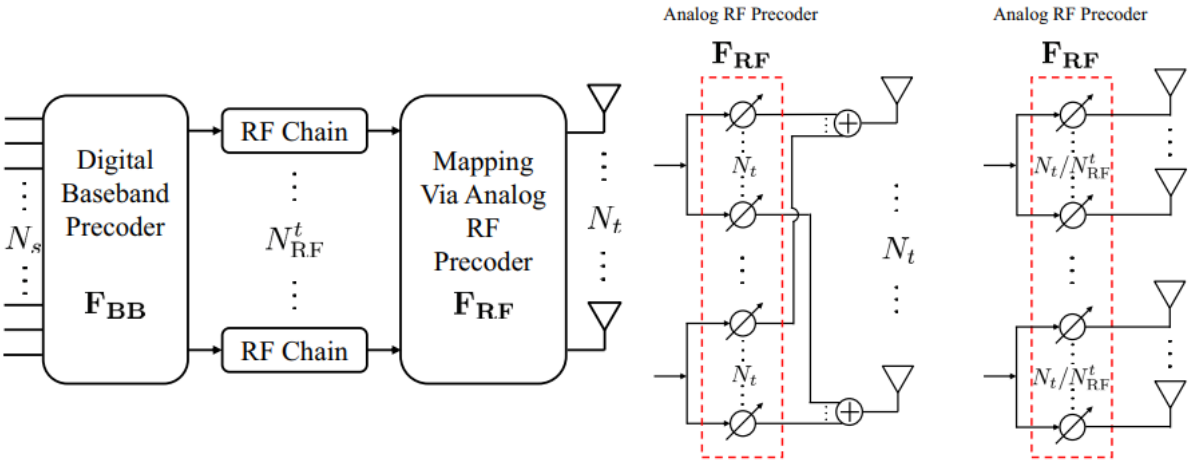
2.4.1 Mapping strategy and hardware Implementation

For mmWave MIMO systems with RF and baseband hybrid precoding, the mapping structure indicates the relations between the antenna elements and RF chains. Mapping structure choice is critical since it determines the design of hardware components and causes performance and implementation complexity differences.

The way RF chains are connected to antenna elements is mostly associated with mapping strategy, which in turn affects the total amount of hardware elements within the analog architecture of the system. There are different types of mapping in hybrid beamforming.

1. Fully connected mapping: For this design, all antenna elements are connected to all RF chains through networks of phase shifters. The architecture has relative advantages as it provides high beamforming gains, is flexible, and can emulate the full potential of fully digital precoding. On the contrary, this method requires many phase shifters and complex hardware and increased power consumption. This thesis focuses on this mapping technique both for the PE-AltMin and MO-AltMin algorithms.
2. Partially connected structure: Each RF chain corresponds to a set of antenna elements, which are sometimes arranged in a modular or subarray arrangement. With such an arrangement you will benefit from a simple design and energy efficiency, as well as reduced use of phase shifters and switches. However, this structure has the disadvantage of lower beamforming gain than the fully connected structure and poor performance when the RF-to-antenna connection decreases.
3. Dynamic or switch-based mapping: The RF chain-antenna link is based upon switch networks which provide the ability for flexible reconfiguration of connections. Dynamic mapping allows adaptation to channel conditions with flexibility and promises hardware efficiency. However, this entails the challenge of deploying speedy and optimized switches together with more complex control systems.
4. Lens array (Beam space) mapping: it uses a lens antenna array to focus energy into beams (beam space MIMO). Each RF chain corresponds to a beam rather than a physical antenna. Advantages of this method are that it reduces the dimension of analog beamforming and is suitable for sparse millimeter wave channels. The disadvantage is performance depending on beam alignment. [1].

The hardware implementation describes the hardware elements used and the connections made between each RF chain-antenna pair. The most widely used method identified in the literature is the single-phase shifter (SPS) architecture, which uses one phase shifter to treat and process each radio frequency elements in the antenna structure. As a result, SPS are fully connected and SPS partially connected structures are the most common types of analog network architecture, as depicted in Figure *2.3* below.



a) millimeter wave MIMO transmitter architecture

b) The Single phase shifter hybrid structure with complete connected elements

c) The single phase shifters hybrid architecture with a sub-connected element

Figure 2.3: Hybrid beamforming architecture *[1]*

2.4.2 Problem Formulation of 5G mmWave Hybrid Beamforming

There are various approaches to the optimization of 5G mmWave hybrid beamforming. Some of them are based on additional limitations and challenges like the unit modulus constraints on the analog beamformer to ease the system modeling in multiple receivers in a single carrier communication system, with the express goal of directly increasing the spectral efficiency in single-user systems through approximations and limits. Directly optimizing the analog digital hybrid beamformer with spectrum effectiveness as the objective equation in multiuser multicarrier systems is very challenging and unachievable because each user's spectral efficiency on individual subcarrier is integrated with one another via the common analog beamformer. Conversely, several studies have demonstrated that optimizing spectral efficiency may be achieved by decreasing the Euclidean distance between the hybrid and purely digital precoders [1] [15].

Therefore, the design goal for the low complexity algorithm as alternative objective function will be formulated [1] by:

$$\underset{\mathbf{F}_{\mathrm{RF}},\mathbf{F}_{\mathrm{BB}}}{\text{Minimize}} \; ||\mathbf{F}_{\mathrm{opt}} - \mathbf{F}_{\mathrm{RF}}\,\mathbf{F}_{\mathrm{BB}}||_{\mathrm{F}}^{2}, \qquad \text{Subject to} \begin{cases} \mathbf{F}_{\mathrm{RF}} \in \mathcal{A} \\ ||\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}||_{\mathrm{F}}^{2} \leq KN_sF \end{cases} \tag{2.6}$$

Where $\mathbf{F}_{\mathrm{opt}} = [\mathbf{F}_{\mathrm{opt}_{1,1}}, \dots, \mathbf{F}_{\mathrm{opt}_{k,f}}, \dots, \mathbf{F}_{\mathrm{opt}_{K,F}}] \in \mathbb{C}^{N_t\, x\, KN_sF}$ is the combined fully digital beamformer and $\mathbf{F}_{\mathrm{BB}} = \left[\mathbf{F}_{\mathrm{BB}_{1,1}}, \dots, \mathbf{F}_{\mathrm{BB}_{k,f}}, \dots, \mathbf{F}_{\mathrm{BB}_{K,F}}\right] \in \mathbb{C}^{N_{\mathrm{RF}}^{t}\, x\, KN_sF}$ is the concatenated baseband digital beamformer.

The transmit power constraint at the base station also represents another serious limitation to be considered. $\mathbf{F}_{\mathrm{RF}}$ represents the analog radio frequency beamformer, which is valid for all the K users, and F subcarriers and the selection thereof is constrained by a candidate set $\mathcal{A}$ as determined by numerous configurations of the analog network. The worst performance achievable by the best hybrid precoder is represented by the fully digital precoder, with the best hybrid performance being an ideal objective of closely matching the performance. As a result, the design problem can be easily approached by attempting to emulate the performance of the conventional optimum baseband (digital) beamformer with the help of the hybrid option. By using this representation, the PE-AltMin algorithm could be used irrespective of any fully digital precoder selected.

For point-to-point systems, optimization of fully digital precoding and combining for the f-th subcarrier requires the first N_s columns of $\mathbf{V}_f$ and $\mathbf{U}_f$ as they are unitary matrices resulting from the SVD of the channel which is given by that is $\mathbf{H}_f = \mathbf{U}_f\mathbf{\Sigma}_f\mathbf{V}_f$.

3 Chapter Three: Evaluation of Low complexity Hybrid Beamforming

3.1 An overview of PE-AltMin Algorithm

One practical way to handle this problem is to take advantage of the fact that the digital precoder is orthogonal. Thanks to this property, we can develop a fairly efficient computational approach to determine the analog precoder even when its elements are restricted to lying on the unit circle (constant-modulus / unit-modulus constraint).

The key insight is pretty neat: because of the orthogonality of the digital part, we can directly pull out the phase information needed for the analog precoder by comparing two different but equivalent beamforming solutions — specifically, the phases from the actual digital beamformer and the unconstrained optimal beamformer that would be ideal if there were no hardware restrictions [1].

It's shown using the simulation results in chapter six that, nevertheless this beamforming technique incurs some performance loss relative to the manifold optimization alternating minimization (MO-AltMin) method, it has significant computational complexity improvement.

3.2 Algorithm Description

3.2.1 Digital Baseband Beamforming Design

It's well known (as demonstrated in [1]) that the columns of the unconstrained optimal precoding matrix $\mathbf{F}_{\text{opt}}$ are mutually orthogonal. This property helps minimize interference among the spatially multiplexed data streams. Building directly on this orthogonality present in the unconstrained solution, we impose the same requirement on the digital precoding matrix — namely, that its columns remain orthogonal.

$$\mathbf{F}_{\text{BB}}^{H}\mathbf{F}_{\text{BB}} = \alpha\mathbf{F}_{\text{DD}}^{H}\alpha\mathbf{F}_{\text{DD}} = \alpha^{2}\mathbf{I}_{N_s}, \tag{3.1}$$

$\mathbf{F}_{\text{DD}}$ is also a unitary matrix with the size relative to $\mathbf{F}_{\text{BB}}$. Given that since the optimal digital beamforming structure is far from established it is self-evident and intriguing to investigate the hybrid precoder when the digital beamformer must be orthogonal. Notably, the orthogonal

constraint enables the analog precoder $\mathbf{F}_{\mathrm{RF}}$ to break out of the product dependency upon $\mathbf{F}_{\mathrm{BB}}$, dramatically simplifying the design of analog precoder.

3.2.2 Hybrid Beamforming Design

As defined in [1] the objective function in equation (*2.6*) may be further rewritten as, by replacing $\mathbf{F}_{\mathrm{BB}}$ with $\alpha\mathbf{F}_{\mathrm{DD}}$, follows

$$\begin{aligned} &||\mathbf{F}_{\mathrm{opt}} - \mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}||_{\mathrm{F}}^2 \\ &= \mathrm{tr}\left(\mathbf{F}_{\mathrm{opt}}^H\mathbf{F}_{\mathrm{opt}}\right) - \mathrm{tr}\left(\mathbf{F}_{\mathrm{opt}}^H\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\right) - \mathrm{tr}\left(\mathbf{F}_{\mathrm{BB}}^H\mathbf{F}_{\mathrm{RF}}^H\mathbf{F}_{\mathrm{opt}}\right) + \mathrm{tr}\left(\mathbf{F}_{\mathrm{BB}}^H\mathbf{F}_{\mathrm{RF}}^H\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\right) \\ &= ||\mathbf{F}_{\mathrm{opt}}||_{\mathrm{F}}^2 - 2\alpha\Re\mathrm{tr}\left(\mathbf{F}_{\mathrm{DD}}\mathbf{F}_{\mathrm{opt}}^H\mathbf{F}_{\mathrm{RF}}\right) + \alpha^2||\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{DD}}||_{\mathrm{F}}^2 \end{aligned} \tag{3.2}$$

Obviously, when $\alpha = \frac{\Re\mathrm{tr}\left(\mathbf{F}_{\mathrm{DD}}\mathbf{F}_{\mathrm{opt}}^H\mathbf{F}_{\mathrm{RF}}\right)}{||\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{DD}}||_{\mathrm{F}}^2}$, the objective function $||\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{DD}}||_F^2$ in (8) bears the minimum value which can be given as $||\mathbf{F}_{\mathrm{opt}}||_{\mathrm{F}}^2 - \frac{\left\{\Re\mathrm{tr}\left(\mathbf{F}_{\mathrm{DD}}\mathbf{F}_{\mathrm{opt}}^H\mathbf{F}_{\mathrm{RF}}\right)\right\}^2}{||\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{DD}}||_{\mathrm{F}}^2}$.

Additionally, the upper bound for the Frobenius norm square $||\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{DD}}||_{\mathrm{F}}^2$ is given as below.

$$\begin{aligned} ||\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{DD}}||_{\mathrm{F}}^2 &= \mathrm{tr}\ \left(\mathbf{F}_{\mathrm{DD}}^H\mathbf{F}_{\mathrm{RF}}^H\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{DD}}\right) \\ &= \mathrm{tr}\left\{\begin{pmatrix}\mathbf{I}_{N_s} & \\ & 0\end{pmatrix}\mathbf{K}^H\mathbf{F}_{\mathrm{RF}}^H\mathbf{F}_{\mathrm{RF}}\mathbf{K}\right\} \\ &\leq tr\{\mathbf{K}^H\mathbf{F}_{\mathrm{RF}}^H\mathbf{F}_{\mathrm{RF}}\mathbf{K}\} \\ &= ||\mathbf{F}_{\mathrm{RF}}||_{\mathrm{F}}^2 \end{aligned} \tag{3.3}$$

Where $\mathbf{F}_{\mathrm{DD}}\mathbf{F}_{\mathrm{DD}}^H = \mathbf{K}\begin{pmatrix}\mathbf{I}_{N_s} & \\ & 0\end{pmatrix}\mathbf{K}^H$ is the singular value decomposition of $\mathbf{F}_{\mathrm{DD}}\mathbf{F}_{\mathrm{DD}}^H$. Equality, therefore, occurs when $N_{RF}^t = N_s$, that is $\mathbf{F}_{\mathrm{DD}}$ is a square matrix. Hence, the objective optimizing model in Equation (*2.6*) is upper bounded by $||\mathbf{F}_{\mathrm{opt}}||^2{}_F - \frac{\left\{\Re\mathrm{tr}\left(\mathbf{F}_{\mathrm{DD}}\mathbf{F}_{\mathrm{opt}}^{\mathrm{H}}\mathbf{F}_{\mathrm{RF}}\right)\right\}^2}{||\mathbf{F}_{\mathrm{RF}}||_{\mathrm{F}}^2}$. To remove the product term, one approach is to incorporate an additional constant, $\left(\frac{1}{2||\mathbf{F}_{\mathrm{RF}}||_{\mathrm{F}}^2} - 1\right)||\mathbf{F}_{\mathrm{opt}}||_{\mathrm{F}}^2 + \frac{1}{2}$, into the bound and then scale the whole expression by a positive constant factor $2||\mathbf{F}_{\mathrm{RF}}||_{\mathrm{F}}^2$. To get rid of the

product term entirely, one straightforward option is to include an additional constant in the bound and scale the resulting expression by a positive constant factor.

Hence, we have,

$$\begin{aligned} &||\mathbf{F}_{\text{opt}}||_{\text{F}}^2 - 2\Re\text{tr}\left(\mathbf{F}_{\text{DD}}\mathbf{F}_{\text{opt}}^H\mathbf{F}_{\text{RF}}\right) + ||\mathbf{F}_{\text{RF}}||_{\text{F}}^2 \\ &= \text{tr}\left(\mathbf{F}_{\text{RF}}^H\mathbf{F}_{\text{RF}}\right) - 2\Re\text{tr}\left(\mathbf{F}_{\text{DD}}\mathbf{F}_{\text{opt}}^H\mathbf{F}_{\text{RF}}\right) + \text{tr}\left(\mathbf{F}_{\text{DD}}\mathbf{F}_{\text{opt}}^H\mathbf{F}_{\text{opt}}\mathbf{F}_{\text{DD}}^H\right) \\ &= ||\mathbf{F}_{\text{opt}}\mathbf{F}_{\text{DD}}^H - \mathbf{F}_{\text{RF}}||_{\text{F}}^2 \end{aligned} \tag{3.4}$$

Direct minimization of the objective function in Equation (3.2) will still lead to high computational complexity, and hence we are going to use an upper bound of Equation (3.4) as an objective function instead of the original one. Besides, one can temporarily relieve the transmit power limitation after updating the hybrids precoders to satisfy the transmit power constraint with the normalization. Consequently, the hybrid beamformer design issue can be expressed as follows using Equation (3.4) as the target function of interest:

$$\begin{aligned} &\underset{\mathbf{F}_{\text{RF}},\,\mathbf{F}_{\text{DD}}}{\text{Minimize}} \; ||\mathbf{F}_{\text{opt}}\mathbf{F}_{\text{DD}}^H - \mathbf{F}_{\text{RF}}||_{\text{F}}^2 \\ &\text{Subject to} \begin{cases} |\mathbf{F}_{\text{RF}}(i,j)| = 1, \forall i,j \\ \mathbf{F}_{\text{DD}}^H\mathbf{F}_{\text{DD}} = \mathbf{I}_{N_s} \end{cases} \end{aligned} \tag{3.5}$$

According to the problem formulation (3.5), we simply need to look for a unitary precoding matrix $\mathbf{F}_{\text{DD}}$ to obtain a corresponding matrix $\mathbf{F}_{\text{BB}}$ with orthogonal columns.

The target function in equation (3.5) simplifies substantially the analog precoder implementation if alternating minimization can be used. More precisely, because it removes the product form with $\mathbf{F}_{\text{BB}}$, the matrix $\mathbf{F}_{\text{RF}}$ has a closed-form solution.

$$\arg\left(\mathbf{F}_{\mathrm{RF}}\right) = \arg\left(\mathbf{F}_{\mathrm{opt}}\mathbf{F}_{\mathrm{DD}}^{H}\right), \tag{3.6}$$

$\arg(\mathbf{A})$ returns a matrix containing the argument or phase of each entry in **A.** The closed-form solution presented above in equation (3.6), that alternatively can be seen as the projection onto the set using the standard Euclidean projection of $\mathbf{F}_{\mathrm{opt}}\mathbf{F}_{\mathrm{DD}}^{H}$ onto the realizable $\mathcal{A}_{\mathrm{f}}$ (set of feasible) analog beamforming weight vectors, thus demonstrates that the phases of an equivalent beamformer $\mathbf{F}_{\mathrm{opt}}\mathbf{F}_{\mathrm{DD}}^{H}$ can be derived from the phases of an analog beamformer $\mathbf{F}_{\mathrm{RF}}$.

We attempt to develop a digital precoder that solves the following problem by treating $\mathbf{F}_{\mathrm{RF}}$ as a fixed parameter for the digital precoder design.

$$\begin{aligned} &\underset{\mathbf{F}_{\mathrm{DD}}}{\text{Minimize}} \quad ||\mathbf{F}_{\mathrm{opt}}\mathbf{F}_{\mathrm{DD}}^{H} - \mathbf{F}_{\mathrm{RF}}||_{\mathrm{F}}^{2} \\ &\text{Subject to} \quad \mathbf{F}_{\mathrm{DD}}^{H}\mathbf{F}_{\mathrm{DD}} = \mathbf{I}_{N_s} \end{aligned} \tag{3.7}$$

Since the optimization problem in Equation (3.7) involves only a single variable $\mathbf{F}_{\mathrm{DD}}$, it can be rewritten as, as it is formulated in [1]

$$\begin{aligned} &\underset{\mathbf{F}_{\mathrm{DD}}}{\text{Maximize}} \quad \Re\mathrm{tr}\left(\mathbf{F}_{\mathrm{DD}}\mathbf{F}_{\mathrm{opt}}^{H}\mathbf{F}_{\mathrm{RF}}\right) \\ &\text{Subject to} \quad \mathbf{F}_{\mathrm{DD}}^{H}\mathbf{F}_{\mathrm{DD}} = \mathbf{I}_{N_s} \end{aligned} \tag{3.8}$$

In accordance with the dual norm definition, we have

$$\begin{aligned} \Re\mathrm{tr}\left(\mathbf{F}_{\mathrm{DD}}\mathbf{F}_{\mathrm{opt}}^{\mathbf{H}}\mathbf{F}_{\mathrm{RF}}\right) &\leq \left|\mathrm{tr}\left(\mathbf{F}_{\mathrm{DD}}\mathbf{F}_{\mathrm{opt}}^{H}\mathbf{F}_{\mathrm{RF}}\right)\right| \\ &\overset{(a)}{\leq} \|\mathbf{F}_{\mathrm{DD}}^{H}\|_{\infty} \cdot \left\|\mathbf{F}_{\mathrm{opt}}^{H}\mathbf{F}_{\mathrm{RF}}\right\|_{1} \\ &= \left\|\mathbf{F}_{\mathrm{opt}}^{H}\mathbf{F}_{\mathrm{RF}}\right\|_{1} \\ &= \sum_{i=1}^{N_s} \sigma_i \end{aligned} \tag{3.9}$$

(a) in equation (3.10) and presented clearly in [1], it is based on "the Hölder's inequality, $\|\cdot\|_\infty$ and $\|\cdot\|_1$ are the abbreviations of the infinite and one Schatten norm" [19]. There can be no equality until the following conditions are satisfied

$$\mathbf{F}_{\mathrm{DD}} = \mathbf{V}_1\mathbf{U}^H \tag{3.10}$$

$\mathbf{U\Sigma V}^H = \mathbf{USV}_1^H$, and $\mathbf{S}$ is a matrix that only has numbers on the main diagonal and zeros everywhere else whose entries are the first N_s nonzero singular values $\sigma_1, \sigma_2, \ldots, \sigma_{N_s}$, and the above expression is the SVD of $\mathbf{F}_{\mathrm{opt}}^H\mathbf{F}_{\mathrm{RF}}$.

PE-AltMin Algorithm: low complexity algorithm for fully connected strategies

---

Input: $\mathbf{F}_{\mathrm{opt}}$

1. Set $k = 0$, and Construct $\mathbf{F}_{\mathrm{RF}}{}^{(0)}$ with random phases;
2. **Repeat**
3. Fix $\mathbf{F}_{\mathrm{RF}}{}^{(k)}$, compute the SVD: $\mathbf{F}_{\mathrm{opt}}^H\mathbf{F}_{RF}{}^{(k)} = \mathrm{U}^{(k)}\mathrm{S}^{(K)}\mathrm{V}_1{}^{(k)^H}$;
4. $\mathbf{F}_{\mathrm{DD}}{}^{(k)} = \mathbf{V}_1{}^{(k)}\mathbf{U}^{(k)^H}$;
5. Fix $\mathbf{F}_{\mathrm{DD}}{}^{(k)}$, and compute arg $\{\mathbf{F}_{\mathrm{RF}}{}^{(k+1)}\} = \arg(\mathbf{F}_{\mathrm{opt}}\,\mathbf{F}_{\mathrm{DD}}{}^{(k)^H})$;
6. $k \leftarrow k + 1$;
7. **Until** the algorithm stopping rule is activated;
8. For the baseband-digital beamformer at the transmitter side, apply unit normalization, $\hat{\mathbf{F}}_{\mathrm{BB}} = \frac{\sqrt{N_s}}{\|\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{DD}}\|_F}\mathbf{F}_{\mathrm{DD}}$.

3.3 Computation Complexity Analysis of PE-AltMin

In this section, we will now see the computational complexity of the phase extraction alternating minimization (PE-AltMin) algorithm. As is stated above, the algorithm's initialization starts by inputting the optimal fully digital beamformer, $\mathbf{F}_{\mathrm{opt}}$. Let us step by step compute the complexities of each step as follows:

- Set $k = 0$, and Construct $\mathbf{F}_{\mathrm{RF}}{}^{(0)}$ with random phases.

At this point, a random matrix of size of $N_{\mathrm{t}} \times N_{\mathrm{RF}}$ is generated and the complexity at this simple initialization will be:

$$O(N_{\mathrm{t}} N_{\mathrm{RF}}) \tag{3.11}$$

- Fix $\mathbf{F}_{\mathrm{RF}}{}^{(k)}$, compute the SVD: $\mathbf{F}_{\mathrm{opt}}^{H} \mathbf{F}_{RF}{}^{(k)} = \mathrm{U}^{(k)} \mathrm{S}^{(K)} \mathrm{V}_1{}^{(k)^H}$;

This step involves the computation of the singular value decomposition (SVD) of $\mathbf{F}_{\mathrm{opt}}^{H} \mathbf{F}_{RF}^{(k)}$.

$$\mathbf{F}_{\mathrm{opt}}^{H} \mathbf{F}_{RF}^{(k)} = \mathbf{U}^{(k)} \mathbf{S}^{(k)} \mathbf{V}_1^{(k)H}$$

The multiplication of $\mathbf{F}_{\mathrm{opt}}^{H} \mathbf{F}_{RF}^{(k)}$ involves the multiplication of matrices of $\mathbf{F}_{\mathrm{opt}}^{H}$ size $(N_{\mathrm{s}} \times N_{\mathrm{t}})$ with $\mathrm{F}_{\mathrm{RF}}^{(k)}$ (size $N_{\mathrm{t}} \times N_{\mathrm{RF}}$) has $N_{\mathrm{s}} \times N_{\mathrm{RF}}$ matrix with a complexity of

$$O(N_s N_t N_{RF}) \tag{3.12}$$

The complexity of SVD computation on the other hand for an $N_{\mathrm{t}} \times N_{\mathrm{RF}}$ matrix is $O(N_s N_{RF}^2)$. This is because for the computational complexity of the SVD of an $m \times n$ matrix, assuming $m \geq n$, is given by $O(mn^2)$. Applying this in our case $m = N_{\mathrm{RF}}$ & $n = N_{\mathrm{s}}$

The total complexity for this step will be the sum of equation (3.11) & (3.12).

$$O(N_t N_s N_{RF}) + O(N_r N_{RF}^2) \tag{3.13}$$

- Compute $\mathbf{F}_{\mathrm{DD}}{}^{(k)} = \mathbf{V}_1{}^{(k)} \mathbf{U}^{(k)^H}$.

At this point, $\mathbf{V}_1{}^{(k)}$ is $N_{\mathrm{RF}} \times N_{\mathrm{s}}$ and $\mathbf{U}^{(k)^H}$ is $N_{\mathrm{s}} \times N_{\mathrm{s}}$. This matrix multiplication has a cost of

$$O(N_{RF} N_s^2) \tag{3.14}$$

- Fix $\mathbf{F}_{\mathrm{DD}}{}^{(k)}$, and compute $\arg\{\mathbf{F}_{\mathrm{RF}}{}^{(k+1)}\} = \arg(\mathbf{F}_{\mathrm{opt}} \mathbf{F}_{\mathrm{DD}}{}^{(k)^H})$;

$\mathbf{F}_{\mathrm{opt}}$ has size of $\mathrm{N_t} \times \mathrm{N}_s$ and $F_{DD}^{(k)^H}$ has $\mathrm{N_s} \times \mathrm{N}_{RF}$.

Therefore, this matrix multiplication has a complexity of: $O(N_t N_s N_{RF})$. On the other hand, extracting phase information using element wise operation has a complexity of $O(N_t N_{RF})$. Hence, the total complexity for the digital beamformer has a complexity of

$$O(N_s N_t N_{RF}) \tag{3.15}$$

- For the baseband-digital beamformer at the transmitter end, scale to unit norm $\hat{\mathbf{F}}_{\mathrm{BB}} = \frac{\sqrt{N_s}}{\|\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{DD}}\|_F}\mathbf{F}_{\mathrm{DD}}$.

Computing the complexity for the Frobenius norm of $\|\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{DD}}\|_{\mathrm{F}}$. This step is divided into two parts: matrix multiplication and Frobenius norm. the matrix multiplication has complexity of $O(\mathrm{N_t N_{RF} N_s})$ and Frobenius norm has $O(N_t N_s)$. The element wise division on the other hand has a complexity of $O(N_{RF} N_s)$.

The total complexity for the normalization:

$$O(N_t N_{RF} N_s) \tag{3.16}$$

Finally, let us sum up all the equations from (3.13) to (3.16) together to get the overall complexity of the PE-AltMin algorithm:

$$O(N_t N_{RF} N_s) + O(N_s N_t N_{RF}) + O(N_s N_{RF}^2) + O(N_{RF} N_s^2) \tag{3.17}$$

Since $N_s \leq N_{RF} \leq N_t$, the most dominant term will be:

$$O(N_t N_{RF} N_s) + O(N_s N_{RF}^2) \tag{3.18}$$

For I iterations the overall complexity for the PE-AltMin algorithm is updated as:

$$O(I(N_t N_{RF} N_s) + O(N_s N_{RF}^2)) \tag{3.19}$$

4 Chapter Four: Comparative Study of Hybrid Beamforming Algorithms

4.1 Fully Digital Beamforming

4.1.1 Introduction

The traditional and standard technology for multi-antenna systems uses fully digital beamforming because each antenna element requires its individual RF chain connection. It lets users fully control the amplitude and phase of signals at each antenna to achieve accurate spatial multiplexing alongside interference suppression capability. The expense, alongside significant power usage associated with individual RF chain deployment for every antenna, prohibits their practical use in massive mmWave MIMO systems.

4.1.2 System Model

Let us consider a single user point to point system with N_t denoting number of transmit antennas, N_r receiving antennas and N_s data streams where $N_s \leq \min(N_t, N_r)$.

The channel is model as

$$y = \mathbf{H}\mathbf{F}_{\text{opt}}s + \mathbf{n} \tag{4.1}$$

Where:

- $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ represents a flat-fading MIMO channel matrix.
- $\mathbf{F}_{\text{opt}} \in \mathbb{C}^{N_r \times N_t}$ is a fully digital precoder.
- $s \in \mathbb{C}^{N_s \times 1}$ represents vector of transmitted data symbols with $\mathbb{E}[\mathbf{s}\mathbf{s}^H] = \frac{\mathbf{1}}{N_s}\mathbf{I}_{N_s}$
- $y \in \mathbb{C}^{N_r \times 1}$ is the received signal and
- $\mathbf{n} \in \mathbb{C}^{N_r \times 1}$ is the noise vector modeled as $CN(0, \sigma^2\mathbf{I})$

4.1.3 Precoder design

To achieve optimal capacity, the precoder is designed using the right singular vectors of the channel matrix are used.

$$\mathbf{H} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{H} \tag{4.2}$$

- Where is $\mathbf{U} \in \mathbb{C}^{N_r \times N_r}$ the left singular vectors.
- $\mathbf{V} \in \mathbb{C}^{N_t \times N_t}$ are the right-handed "singular vectors" or noninvertible matrix.
- $\boldsymbol{\Sigma}$ is the diagonal matrix of the "singular values"

So, the first N_s singular vectors are selected for the beamforming, such that

$$\mathbf{F}_{\text{opt}} = \mathbf{V1}, \mathbf{W}_{\text{opt}} = \mathbf{U1} \tag{4.3}$$

To manage the power constraint, the precoder is normalized.

$$\hat{\mathbf{F}}_{\text{opt}} = \sqrt{\frac{N_s}{\left\|\mathbf{F}_{\text{opt}}\right\|_F^2}}\mathbf{F}_{\text{opt}} \tag{4.4}$$

4.2 MO-AltMin Based Beamforming

A fully-connected mapping is often applied to mmWave MIMO systems, where individual radio frequency chain is directly utilized to all antenna structures. The mapping structure forces each entry of the analog beamforming matrix to lie on the unit circle (i.e., have magnitude 1). This per-entry constraint corresponds to a Riemannian manifold.

As clearly presented in [1] and [20] the Frobenius norm in Equation (2.1) can be set precisely zero in the event that $N_{\text{RF}}^{t} \geq 2N_s$That is, under this special condition, hybrid precoders attain the performance of the fully digital precoder and are equivalent to the conventional optimum precoders obtained in [20].

Thus, we will concentrate on the region where $N_s \leq N_{\text{RF}}^{t} < 2N_s$.

4.2.1 Digital Baseband Beamformer Design

Let us first consider designing the digital beamformer F_{BB} with a fixed analog beamformer F_{RF}. Therefore, problem formulation equation (3.5) as given by [1]can be redefined as

$$\underset{\mathbf{F}_{BB}}{\text{Minimize}} \quad \left\|\mathbf{F}_{opt} - \mathbf{F}_{RF}\,\mathbf{F}_{BB}\right\|_F, \tag{4.5}$$

Which has a widely known least squares solution presented as

$$\mathbf{F}_{BB} = \mathbf{F}_{RF}^{\dagger}\mathbf{F}_{opt} \tag{4.6}$$

4.2.2 Analog RF Beamformer Design through Manifold Optimization

With full connectivity architecture, the feasible set $\mathcal{A}$ of the analog beamformer in equation (2.1) can be particularly determined $\mathcal{A}_f$ as $|F_{RF}(i,j)| = 1$ elementwise due to each RF chain is connecting to all antenna structures. In the step below, we are going to fix $\mathbf{F}_{BB}$ and find an analog precoder that optimizes the following problem. In this step, since the square of the Frobenius norm will not only make the problem objective function quadratic in nature but also make it smooth, this will not affect the solution [1].

$$\begin{aligned} &\underset{\mathbf{F}_{RF}}{\text{Minimize}} \quad ||\mathbf{F}_{opt} - \mathbf{F}_{RF}\,F_{BB}||_F^2, \\ &\text{Subject to } |F_{RF}(i,j)| = 1, \forall i,j. \end{aligned} \tag{4.7}$$

The major barriers to the problem we have here are the imposed unit modulus constraints that are inherently non-convex. In the next section, there is an effective manifold optimization algorithm for finding the near-optimal solution of the problem in Equation (4.7) [1]. A manifold $\mathcal{M}$, as depicted in Figure 4.1, is a topological space that is like the Euclidean space in the vicinity of each of its points. In the other case, each point of a manifold has its neighborhood, which is every point has a neighborhood homeomorphic to an open subset in the Euclidean space.

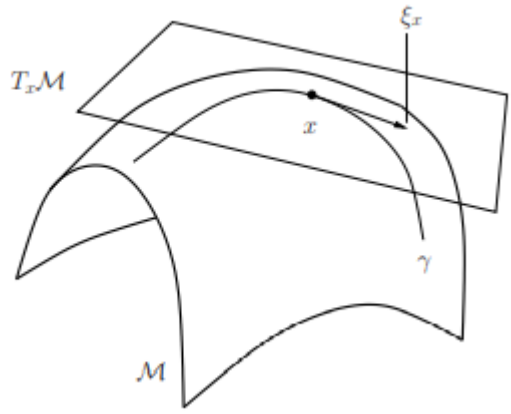


Figure 4.1: The Tangent space and tangent vector of a Riemannian manifold

The tangent space $T_x\mathcal{M}$ at a randomly mapped point x on the manifold $\mathcal{M}$ is formed by all the tangent vectors ξ_x that arise from smooth curves γpassing through x. In the majority of practical applications, the manifolds under consideration belong to an important subclass of topological manifolds known as Riemannian manifolds.

A Riemannian manifold comes equipped with an inner product defined on each tangent space $T_x\mathcal{M}$; this additional structure is called the Riemannian metric. The metric makes it possible to rigorously define lengths of curves, distances between points, and angles between tangent vectors on the manifold. Thanks to the Riemannian metric, one can also carry out differential calculus (including gradients, Hessians, covariant derivatives, and curvature computations) in a geometrically meaningful way on the manifold [1].

It is possible to define the gradients of cost functions based on the detailed geometry of the Riemannian manifold. Optimization with Riemannian manifolds can, however, locally be rewritten as optimization on a Euclidean space with smooth constraints. Consequently, on the stated Riemannian manifolds, one can find a well-developed conjugate gradient algorithm in the context of the Euclidean spaces. The following will provide a quick introduction to equivalent.

We first applied the Euclidean metric to the complex plane $\mathbb{C}$.

$$\langle x_2, x_2 \rangle = \Re\{x_1^* x_2\}, \tag{4.8}$$

As it is clearly presented in [1], it is the same as considering $\mathbb{C}$ as the $\mathbb{R}^2$ "with the canonical inner product". Then, we can re-write the complex circle in the form of

$$\mathcal{M}_{\mathrm{cc}} = \{x \in \mathbb{C}: x^* x = 1\} \tag{4.9}$$

The tangent vectors describe directions in which a particular point x on the manifold $\mathcal{M}_{\mathrm{cc}}$ can travel. As a result, the tangent space at $x \in \mathcal{M}_{\mathrm{cc}}$ may be expressed as

$$T_x\mathcal{M}_{\mathrm{cc}} = \{z \in \mathbb{C}: z^* x + x^* z = 2\langle x, z\rangle = 0\} \tag{4.10}$$

If $m = N_t N_{\mathrm{RF}}^t$, then "the vector $\mathbf{x} = \mathrm{vec}(\mathrm{F_{RF}})$ creates a complex circle manifold $\mathcal{M}_{\mathrm{cc}}^m = \{\mathbf{x} \in \mathbb{C}^m: |\mathbf{x}_1| = |\mathbf{x}_2| = \cdots = |\mathbf{x}_{\mathrm{m}}| = 1\}$. In the complex plane, which is a Riemannian sub-manifold of $\mathbb{C}^m$ with the product geometry, the search space of the optimization problem in Equation (31) is therefore over a product of m circles. Thus, [21] [1] may be used to represent the tangent space at a given point $x \in \mathcal{M}_{\mathrm{cc}}^m$:"

$$T_x\mathcal{M}_{\mathrm{cc}}^m = \{\mathbf{z} = \mathbb{C}^m: \Re\{\mathbf{z} \circ \mathbf{x}^*\} = \mathbf{0}_m\} \tag{4.11}$$

The direction of the largest drop of a function is represented by one of the tangent vectors connected to the negative Riemannian gradient, which is analogous to the Euclidean space. Authors in [1] states that "since the complex circle manifold $\mathcal{M}_{\mathrm{cc}}^m$ is a Riemannian submanifold of $\mathbb{C}^m$ [22], the Riemannian gradient at x is a tangent vector $\mathrm{grad}\, f(\boldsymbol{x})$ given by the orthogonal projection of the Euclidean gradient $\nabla f(\boldsymbol{x})$ onto the tangent space $T_x\mathcal{M}_{\mathrm{cc}}$".

$$\begin{aligned} \mathrm{grad}\, f(\mathbf{x}) &= \mathrm{Proj_x} \nabla f(\mathbf{x}) \\ &= \nabla f(\mathbf{x}) - \Re\{\nabla f(\mathbf{x}) \circ \mathbf{x}^*\} \circ \mathbf{x} \end{aligned} \tag{4.12}$$

In Equation (4.7), the cost function's Euclidean gradient is.

$$\nabla f(\mathbf{x}) = -2\left(\mathbf{F}_{\mathrm{BB}}^* \otimes \mathbf{I}_{N_t}\right)\left[\mathrm{vec}\left(\mathbf{F}_{\mathrm{opt}}\right) - \left(\mathbf{F}_{\mathrm{BB}}^T \otimes \mathbf{I}_{N_t}\right)\mathbf{x}\right] \tag{4.13}$$

To solve this Euclidean gradient, some complex-valued matrix derivative methods, there found more in [22].

"Retraction", which translates a point considered as a vector from the "tangent space onto the manifold itself", is another essential component of manifold optimization. When traveling along a tangent vector, it establishes the destination on the manifold. The retraction $\alpha\mathbf{d}$ of a tangent vector at point $\mathbf{x} \in \mathcal{M}_{cc}^{m}$ may be represented as [1]:

$$\begin{aligned} &\text{Retr}_{\mathbf{x}}: T_x\mathcal{M}_{cc}^{m} \rightarrow \mathcal{M}_{cc}^{m}: \\ &\alpha\mathbf{d} \mapsto \text{Retr}_{\mathbf{x}}(\alpha\mathbf{d}) = \text{vec}\left[\frac{(\mathbf{x}+\alpha\mathbf{d})_i}{|(\mathbf{x}+\alpha\mathbf{d})_i|}\right]. \end{aligned} \tag{4.14}$$

Algorithm 1: Applying Manifold Optimization, Conjugate Gradient Algorithm for Analog Precoding [1]

Input: $\mathbf{F}_{\text{opt}}, \mathbf{F}_{\text{BB}}, \mathbf{x}_0 \in \mathcal{M}_{cc}^{m}$

1. $\mathbf{d}_0 = -\text{grad}\, f(\mathbf{x}_0)$ and $k = 0$;
2. **Repeat**
3. "Choose Armijo backtracking line search step size α_k";
4. "Find the next point $\mathbf{x}_{k+1}$ using retraction in Equation (4.14): $\mathbf{x}_{k+1} = \text{Retr}_{\mathbf{x}_k}(\alpha_k\mathbf{d}_k)$";
5. Determine Riemannian gradient $\mathbf{g}_{k+1} = \text{grad}\, f(\mathbf{x}_{k+1})$ according to Equations (4.12) and (4.13);
6. "Determine the vector movements form $\mathbf{x}_k$ to $\mathbf{x}_{k+1}$ of $\mathbf{g}_k^+$ and $\mathbf{d}_k^+$ of gradient $\mathbf{g}_k$ and conjugate direction $\mathbf{d}_k$";
7. Choose Polak-Ribiere parameter $\mathbf{d}_{k+1} = -\mathbf{g}_{k+1} + \beta_{k+1}\mathbf{d}_k^+$;
8. $k \leftarrow k + 1$;
9. **Until** a condition for ending triggers;

Xianghao et. al showed that "along with the tangent space, a conjugate gradient approach based on a line search, a traditional algorithm in the Euclidean space, can be created to build the analog precoder using the Riemannian gradient and retraction of the complex circle manifold" $\mathcal{M}_{cc}^{m}$, as seen in the methodology above.

The Polak-Ribiere parameter and the well-known "Armijo backtracking line search step" are used in the algorithm to ensure that the desired function is not rising with each iteration [23]. Furthermore, the reason we introduce a "transport" map between tangent spaces is that operations in steps 7 and 8 act on tangent vectors residing in two separate tangent spaces, which are not the same vector space and therefore have no canonical way to interact directly, $T_{x_k}\mathcal{M}_{\text{cc}}^{m}$ and $T_{x_{k+1}}\mathcal{M}_{\text{cc}}^{m}$, that could not be added together straightforward. A tangent vector d transfer from $\mathbf{x}_k$ to $\mathbf{x}_{k+1}$ may be expressed as

$$\begin{aligned} \text{Transp}_{\mathbf{x}_k}: T_x\mathcal{M}_{\text{cc}}^{m} \rightarrow T_{x_{k+1}}\mathcal{M}_{\text{cc}}^{m}: \\ \mathbf{d} \mapsto \mathbf{d} - \Re\{\mathbf{d} \circ \mathbf{x}_{k+1}^{*}\} \circ \mathbf{x}_{k+1}, \end{aligned} \tag{4.15}$$

This is accomplished in Step 6.

The Manifold optimization algorithm is guaranteed for coverages to the critical point, which is the point where the gradient of the desired function is zero.

4.2.3 Hybrid Beamformer Design

Having the Manifold optimization algorithm, "the hybrid beamformer design through alternating minimization for the fully connected mapping is described in the MO-AltMin Algorithm by solving problems of Equations (4.5) and (4.7) iteratively [1]. To satisfy the power constraint in Equation (3.5), we normalize $\mathbf{F}_{\text{BB}}$ by a factor of $\frac{\sqrt{N_s}}{\|\mathbf{F}_{\text{RF}}\,\mathbf{F}_{\text{BB}}\|_F}$ at Step 7".

MO-AltMin Algorithm: "Manifold Optimization Based Hybrid Precoding for the Fully connected mapping [1]"

Input: $\mathbf{F}_{\text{opt}}$

1. "Construct $\mathbf{F}_{\text{RF}}^{(0)}$ with random phases and set k = 0";
2. **Repeat**
3. **Fix** $\mathbf{F}_{\text{RF}}^{(k)}$, and $\mathbf{F}_{\text{BB}}^{(k)} = \mathbf{F}_{\text{RF}}^{(k)\dagger}\mathbf{F}_{\text{opt}}$;
4. Optimize $\mathbf{F}_{\text{RF}}^{(k+1)}$ using Algorithm 1 where $\mathbf{F}_{\text{RF}}^{(k)}$ is fixed;
5. $k \leftarrow k + 1$;

6. **Until** a stopping criterion triggers;
7. For the digital-baseband beamformer at the base station side, apply unit normalization $\hat{\mathbf{F}}_{\mathrm{BB}} = \frac{\sqrt{N_S}}{\|\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\|_F}\mathbf{F}_{\mathrm{BB}}$

Lemma 1: according to [1] "its shown that if the Euclidean distance before normalization is $\left\|\mathbf{F}_{\mathrm{opt}} - \mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\right\|_F \leq \delta$, then after normalization we have $\left\|\mathbf{F}_{\mathrm{opt}} - \mathbf{F}_{\mathrm{RF}}\hat{\mathbf{F}}_{\mathrm{BB}}\right\|_F \leq 2\delta$".

Proof of Lemma 1 is formulated in [1] as: Define the normalization factor $\frac{\sqrt{N_S}}{\|\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\|_F} = \frac{1}{\lambda}$ and therefore, $\|\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\|_F = \lambda\sqrt{N_S} = \lambda\left\|\mathbf{F}_{\mathrm{opt}}\right\|_F$.

From the norm inequality, we have

$$\left\|\mathbf{F}_{\mathrm{opt}} - \mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\right\|_F \geq \left|\left\|\mathbf{F}_{\mathrm{opt}}\right\|_F - \|\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\|_F\right| = |1-\lambda|\left\|\mathbf{F}_{\mathrm{opt}}\right\|_F, \tag{4.16}$$

Which is the same as $\left\|\mathbf{F}_{\mathrm{opt}}\right\|_F \leq \frac{1}{1-\lambda}\delta$.

When $\lambda \neq 1$, which indicates $\left\|\mathbf{F}_{\mathrm{opt}} - \mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\right\|_F \neq 0$,

$$\begin{aligned} &\left\|\mathbf{F}_{\mathrm{opt}} - \mathbf{F}_{\mathrm{RF}}\hat{\mathbf{F}}_{\mathrm{BB}}\right\|_F \\ &= \left\|\mathbf{F}_{\mathrm{opt}} - \mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}} + \left(1-\frac{1}{\lambda}\right)\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\right\|_F \\ &\leq \left\|\mathbf{F}_{\mathrm{opt}} - \mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\right\|_F + \left|1-\frac{1}{\lambda}\right|\|\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\|_F \\ &\leq \delta + |\lambda \quad 1|\left\|\mathbf{F}_{\mathrm{opt}}\right\|_F \leq \delta + \frac{|\lambda \quad 1|}{|\lambda \quad 1|}\delta = 2\delta \end{aligned} \tag{4.17}$$

From lemma 1, as it is articulated using the authors [1] "we can see that if we can make the Euclidean distance between the optimal digital precoder and the hybrid precoders sufficiently small when ignoring the power constraint in Equation (*2.6*), the normalization step will also achieve a small distance to the optimal digital beamformer."

Since the "objective function" in Equation (*2.6*) is minimized at steps 3 and 4, each computational optimization steps will never make it worse. Furthermore, "the objective function" is non-decreasing. By combining these two properties, we can ensure that the "MO-AltMin algorithm" will eventually reach a solution that is both feasible and optimal.

However, the MO-AltMin algorithm has high computational complexity because it uses a "line search algorithm (Algorithm 1)" to update the analog beamformer in each iteration, which causes "the nested loops in the algorithm to slow down the overall computational procedure". Additionally, the "Kronecker products", Equation (*2.6*), will produce two-matrices of dimension $N_{\mathrm{RF}}^{t}N_t \times N_s N_t$, which increase with the size of the antenna and, as stated in [1], result in an exponential increase in the MO-AltMin algorithm computational complexity. "Despite this high complexity, we observe that the MO-AltMin algorithm based on the manifold optimization directly solves the hybrid precoder design problem under the unit modulus constraints, which could improve spectral efficiency when compared to other algorithms".

4.3 Complexity Analysis of Fully Digital and MO-AltMin Algorithms

In this section we will calculate the computational complexity of the fully digital and MO-AltMin beamforming algorithms for 5G millimeter wave MIMO systems, mainly by focusing on matrix operations and their respective computational costs with the Big O notation. The evaluation of the PE-AltMin algorithm has covered in chapter 3 of subsection 3.3.

Big O notation is a mathematical concept which is used to describe the efficiency as well as the performance of an algorithm. It specifically measures the worst-case scenario of an algorithm's time and space complexity, helping to predict how the algorithm scales with increasing input size.

4.3.1 Complexity of Fully digital (Optimal) beamforming

In fully digital (Optimal) beamforming, the precoder $\mathbf{F}_{\mathrm{opt}}$ is computed directly as: $\mathbf{F}_{\mathrm{opt}} = \mathbf{V}_{\mathrm{opt}}$.

Where $\mathbf{V}_{\mathrm{opt}}$ is derived from the "singular value decomposition (SVD) of the channel matrix H."

Step 1: Compute SVD of $\boldsymbol{H}$ using $\boldsymbol{H} = \mathbf{U\Sigma V}^{H}$**.** Where: $\boldsymbol{H} \in \mathbb{c}^{N_r \times N_t}$,$\mathbf{U} \in \mathbb{c}^{N_r \times N_r}$, $\sum \in \mathbb{c}^{Nr \times N_t}$ and $\mathbf{V} \in \mathbb{C}^{N_t \times N_t}$.

Therefore, the complexity of computing the SVD is $O(\min(N_r, N_t)N_r N_t^2)$. For specific scenario where $N_t > N_r$, the complexity is reduced to:

$$O(N_r N_t^2) \tag{4.18}$$

Step 2: extracting V_{opt} for beamforming.

As we are taking the first N_s columns of $\boldsymbol{V}$, this operation has negligible complexity. Therefore, the complexity of the optimal fully-digital beamformer is represented by equation (4.18).

4.3.2 Computational analysis of manifold optimization alternating minimization ("MO-AltMin") algorithm.

Algorithm: "Manifold Optimization Based Hybrid Precoding for the Fully-connected mapping"

Input: $\mathbf{F_{opt}}$

1. "Construct $\mathbf{F}_{\text{RF}}^{(0)}$ with random phases and set k = 0";
2. **Repeat**
3. Fix $\mathbf{F}_{\text{RF}}^{(k)}$, and $\mathbf{F}_{\text{BB}}^{(k)} = \mathbf{F}_{\text{RF}}^{(k)\dagger}\mathbf{F}_{\text{opt}}$;
4. Optimize $\mathbf{F}_{\text{RF}}^{(k+1)}$ using Algorithm 1 where $\mathbf{F}_{\text{RF}}^{(k)}$ is fixed;
5. $k \leftarrow k + 1$;
6. **Until** an algorithm ending rule is activated;
7. "For the digital precoder at the transmitter end, normalize $\hat{\mathbf{F}}_{\text{BB}} = \frac{\sqrt{N_s}}{\|\mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}\|_F}\mathbf{F}_{\text{BB}}$"

- **Construct the analog beamformer, $\mathbf{F}_{\mathbf{RF}}^{(0)}$ with unit-modulus elements by setting $N_t \times N_{RF}$ elements randomly**.

The complexity at step 1 of the MO-AltMin algorithm will be negligible. But the process will repeat until convergence or looped I iterations times.

$$O(N_t N_{RF}) \tag{4.19}$$

- **Compute the digital beamformer $\mathbf{F}_{\mathbf{BB}}^{(k)}$ by fixing $\mathbf{F}_{\mathbf{RF}}^{(k)}$ using the singular value decomposition:**

Given $\mathbf{F}_{\mathrm{RF}} \in^{N_t \times N_{RF}}$ which is fixed in this case, and the unconstrained optimal precoder $\mathbf{F}_{\mathrm{opt}} \in^{N_t \times N_s}$, compute the pseudo-inverse $\mathbf{F}_{\mathrm{RF}}^{\dagger}$ using the SVD for numerical stability.

$$\mathbf{F}_{\mathrm{RF}} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{H} \tag{4.20}$$

Where $\mathbf{U} \in \mathbb{C}^{N_t \times N_{RF}}$, and $\mathbf{V} \in \mathbb{C}^{N_{RF} \times N_{RF}}$ are unitary matrices and $\sum \in \mathbb{C}^{N_{RF} \times N_{RF}}$ is diagonal with singular values. The pseudo-inverse is given by:

$$\mathbf{F}_{\mathrm{RF}}^{\dagger} = \mathbf{V}\boldsymbol{\Sigma}^{-1}\mathbf{U}^{H} \tag{4.21}$$

The complexity of these operations is therefore given by:

- SVD of $\mathbf{F}_{\mathrm{RF}}$ has a complexity of $O(N_t N_{RF}^2)$
- Inverse of the diagonal matrix $O(N_{RF})$
- Matrix multiplications of $\mathbf{V}\boldsymbol{\Sigma}^{-1}$ and $\mathbf{V}\boldsymbol{\Sigma}^{-1}\mathbf{U}^{H}$ has $O(N_{RF}^3)$ and $O(N_t N_{RF}^2)$, respectively.

Therefore, the total complexity of $\mathbf{F}_{\mathrm{RF}}^{\dagger}$ becomes:

$$O(N_t N_{RF}^2) + O(N_{RF}^3) + O(N_t N_{RF}^2) \approx O(N_t N_{RF}^2) \tag{4.22}$$

$$\text{since } N_t > N_{RF}$$

Multiplication of $\mathbf{F}_{\mathrm{RF}}^{\dagger}$ with $\mathbf{F}_{\mathrm{opt}}$, from step 3 of the algorithm with dimensions of $\mathbf{F}_{\mathrm{RF}}^{\dagger} \in \mathbb{C}^{N_{RF} \times N_t}$ and $\mathbf{F}_{\mathrm{opt}} \in \mathbb{C}^{N_t \times N_s}$ has a complexity of $O(N_{RF} N_t N_s)$.

So, the total complexity of step 3 of the algorithm becomes:

$$O(N_t N_{RF}^2 + N_{RF} N_t N_s) \tag{4.23}$$

For $N_{RF} > N_s$, this simplifies to $O(N_t N_{RF}^2)$

- **Update the analog precoder via Riemannian Conjugate Gradient (RCG)**

The aim of this step is to optimize $\mathbf{F}_{RF} \in \mathbb{C}^{N_t \times N_{RF}}$ with $|[\mathbf{F}_{RF}]_{i,j} = 1$, minimize equation (31)

$$f(\mathbf{F}_{RF}) = \left\| \mathbf{F}_{\mathrm{opt}} - \mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}} \right\|_F^2 \tag{4.24}$$

The inputs of equation (4.24) are given as follows:

- $\mathbf{F}_{\text{opt}} \in \mathbb{C}^{N_t \times N_s}$, the fixed digital precoder $\mathbf{F}_{\text{BB}} \in \mathbb{C}^{N_{RF} \times N_s}$ and $\mathbf{F}_{\text{RF}}^{(0)}$ is initialized using a random unit-modulus matrix. Equation (48) is computed using the following steps with the respective complexity.

Step1: Compute the Euclidean gradient is computed using equation (4.13) as stated in [1] and [22] as

$$\nabla f(\mathbf{F}_{RF}) = -2\left(\mathbf{F}_{\text{opt}} - \mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}\right)\mathbf{F}_{\text{BB}}^{H} \tag{4.25}$$

Operations of equation (4.25) involves computations of Floating Point Operations (FLOPs) of $2N_{\text{t}}N_{RF}N_{\text{s}} - N_{\text{t}}N_{\text{s}}$ for the residual matrix $\mathbf{F}_{\text{opt}} - \mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}$ and $2N_{\text{t}}N_{RF}N_{\text{s}}$ for multiplication of the residual matrix with the Hermitian of the fixed digital precoder. Hence, the total complexity for this step is formulated as $O(N_{\text{t}}N_{RF}N_{\text{s}})$.

Step 2: project gradient onto the tangent space

The tangent space projection enforces the unit-modulus constraint.

$$\xi = \nabla f(\mathbf{F}_{RF}) - Re\{\nabla f(\mathbf{F}_{RF}) \odot \mathbf{F}_{RF}^{*}\} \odot \mathbf{F}_{RF} \tag{4.26}$$

Complexities of operations of equation (58) are element wise product $\nabla f \odot \mathbf{F}_{RF}^{*}$ ($O(N_{\text{t}}N_{RF})$), taking the real part also has $O(N_{\text{t}}N_{RF})$, second element-wise product $O(N_{\text{t}}N_{RF})$ and subtraction from ∇f has $O(N_{\text{t}}N_{RF})$. The total complexity of this step therefore becomes $ON_{\text{t}}N_{RF}$.

Step 3: compute search directions (Conjugate Gradient) (step 6 of algorithm 1)

Update the search direction $\mathbf{d}_k$ using the Polak-Ribiere formula:

$$\beta_k = \frac{\langle \xi_k, \xi_k - \xi_{k-1} \rangle}{\langle \xi_{k-1} - \xi_{k-1} \rangle}, \mathbf{d}_k = -\xi_k + \beta_k \mathbf{d}_{k-1} \tag{4.27}$$

Operations in equation (4.27) are: Iner product of $\langle \xi_k, \xi_k \rangle$ has complexity of $v$ector subtraction $\langle \xi_k - \xi_{k-1} \rangle$, inner product $\langle \xi_k, \xi_k - \xi_{k-1} \rangle$ and scalar division and vector updating all have complexity of $O(N_{\text{t}}N_{RF})$. Meaning the total complexity is $O(N_t N_{RF})$.

Step 4: retraction or projection back to the manifold

This process is computed using equation (4.14) to retract the update onto the unit-modulus manifold.

$$\mathbf{F}_{\mathrm{RF}}^{(k+1)} = \mathrm{Retr}(\mathbf{F}_{\mathrm{RF}}^{(k)} + \alpha_k \mathbf{d}_k = \frac{[\mathbf{F}_{\mathrm{RF}}^{(k+1)} + \alpha_k \mathbf{d}_k]_{i,j}}{|[\mathbf{F}_{\mathrm{RF}}^{(k+1)} + \alpha_k \mathbf{d}_k]_{i,j}|} \tag{4.28}$$

Operations of this equation are vector addition of $\mathbf{F}_{\mathrm{RF}}^{(k+1)} + \alpha_k \mathbf{d}_k$ and element-wise normalization with same complexity of $O(N_t N_{RF})$.

Step 5: choose Armijo line search step size α_k

The aim of this step is to find α_k satisfying Armijo conditions:

$$f(\mathbf{F}_{RF} + \alpha_k \mathbf{d}_k) \leq f(\mathbf{F}_{RF}) + c\alpha_k \mathrm{Re}\langle \xi_k, \mathbf{d}_k \rangle \tag{4.29}$$

Operations of the backtracking steps of equation (4.29) are computing $\mathbf{F}_{RF} + \alpha_k \mathbf{d}_k)$ with complexity of $O(N_t N_{RF})$, evaluation of $f(\cdot)$ with complexity of $O(N_t N_{RF} N_s)$ and inner product of $\langle \xi_k, \mathbf{d}_k \rangle$ has complexity of $O(N_t N_{RF})$. The total per line search iteration has complexity of $O(N_t N_{RF} N_s)$.

For the RCG iteration, from step 1 to 5, the dominant term is only the Euclidean gradient (if $N_s \geq 1$) of step 1. Therefore, the overall complexity of K RGG iteration is given by:

$$O(K N_t N_{RF} N_s). \tag{4.30}$$

- **Computing Power normalization (step 7)**

The digital precoder at the transmitter side is normalized as $\hat{\mathbf{F}}_{\mathrm{BB}} = \frac{\sqrt{N_s}}{\|\mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}}\|_F} \mathbf{F}_{\mathrm{BB}}$, so that to satisfy the total transmit power constraint. This involves operations of norm computation and scaling with complexity $O\,(N_{RF} N_s)$. $O(N_{RF} N_s)$, respectively.

The total complexity of for I MO-AltMin and K RGC iterations is formulated as:

$$O(I(N_t N_{RF}^2 + (K N_t N_{RF} N_s)) \tag{4.31}$$

Table 4:1: Summary of Comparison of Computational Complexity of different algorithms

	Algorithm	Computational steps	Computational complexity per iteration	Overall computational complexity for I iteraions
1.	Optimal digital precoder	- Compute SVD of $H = USV^H$ Select the first N_s columns of V as $\boldsymbol{F}_{\mathrm{opt}}$	$O(N_r N_t^2)$	$O(N_r N_t^2)$

2.	PE-AltMin (Fully-connected)	- Construct $\boldsymbol{F}_{\mathrm{RF}}^{(0)}$ with random phases. - Compute SVD of $\boldsymbol{F}_{\mathrm{opt}}^{H}\boldsymbol{F}_{\mathrm{RF}}^{(k)} = U^{(k)}S^{(k)}\left(V_1^{(k)}\right)^{H}$ - Solve $\boldsymbol{F}_{\mathrm{DD}}^{(k)} = V_1^{(k)}(U^{(k)})^{H}$ - Solve $\arg\left(\mathrm{F}_{\mathrm{RF}}^{(k+1)}\right) = \arg\left(\boldsymbol{F}_{opt}\,{\boldsymbol{F}_{DD}}^{(k)^H}\right)$; until convergence, iterative	$O((N_tN_{RF}N_s) + (N_sN_{RF}^2))$	$O(I(N_tN_{RF}N_s) + (N_sN_{RF}^2))$
3.	MO-AltMin	- Construct $\mathrm{F}_{\mathrm{RF}}^{(0)}$ with unit modulus elements with random $N_t \times N_{RF}$ - Compute $\mathrm{F}_{\mathrm{BB}}^{(k)}$ by fixing $\mathrm{F}_{\mathrm{RF}}^{(k)}$ - Update the analog precoder via RCG	$O(N_tN_{RF}^2 + (KN_tN_{RF}N_s))$	$O(I(N_tN_{RF}^2 + (KN_tN_{RF}N_s))$

5 Chapter Five: Performance Evaluation

5.1 Simulation Setup

In this chapter, we carry out numerical simulations to assess how well the low-complexity algorithm based on alternating minimization performs.

The channel is modeled using 5 clusters ($N_{cl} = 5$), with each cluster containing 10 rays ($N_{\text{ray}} = 10$). The average power per cluster is normalized to $\sigma_{\alpha,i}^2 = 1$. Data streams are transmitted from a system equipped with $N_t = 144$ antennas at the transmitter to a receiver with $N_r = 36$ antennas; both ends employ a uniform spherical phased array (USPA).

The angle of departure (AoD) at the transmitter and the angle of arrival (AoA) at the receiver are drawn from a Laplacian distribution. The mean angles of these distributions are chosen uniformly at random over the full [0, 2π] range, while the angular spread is fixed at 10 degrees. Antenna elements within the USPA are spaced half a wavelength apart.

All simulation results are averaged over 1000 independent channel realizations. For every run of the alternating minimization procedure, the initial phases of the RF precoder $\mathbf{F}_{\text{RF}}$ are randomly initialized according to a uniform distribution over [0, 2π].

5.2 Results and Discussion

For analysis and simulation of computational complexity and spectral efficiency, the system parameters we use are: The channel is modeled so as to have $N_{cl} = 5$ clusters, $N_{\text{ray}} = 10$ rays, and the average power of clusters is $\sigma_{\alpha,i}^2 = 1$. Data streams are sent from a transmitter with $N_t = 144$ to a receiver with $N_r = 36$ antennas, both equipped with USPA. The azimuth and elevation AoDs and AoAs follow the Laplacian distribution with uniformly distributed mean angles over [0, 2π] and angular spread of 10 degrees. The antennas elements in the USPA are separated by half wavelength distance and all simulations are averaged over 1000 channel realization and all the alternating minimization algorithms, $\mathbf{F}_{\text{RF}}$ follows a uniform distribution over [0, 2π].

For the evaluation of computational complexity and spectral efficiency through numerical analysis and simulations, the following system configuration is adopted. The wireless channel is

represented by using a clustered model with $N_{cl} = 5$ clusters, each containing $N_{\text{ray}} = 10$ propagation rays, and the average power associated with every cluster is normalized such that $\sigma_{\alpha,i}^2 = 1$. Transmission occurs from a transmitter employing $N_t = 144$ antennas to a receiver equipped with $N_r = 36$ antennas. Both the transmitter and receiver utilize a uniform spherical phased array (USPA). The azimuth and elevation angles of departure (AoDs) at the transmitter and angles of arrival (AoAs) at the receiver are generated according to a Laplacian distribution. The mean angle for each distribution is drawn uniformly at random from the interval [0, 2π], and the angular spread is set to 10 degrees. The antenna elements within the USPA are spaced at half-wavelength intervals. To ensure statistical reliability, all simulation outcomes are averaged across 1000 independent channel realizations. Furthermore, in each execution of the alternating minimization procedure, the initial RF precoding matrix $\mathbf{F}_{\text{RF}}$ is initialized by drawing its phases independently and uniformly from the range [0, 2π].

5.2.1 Comparison of computational complexity

For evaluation of complexity using the Big O analysis, we give emphasis on FLOPs (Floating Point Operations) that refers to the number of floating-point addition, subtraction, multiplication, or division operations performed by an algorithm to measure the computational cost of the fully-digital, PE-AltMin, and MO-AltMin algorithms.

Table 5:1:complexity comparison of fully-digital, PE-AltMin and MO-AltMin algorithms

Algorithm	**Arithmetic workload, FLOPs**	**Scaling**	**Hardware cost**
Fully Digital	186,624	$O\,(N_r^2 N_t)$	$N_t = 144$ RF chains
PE-AltMin	48,640	$O\,(T\,(N_s N_{RF}^2) + (N_t N_{RF} N_s))$	$N_{RF} = 8$ RF chains
MO-AltMin	322,560	$O\,(T\,(N_t N_{RF}^2) + K(N_t N_{RF} N_s))$	$N_{RF} = 8$ RF chains

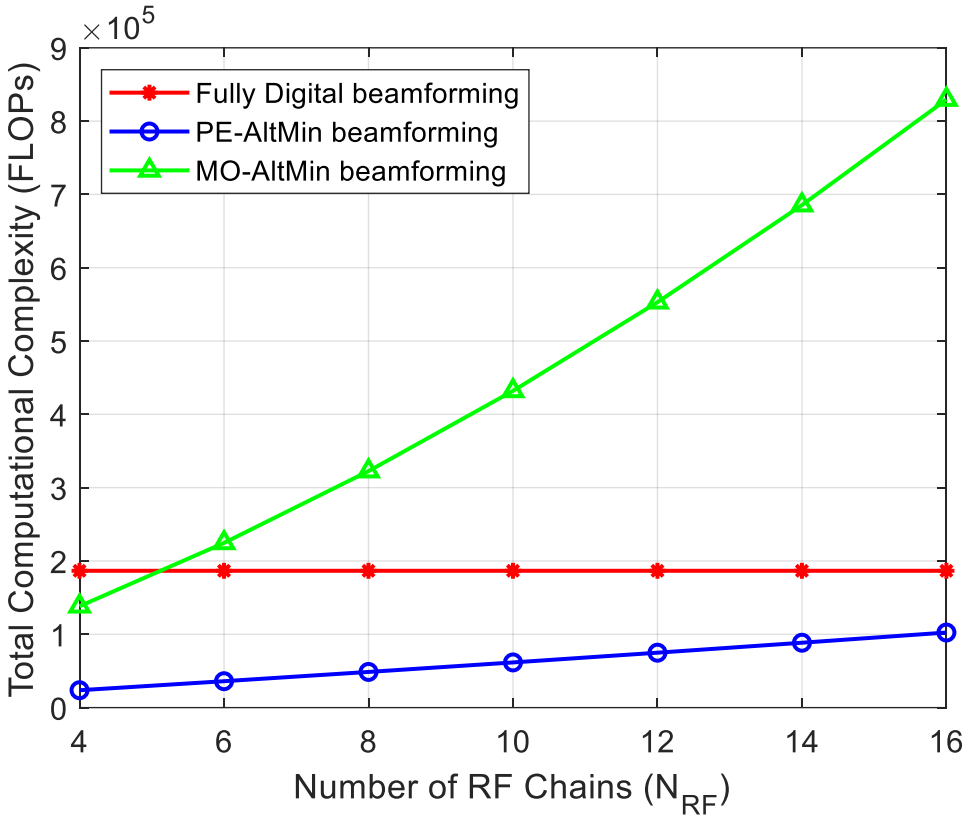


Figure 5.1: Complexity vs Number of RF chains, given $N_s = 4,\ N_t = 144, N_r = 36$, RCG iterations = 5, and 10 algorithm iterations.

Figure 5.1 shows that PE-AltMin algorithm has a slight quadratic growth with N_{RF} and MO-AltMin has a steeper quadratic growth due to RCG to optimize $\mathbf{F}_{\mathrm{RF}}^{(k+1)}$ using algorithm 1.

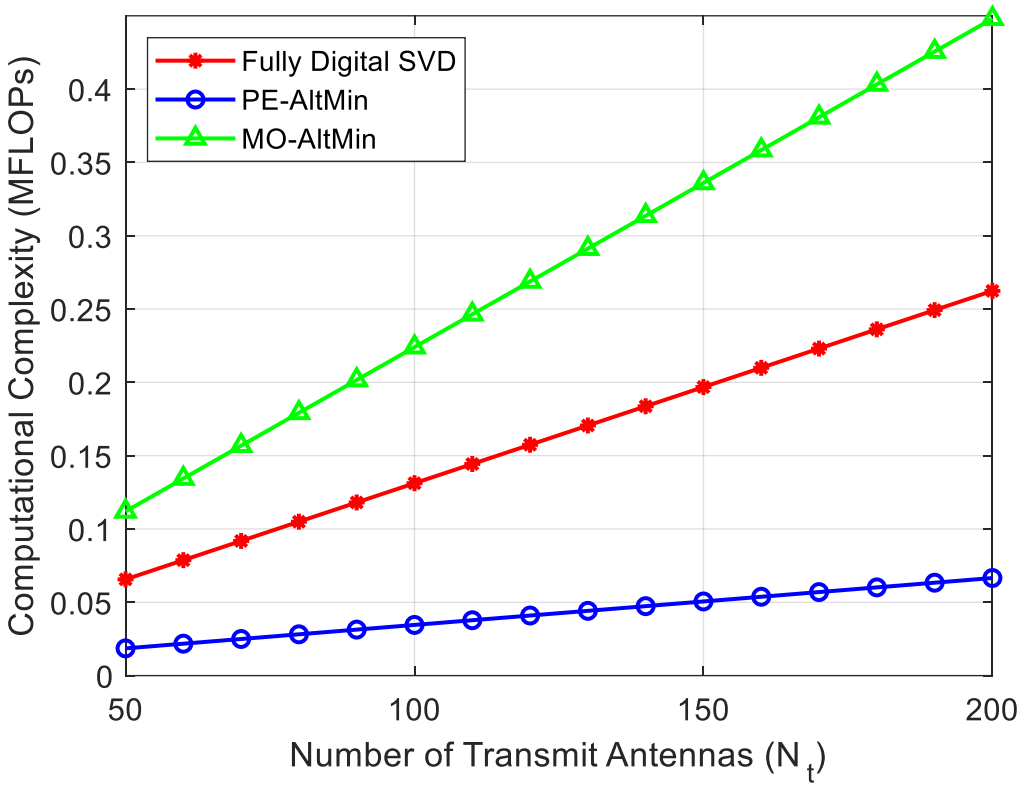


Figure 5.2: Complexity vs Number of transmit antenna, given $N_s = 4$ $N_{RF} = 8$, $N_r = 36$, RCG iterations = 5, and 10 algorithm iterations.

Figure 5.2 shows that all the three algorithms have linear growth computational complexity with the number of transmitter antennas growing. The dominant term in the PE-AltMin is $O(N_t N_{RF} N_s)$ with 48.6 MFLOPs (millions, 10^6 FLOPs) at $N_t = 144$, that is 3.8 times cheaper than 188.6 MFLOPs of fully digital (used as a baseline for comparison). MO-AltMin at $N_t = 144$ on the contrary has 322.6 MFLOPs, which makes it the most complex.

Fully-digital beamformer is impractical for mmWave massive MIMO systems due to very high hardware costs, 144 RF chains in our case. PE-AltMin technique is best due to low arithmetic overload as well as low hardware cost.

Iteration is linearly related to computational complexity as shown in *Figure 5.3*, where PE-AltMin increases linearly and MO-AltMin showing steeper linear increase. More iteration will lead to accuracy but requires more processing time.

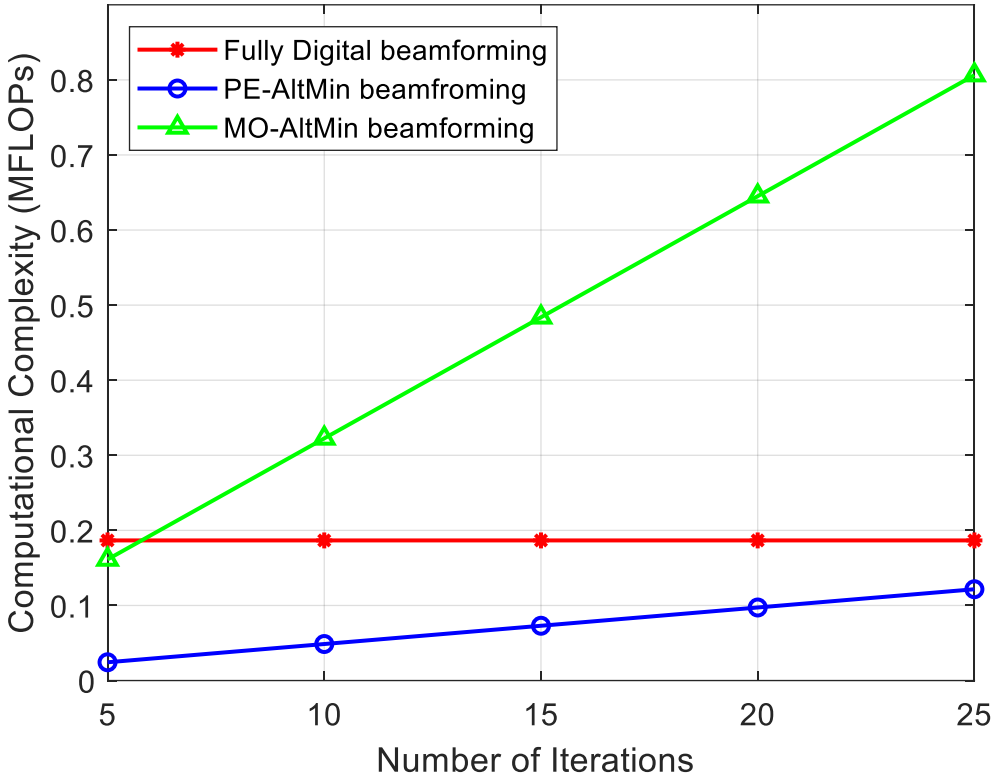


Figure 5.3: Complexity vs Number of Iterations, given $N_s = 4$, $N_{RF} = 8, N_t = 144, N_r = 36$, RCG iterations = 5, and 10 algorithm iterations

5.2.2 Spectral Efficiency Evaluation

Figure *5.4* illustrates the impact of the number of RF chains in different algorithms in the region $N_{RF}^t = N_{RF}^r = N_{RF} \in [4, 20]$. From the plot we find that the PE-AltMin algorithm has a small gap compared with the MO-AltMin algorithm. The reason is the low complexity algorithm tries to minimize the upper bound instead of the original objective function. As discussed, in the hybrid design of the low complexity beamformer for the fully-connected mapping section, the upper bound gets tight when $N_{RF} = N_s$ and gets looser when N_{RF} increases, which consequently can affect the gap between the two algorithms. As it can be seen from the spectral efficiency plot the most gain occur up to $N_{RF} = 8$ region, beyond this region improvements diminish due to the limited resolution of phase shifter in the analog beamforming, interference saturation and algorithmic convergence.

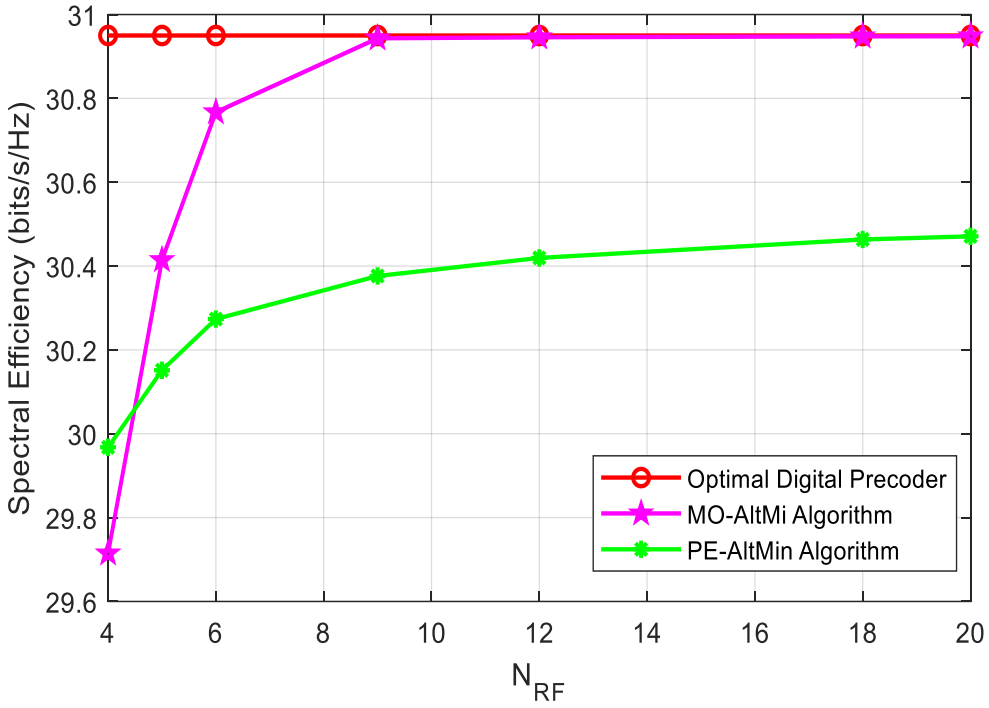


Figure 5.4: Spectral efficiency achieved by different algorithms given $N_s = 4$, $SNR\ (dB) = 0$ and $N_{RF}^t = N_{RF}^r = N_{RF}$

5.2.3 Energy Efficiency Evaluation

In this section, we will see energy efficiency with respect to the achievable sum rate and number of RF chains. The achievable sum rate is given by equation (2.3). according to the energy consumption model in [24] the energy efficiency (EE) in wireless communication systems can be defined as

$$\mathrm{EE} = \frac{R}{P_{\mathrm{total}}} = \frac{R}{P_{\mathrm{BB}} + P_{\mathrm{RF}} + P_{\mathrm{PS}} + P_{\mathrm{PA}}} = \frac{\mathrm{R}}{P_{\mathrm{t}} + N_{\mathrm{RF}}P_{\mathrm{RF}} + N_{\mathrm{PS}}P_{\mathrm{PS}} + P_{\mathrm{PA}}} \tag{5.1}$$

Where $P_{total} \cong P_{\mathrm{t}} + N_{\mathrm{RF}}P_{\mathrm{RF}} + N_{\mathrm{PS}}P_{\mathrm{PS}} + P_{\mathrm{PA}}$ is the total energy consumption, $P_{\mathrm{t}} = P_{\mathrm{BB}}$ is the transmitted power (baseband processing power), P_{RF} is the RF energy consumption, P_{PS} is energy consumed by phases shifters including energy used for excitation and compensation of insertion

loss [25], N_{PS} and N_{RF} are number of phase shifters and RF chains. For the simulation we used practical values from [25] and [1] $P_{\mathrm{RF}} = 250$ mW, $P_{\mathrm{PS}} = 1$ mW, and $P_{\mathrm{t}} = 1$ W, $P_{\mathrm{A}} = 1$ W.

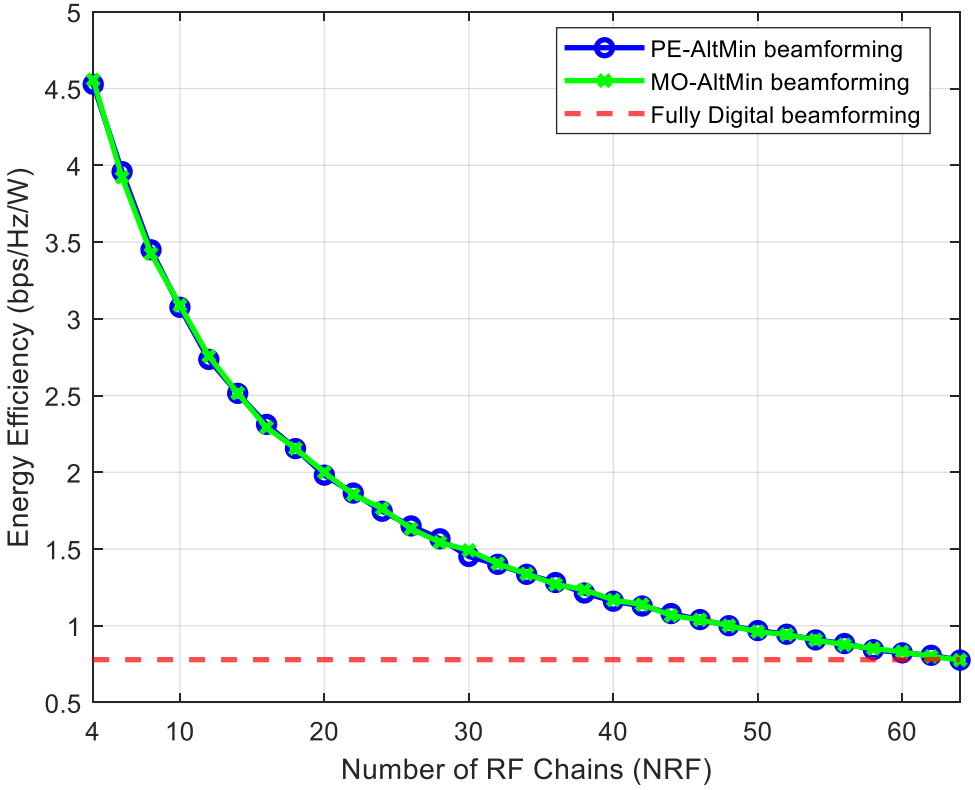


Figure 5.5. Energy efficiency comparison against the number of RF chains

Figure 5.5 shows the energy efficiency of three algorithms against the number of RF chains used where SNR $= 0$ dB. We see that both PE-AltMin and MO-AltMin algorithms can achieve higher energy efficiency than the fully digital precoder when RF chains are limited. The reason for similar EE of the two algorithms is both are implemented for the fully connected structure and as a result have the same number of RF chains, PS and similar hybrid precoding technique.

6 Chapter Six: Conclusion and Future Work

6.1 Conclusion

This thesis has evaluated the performance of the low complexity hybrid beamforming technique for 5G millimeter wave MIMO systems called PE-AltMin algorithm based on computational complexity, spectral efficiency and energy efficiency by comparing with the MO-AltMin algorithm and the fully-digital beamforming techniques. Both beamforming techniques are implemented for the fully connected hybrid beamforming structure where each RF chain is connected to all antennas. Fully-digital beamforming is used only for bench marking.

In both the PE-AltMin and Mo-AltMin hybrid beamforming design the complexity variation comes from the analog beamformer because the digital beamformer is solved using the closed form solution. The later technique updates the analog beamforming using the manifold optimization technique where in each iteration gradient descent search is performed to update it. Whereas the former updates the analog beamformer by extracting the phases of the product of unconstrained optimal precoder and the digital precoder.

It's shown in chapter five that for better study we have simulated and showed the relationship and effect between efficiency measuring metrics with respect to antenna parameters, and iteration. The results show that PE-AltMin offers lower computational complexity with comparable energy efficiency and slight spectral efficiency loss. On the other hand, MO-AltMin achieves higher spectral efficiency with very high computational overhead.

Therefore, the PE-AltMin algorithm provides favorable tradeoff for better computational resources utilization where hardware simplicity matters and real-time deployment are prioritized, whereas MO-AltMin algorithm is preferable when computational resources are not the challenge and performance margins matter the most.

6.2 Recommendation for Future Work

Based on the findings, it's advisable to restrict the number of RF chains ($N_{\mathrm{RF}} < 8$) to mitigate complexity, as the SE gain diminishes (has no improvement) far from this region. The future study

is going to focus on optimizing RF chain allocations with machine learning methods to improve the computational complexity of the PE-AltMin algorithm.

The low complexity of the PE-AltMin is due to adoption of minimizing the upper bound, rather than the objective function, as presented in equation (*3.4*), the effectiveness of this algorithm is restricted to $N_s \leq N_{\mathrm{RF}}^t \leq 2N_s$. This upper bound condition is not feasible for practical scenarios and further works can be extended to mitigate the spectral efficiency loss for a system where transmitter number of RF chains are not bounded.

References

[1] Xianghao Yu et al, "Alternating Minimization Algorithms for Hybrid Precoding in millimeter Wave MIMO Systems," *Communication Systems,* vol. 01, 2018.

[2] G. Andrews et al, "What will 5G be?," *Areas Commun.,* vol. 32, no. 6, pp. 1065-1082, 2014.

[3] J. Zhang et al, "Massive hybrid antenna array for millimeter-wave cellular communications," *IEEE Wireless Commun.,* vol. 22, no. 1, pp. 79-87, Feb 2015.

[4] Rappaport et al, "Millimeter wave mobile communications for 5G cellular: It will work!," *IEEE Access,* vol. 1, pp. 335-349, 2013.

[5] Zhang et al, "Beam alignment and tracking for millimeter wave communications: State-of-the-art and challenges," *IEEE Communications Magazine,* vol. 57(1), pp. 34-40, 2019.

[6] Heath et al, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE Journal of Selected Topics in Signal Processing,* vol. 10, pp. 436-453, 2016.

[7] Gao et al, "Low RF-complexity technologies to enable millimeter-wave MIMO with large antenna arrays," *IEEE Communications Magazine,* vol. 56(4), pp. 211-217, 2016.

[8] Soharbi et al, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE Journal of Selected Topics in Signal Processing,* vol. 10(3), pp. 501-513, 2016.

[9] T. E. Bogale et al, "Massive MIMO and millimeter wave for 5G wireless HetNet: Potential benefits and challenges," *IEEE Vehicular Technology Magazine,* vol. 11(1), pp. 64-75, 2016.

[10] A. Alkhateeb et al, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE Journal of Selected Topics in Signal Processing,* Vols. 831-846, p. 8(5), 2014.

[11] Ye Wang & Weixia Zou, "Low Complexity Hybrid Precoding Algorithm in Millimeter Wave MIMO Systems," *IEEE/CIC International Conference on Communications in China (ICCC Workshops), Beijing, China, 2018,* pp. 16-20, 2018.

[12] A. Gupta and &. K. Jha, "A Survey of 5G Network: Architecture and Emerging Technologies," pp. 1206 - 1232, 28 July 2015.

[13] Z. Q Luo et al, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process,* vol. 27, pp. 20 - 30, May 2010.

[14] Swanima et al, "Performance Evaluation of a Millimeter Wave MIMO Hybrid Beamforming System," *IEEE,* vol. 10, pp. 1938-1946, Oct 2016.

[15] O. El Ayach, et. al., Spatially sparse precoding in millimeter wave MIMO systems, vol. 13, IEEE Trans. wireless Commun., March 2014, pp. 1499-1513.

[16] Jialin Cai et al, "A low complexity Hybrid precoding scheme for massive MIMO system," *International Symposium on Communications and Information Technologies (ISCIT),* vol. 1, 2016.

[17] N. Farsad and A. Goldsmith, "Neural Network-Based Beamforming for MmWave Massive MIMO," *IEEE Trans. Signal Process,* 2020.

[18] Saleh Valenzuela et al, "A Statistical Model for Indoor Multipath Propagation," *IEEE Journal on Selected Areas in Communications,* vol. 05, no. 02, pp. 128-137, 1987.

[19] R. A. Horn et al, Matrix Analysis, Cambridge, U.K.: Cambridge University Press, 2012.

[20] E. Haung et al, "On achieving optimal rate of digital precoder by RF-baseband codesign for MIMO systems," *in Proc. 80th IEEE Veh. Technol. Conf. (VTC Fall),* pp. 1-5, Sept 2014.

[21] R. Mahony et. al., "Optimization Algorithms on Matrix Maniflods," 2009.

[22] A. Hjorungnes, "Complex-valued Matrix Derivatives," 2011.

[23] D. P. Bertsekas , "Nonlinear programming," *Athena scientific,* 1999.

[24] S. Cui et al., "Energy-constarined modulation optimization," *IEEE Trans. Wireless Commun,* vol. 61, no. 5, pp. 1880-1890, 2013.

[25] C. A. Ballins , Antenna Theory: Analysis and Design, Hoboken, NJ, USA: Wiley, 2012.