Pilot Assignment and Cluster Formation in Cell-Free Massive MIMO Networks

Authors

Christodoulos Tsialamanis
Paschalina Foti
Wenjian Ma
Fikiri Mwembe
Sharayesh Bhosale
Swadesh Ranjan Nayak

Communication Systems Design Project Report

Communication Systems
School of Electrical Engineering and Computer Science
KTH Royal Institute of Technology
Stockholm, Sweden

11 January 2023

Examiner: Dejan Manojlo Kostic
Supervisors: Ki Won Sung, Mahmoud Zaher

Abstract

The wireless communication sector has seen a significant increase in demand as more and more users have started using this technology and a lot of service providers have started shifting their services on this channel. This has resulted in a need for better ways of transferring data among the users in an efficient way. As we know that bandwidth is a limited resource in this sector and the number of users has skyrocketed as compared to a few years in the past, the Spectral Efficiency and the channel capacity are the main focus of this industry. Cell-free massive MIMO technology has been utilized in recent years to increase the spectral efficiency of 5G and beyond systems. Massive MIMO has also gained popularity due to its ability to solve issues like interference a user experiences at the edge of a cell in cellular system and provide uniform services (data rates) to all the users. A network scenario has been generated including a number of multi-antenna access points (APs) and single-antenna UEs. We consider imperfect channel state information (CSI) obtained via pilot training and utilize linear precoding schemes based on the locally available CSI at each AP. The network has been simulated to obtain the Spectral Efficiency which is the primary performance metric. A Pilot Assignment and Clustering algorithm has been implemented which ensures that the pilot sequence which provides the most degree of separation between the two users sharing the same pilot sequence is selected. Even if there are more number of pilot sequences which provide desired degree of separation, the sequence with the least amount of interference is used in this particular instance. This setup results in better performance regarding the SE and data rates of the system as compared to other pilot assignment and cluster formation algorithms which were previously being used in different papers.

Acknowledgements

We would like to acknowledge our TA Mahmoud Zaher and our adviser Professor Ki Won Sung for guiding us and clearing our doubts regarding different aspects of this project.

Contents

List of Figures

List of Tables

List of Acronyms and Abbreviations

3GPP	Third Generation Partnership Project
5G	Fifth Generation Mobile Network
AP	Access Point
CDF	Cumulative Distribution Function
CSI	Channel State Information
DL	DownLink
DPR	Dynamic Power Reuse
IB-KM	Interference Based K-means
IGS	Iterative Grid Search
TA	Teaching Assistant
MATLAB	Matrix Laboratory
MIMO	Multiple Input Multiple Output
MMSE	Minimum Mean Square Error Estimation
MR	Maximum Ratio
N	Number
QoS	Quality of Service
RZF	Regularized Zero Forcing
SE	Spectral Efficiency
SINR	Signal to Interference plus Noise Ratio
SNR	Signal to Noise Ratio
TDD	Time Division Duplex
UE	User Equipment
UL	UpLink
ULA	Uniform Linear Array

Chapter 1

Introduction

Ever since the wireless communication has been prioritized as the means of communication for pretty much every field of our society, there has been a constant demand for increase in the quality of services provided while decreasing the drawbacks these systems face. The first generation of mobile communication was based on analog communication and provided voice only services. A new type of network was developed called the cellular network. All of the coverage area was divided into cells which were assigned different frequencies to operate. A user in a particular cell could use the resources available in that cell. The following generation saw a massive change in its techniques as it shifted to digital communication which provided encrypted voice and text services to its customers. This platform then improved in terms of transmission which let to services like video calling and mobile internet which started the massive demand in this sector. The fourth and fifth generation started using different multiplexing techniques which provided higher data rates to its users. All of this progress had its drawbacks in the form of expensive hardware, interference issues from adjacent cells, uneven data rates to its users, etc.
New technologies, such as massive multiple-input-multiple-output (MIMO) [1] are needed to greatly expand cellular capacity and enhance network spectral efficiency. This project builds on the idea that a cell-free massive MIMO architecture can be considered to improve network performance and ensure a consistent user experience across the entire network area, since it can eliminate poor cell-edge performance and increase the spectral efficiency of a channel [2].

1.1 Background

The wireless communication sector has seen a massive increase regarding the number of users being dependent on the available platform. This has also resulted in the service providers moving their services from an analog platform to a digital wireless platform. The bandwidth is a limited resource and must be used efficiently in order to handle such vast increase in the demand of wireless systems. Massive MIMO networks have been developed in order to tackle this problem by using the available bandwidth to serve multiple users simultaneously while reducing the issues like interference from adjacent cells which is experienced in a traditional cellular network system. This project focuses on the issue of Pilot Contamination caused in Cell-Free Massive MIMO networks due to multiple users sharing a same pilot sequence.

1.2 Problem

The Massive MIMO technology has been proved to increase the spectral efficiency of a given wireless communication system. However, there have been a few issues in this system near the cell edges. Users from one cell may experience interference from adjacent cell areas which can lead to packet drops and/or termination of communications. They may also experience low Signal to Interference plus Noise ratio (SINR) which may lead to lower data rates than somewhere inside the cell. As the wireless communication field is expected to grow further in terms of number of users and the number of services available, can this problem of such a system be tackled easily? And if so, How?

1.3 Purpose

The purpose of this project is to tackle the issue of Pilot Contamination in Cell-Free Massive MIMO networks by designing a pilot assignment algorithm and cluster formation to increase the achievable Spectral Efficiency of the channel. If the desired goals are achieved, this project can help all the mobile operators to reuse the pilots more efficiently and provide relatively high data rates to most of their users with high probability. It can also help in using the available bandwidth efficiently and increasing the scalability of the system. This increases the sustainability of the project as most of the resources are being used efficiently.

1.4 Hypothesis

The new system is expected to further increase the Spectral Efficiency (SE) of a wireless network system and solve the poor cell-edge performance in traditional cellular networks.

1.5 Goals

The goal of this project is to eliminate the Pilot Contamination caused in Cell-Free MIMO networks. This has been divided into the following sub-goals:
1) Build a system level simulator for a cell-free massive MIMO network.
2) Implement existing pilot assignment and cluster formation algorithms.
3) Develop a new algorithm and compare results.

1.6 Research Methodology

In the project, the empirical method [3] will be used due to time constraints preventing a detailed mathematical analysis. Empirical research relies on real-world experiences and observations to gain knowledge and test predictions, focusing on people and events. It involves collecting and analyzing data to understand relationships between various factors, leading to the development of knowledge and well-formed theories. Empirical research seeks knowledge through evidence from experiments, observations, or experiences, with data analyzed using quantitative or qualitative methods to explain inherent situations.

1.7 Delimitations

This project will focus on developing a new algorithm for the Pilot Assignment and cluster formation in Cell-Free MIMO networks to reduce the effects of Pilot Contamination. Simulations of a static environment will be performed on MATLAB for channel estimation and pilot assignment. No actual data transmission will be performed. Either a new pilot assignment algorithm or a new cluster formation technique will be developed in order to increase the performance of a system and reduce the pilot contamination issue due to time constraints.

1.8 Ethical considerations and sustainable development

In the development and implementation of pilot assignment and clustering formation algorithms for cell-free massive MIMO networks, there are a few sustainability aspects to address. In this project, it is expected that the investigated algorithms will contribute in the reduction of the power consumption at the APs, thereby minimizing the ecological footprint of the network infrastructure. In addition, the prospect of increasing the spectral efficiency implies that more data can be transmitted over the available bandwidth using less energy. This also results in reduced power consumption in network operations, contributing to overall energy efficiency and lower carbon emissions.

Regarding ethical considerations, this project does not directly associate with any, except for the profound ones that relate to privacy and security concerns that wireless network design entails. However, our contribution is not expected to address these issues, since our work aims to optimize the physical layer network without directly delving into privacy and security challenges.

1.9 Structure of the report

The first section of this report contains the introduction of this project. It discusses the problems of wireless system being used for mobile communication and the purpose/goals of this project. The next section dives into the details of the cellular system and MIMO system. This gives an insight of the background of this project. The next section named 'Method' discusses the approach of this project. It discusses all the concepts and methods being used to solve the problems of the previous communication system. The next section analyses the results of the project and compares it to that of the previous work done by different people to tackle the same issues. The work of this report is then concluded in the next section and the future work/scope is discussed to further improve the project.

Chapter 2

Background

Ever since the communication system has been developed, the connectivity has increased between people living in different regions of the world. The means of communication were wired which consisted of two telephones physically connected to each other using wires. This type of communication had some drawbacks which were solved by inventing wireless transmission systems. Radio waves were used to transmit a message from one place to the other. This type of propagation had some drawbacks like signal fading and introduction of noise through the channels. This system saw an increase in the number of users in a short span of time. Cellular networks were introduced to tackle this issue.

2.1 Cellular Networks

The bandwidth is a limited resource and needs to be used efficiently as the number of users keep increasing day by day. The cellular networks were designed to increase the number of users to be served in a given area. An area was divided into cells which were allocated a particular set of frequencies from the available bandwidth to operate and reduce interference.

2.1.1 Network Architecture

The area where service needs to be provided was subdivided into smaller cells which were assigned a specific set of frequencies. A specific Access Point (AP) called as the Base Station served the users present in its specified area. A specific set of frequencies were assigned to a particular cell which helped to utilize the available bandwidth. Adjacent cells were assigned different frequency sets in order to reduce the interference caused by using the same frequency to transmit/receive a signal. These frequency sets were being reused throughout

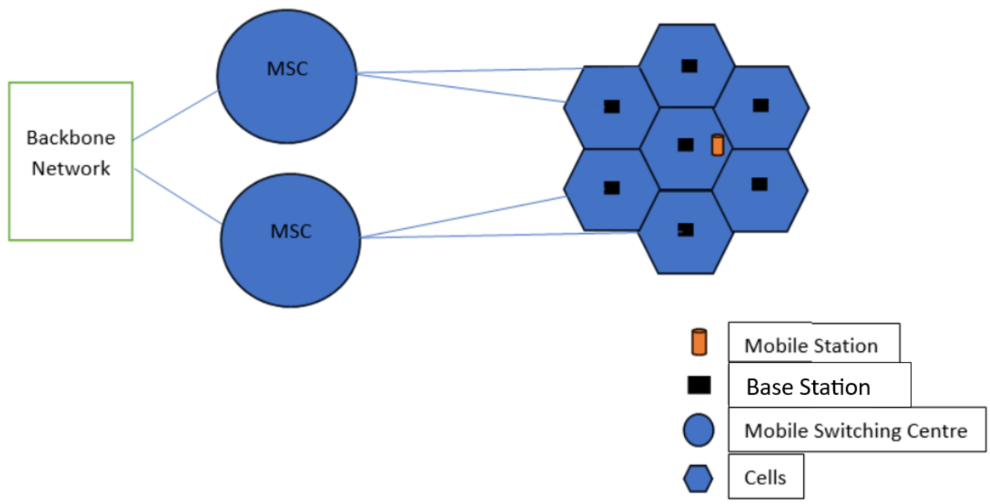


Figure 2.1: A basic Cellular System

different cells given the fact that no two adjacent cells used the same sets of frequencies for signals.This method is known as 'Frequency Reuse' method This technique helped to deal with the increasing number of users by assigning same sets of frequencies throughout the area thus, increasing the spectral efficiency. This system also had its drawbacks. The users near the border of a cell experienced interference from adjacent cells. The users near the AP had the best data rates and the users which were away from the AP had the lowest data rates. This caused the overall data rates to decrease in the system.

2.1.2 Mobility Management

Another drawback of the cellular system was the hand-off that takes place when one user equipment travels from one cell to the other. This means that when a user equipment was actively using the services while travelling, the base station of that cell had to hand-off that particular UE to the next cell base station. This is a seamless process when the new base station has the same available frequency channel that is being used by the UE in the previous cell. If this channel was not available, the UE would experience some packet drops in the form of delay or the service would be terminated if the channel was not available in a given time frame. These were the main issues of the cellular system that needed to be addressed. The service providers started using the wireless system to provide their services to the customers as it was a better means of transferring data. The constant increase in the number of users and the service providers shifting their services to wireless means lead to a new technique of wireless transmission technique known as Wireless Massive Input Massive Output (MIMO) Networks [2].

2.2 Multiple Input Multiple Output (MIMO)

Multiple Input Multiple Output (MIMO) refers to a data transmission technology which uses a large number of antennas to serve multiple users at the same time. The users transmit and receive data at the same channel [1]. This technique was developed to increase the number of users that can be served efficiently on a given channel bandwidth. The users were at the same frequency and time domain but separated at the spatial domain. This helped to utilise the available bandwidth more efficiently thus, increasing the spectral efficiency. MIMO technology was combined with the cellular system to tackle the issue of increase in the demand of wireless networks.

Multipath is a phenomenon that occurs when a radio wave signal bounces off of different objects in the environment like buildings, trees, vehicles, etc when travelling to the receiver end. This was previously viewed as an issue as these different paths lead to interference in the transmitted signal. MIMO technology uses multiple transmitters and receivers which use multipath propagation and add a spatial dimension to these signals which increases the range and the data rates of these signals. In this manner, MIMO technology was introduced to the cellular networks to increase the data rates and to accommodate the increasing number of users. However, cellular network experienced an interference to its users at the edge of its cells which lead to lower spectral efficiency and non-uniform data rates all over the covered area. This was needed to be addressed and resolved as this technology developed. Cell-free MIMO technology was developed to tackle this problem. Access points with multiple antennas were used to serve its users. All APs served all the UEs in a traditional Cell-free MIMO network. This ensured that the interference issue caused at the border of a cell was eliminated. It improved the spectral efficiency of the system and increased the overall data rates. User centric approach was developed in which only a small number of APs served a particular UE. This approach allowed to reduce the CPU backhaul computations. This approach also provided better performance in terms of scalabilty, spectral efficiency and higher data rates.

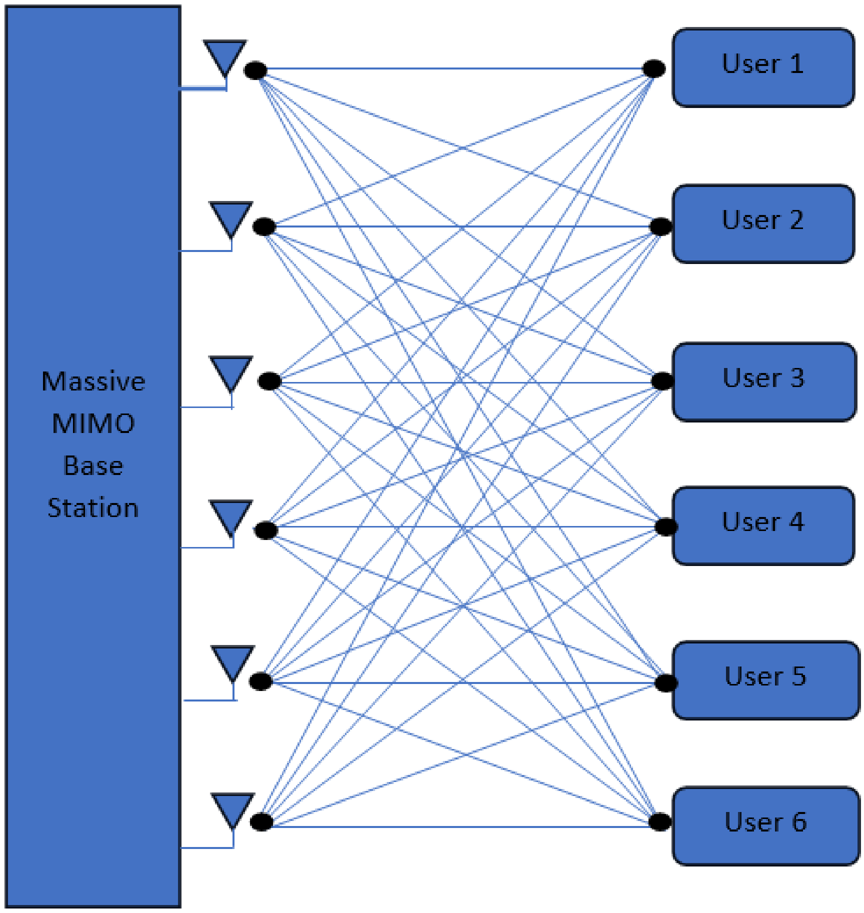


Figure 2.2: A basic Multiple Input Multiple Output System

2.3 User-centric Cell Free Massive MIMO

In traditional MIMO networks, an AP has a number of antennas which serve all the users. The number of APs in this system is much greater than the number of User Equipment (UE). All the APs serve all the UEs in a given area. All the UEs transmit and receive their data in the same bandwidth at the same time. MIMO network is a multi-carrier network which consists of a number of cells which use synchronous TDD for their data transmission.

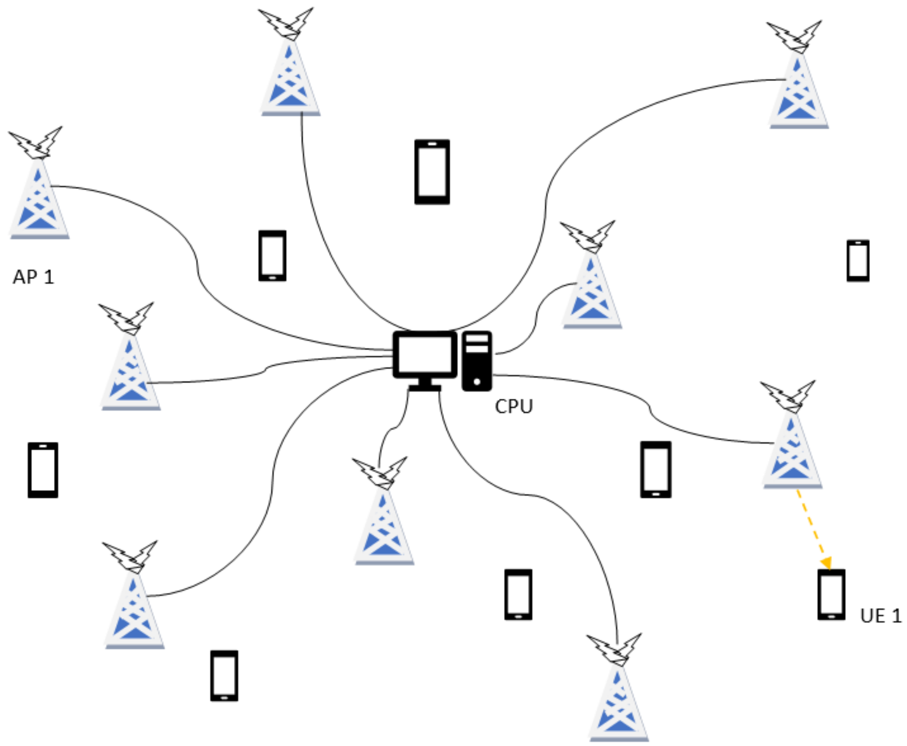


Figure 2.3: A Cell-Free Massive MIMO Network

In a MIMO network system, channel estimation primarily refers to the process of estimating the characteristics of the wireless communication channel through which the signals propagate between the Access Points (APs) and the User Equipment (UE). It involves deducing the conditions and properties of the channel, including signal quality, attenuation, delay, phase shifts, and other factors that impact the transmission of data. This part aims to shed light on the significance of channel estimation, the choice of Minimum Mean-Squared Error (MMSE) estimation, and the methodology involved in channel estimation. The Channel Estimation should be done accurately for the following reasons:

Signal Quality and Reliability: It directly affects the quality and reliability of the received signal. By understanding the channel's characteristics, we can mitigate issues such as fading, interference, and noise, thus enhancing signal quality.

Spectral Efficiency: Effective channel estimation leads to improved spectral efficiency, enabling the transmission of more data within the available bandwidth. [2]

2.3.1 Pilot Assignment

Pilot sequences are orthogonal bit sequences, that are transmitted by the UE to the AP (uplink pilots) in order to provide a reliable way to estimate the state of a given channel [4]. Due to the limitations imposed by the tradeoff between computational overhead and performance, the need to reuse these pilot sequences has emerged. This happens because there is a limited number of bits that can be used, without increasing the computational and transmissional overhead too much [5]. This means that one pilot sequence might need to be assigned to more than one user in a given topology, because it is highly likely that the total number of pilot sequences that can be used is lower than the number of UEs that need to be served at a given timepoint. As a result, the issue of pilot contamination is introduced, where pilot sequence transmissions are no longer orthogonal, because they are being reused [5]. How can some of the negative effects of pilot contamination, like the drop in overall network performance, be alleviated? There is extensive literature surrounding the topic of how one can deal with pilot contamination/reuse; a lot of different algorithms have been proposed such as, the Dynamic Power Reuse algorithm (DPR) [5] , the Scalable Pilot Assignment algorithm [6] and the Weighted-Counting Pilot Assignment algorithm [7]. As mentioned previously, these algorithms can only alleviate the effects of pilot contamination and do not completely eliminate them; they do not provide the performance of orthogonal pilot use, but they can increase the maximum achievable uplink sum-rate of a given wireless system, when compared to a random reuse of pilot sequences [5].

2.3.2 Clustering

In addition to the Pilot Assignment Algorithms described above, AP/UE Clustering is another technique being implemented to increase the overall performance of Massive MIMO networks (especially for cases of UEs located near a cell's edge), by replacing the standard cellular configuration, with a Cell-Free approach [8]. However, when Cellular networks are not used, then there needs to be a way to decide which APs are serving which UEs. This is mainly done using a metric like the average channel gains of each UE-AP, based on which the UE forms a cluster of APs that are best suited to serve that particular UE. The main goal that clustering aims to achieve is to make a solution scalable in the real world, because it would not be realistic to assume that all UEs are served by all the APs in a given network, while also increasing edge case performance. AP Clustering can also be used in tandem with a Pilot Assignment algorithm, in order to reduce pilot contamination. This way, better performance can be achieved as demonstrated in the work of Zhong, Zhu and Lim [9].

2.3.3 In essence

In essence, the performance of a network is best, when pilot assignment and clustering are working in tandem with each other. This way, we can ensure the biggest available degree of separation between 2 UEs that are using the same pilot sequence (by guaranteeing for example that they are served by different APs and avoiding to reuse the same pilot sequence within one cluster), thus reducing the effect of pilot contamination and improving the overall spectral efficiency of the network.

2.4 Related Work

The reference [9] proposed an algorithm for this Pilot Assignment among the users in a way that can increase the SE and the scalability of a given system. Both centralized and distributed methods have been used to evaluate the performance of the proposed algorithm. It is assumed that each AP can serve a limited number of UEs and the UEs have to compete for the APs based on the channel conditions. As the number of Pilot sequences are limited, the UEs with better channel conditions are assigned the pilots orthogonally. Assuming these UEs are K, the K+1th UE will be assigned the pilot sequence from the limited sequences which introduces the least amount of interference to its serving APs. This is done for each UE left after assigning the sequences. This ensures the reuse of the pilots in an efficient way. This algorithm helped increase the Spectral Efficiency (SE) and the scalability of the available channel in both centralized and distributed approach according to the paper " Scalable Pilot Assignment for User-Centric Cell-Free Massive MIMO Networks by Zhihan Ren, Angela Doufexi, Mark A. Beach from Department of Electrical and Electronic Engineering, University of Bristol, Bristol, UK" [9].

A cell free massive MIMO system experiences a few difficulties including the crucial channel training overhead in a dense network. A Dynamic pilot reuse system is proposed in order to allow two users to share the same pilot sequence[6]. This pilot reuse scheme also ensures that the maximum sum-rate is obtained while considering the signal to interference plus noise ratio (SINR) requirements of users are also fulfilled. The SINR for a user sharing a pilot sequence is calculated by using minimum mean squared error detection (MMSE) and channel estimation. A simple pilot algorithm is developed using iterative grid search (IGS) to decide the users that will share the same pilot sequence based on the degree of separation (distance) between them. The simulation results proved that this technique provided better sum-rate performance as compared to other pilot assignment algorithms.

A cell free massive MIMO system is being considered which consists of APs with multiple antennas ans UEs with single antennas. A new pilot assignment algorithm based on the weighted count is proposed to reduce the pilot contamination issue [7]. This algorithm considers the previous geographic information of a user to maximise the weighted distance for pilot reuse. It also considers pilot power control and AP power control in the pilot training phase and downlink phase simultaneously. The simulation results prove that this method can improve the system performance significantly.

Cell-free massive MIMO networks have been suggested as a way to suppress inter-cell interference by eliminating cell boundaries while providing uniform performance throughout the coverage area. However, there is a question of how to form clusters that provide better performance. User-centric clusters have been the best candidates for cluster formation that guarantees good quality of service (QoS) regardless of the cluster size while solving scalability issues; see C.F. Mendoza, S. Schwarz, and M. Rupp [8].

A proper pilot assignment is critical in cell-free massive MIMO networks as it is used to acquire perfect channel state information (CSI). However, challenges arise when it comes to pilot reuse as it leads to pilot contamination, which reduces the quality of channel estimation and makes it harder to reject interference between pilot-sharing UEs. As shown by E. Björnson in their paper “Structured Massive Access for Scalable Cell-Free Massive MIMO Systems” [10], two pilot assignment schemes, namely user-group and interference-based K-means (IB-KM), perform better in suppressing the mutual interference from the pilot sharing UEs.

Chapter 3

Method

In this chapter, an analysis of the system model as well as the investigated and novel algorithms is presented. We begin with a detailed description of our system model that entails the network architecture and channel modeling. Next, we continue by breaking down the channel estimation method, precoding schemes, power allocation method and spectral efficiency calculation method employed in our project. Consequently, we describe the modifications that were applied on an existing pilot assignment algorithm presented in [6] to produce one of the algorithms that we investigated in this project. Finally, we end this chapter by elaborating on novel ideas that can be exploited in the algorithm design, by including the ubiquitous k-means clustering algorithm. Before delving into the technical parts of our method, it is crucial to reming our readers that regarding our research methodology, the empirical research method, described in 1.6, is utilized in our project.

MATLAB has been used to perform simulations in all of our phases. A basic scenario containing K number of users and N number of APs has been created. The wrap-around technique has been used to eliminate the cell-edge problems.

3.1 System Model

3.1.1 Network Architecture

We consider a wireless network in a specified geographical area of certain size with K geographically distributed UEs, each equipped with one antenna and L geographically distributed APs, each equipped with N antennas. In the centralized approach, the APS are jointly serving UEs. However, in the User-Centric approach, each AP is serving a subset of UEs.

A wireless channel is regarded to be a linear but time-variant system [2], making it difficult and nearly impossible to estimate its properties. However, we can consider a short time interval, in which the channel remains time-constant. In addition, we can also consider a frequency range in which the channel's frequency response remains constant. By combining the time and frequency intervals, we can consider a time-frequency block, called the *channel coherence block*, in which the channel between is constant and frequency-flat, meaning that it can be described by only one scalar coefficient. That way we can estimate the channel properties in each coherence block.

Taking into account the standard Cellular Massive MIMO TDD protocol, in each coherence block there are τ_c transmission symbols comprised by τ_p symbols for uplink pilots, τ_u symbols for uplink data and τ_d symbols for downlink data. In our project, we only examine the downlink, which is why $\tau_u = 0$.

For the communication to be viable, there is a need for the UEs to transmit pilot signals towards the APs, so that the APs can use them as reference signals, or "common knowledge" signals. By exploiting the pilot signals, the APs can perform an improved channel estimation, optimize their precoding schemes and transmit the downlink data more efficiently toward the UEs. The APs are also capable of determining the power that is going to be used for the transmission towards one specific UE.

For our project, as a last step in our system, we estimate the spectral efficiency calculation at the AP to characterize the performance of our system. In Fig. 3.1, a flowchart of how our system works is illustrated. Note, that the names in the blocks also correspond to how our network simulator functions work.

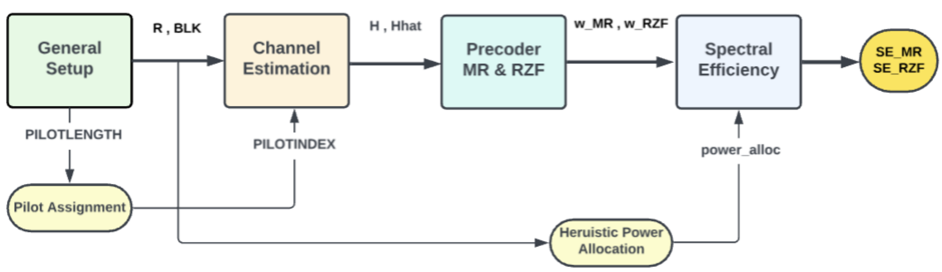


Figure 3.1: System flowchart

3.1.2 Channel Modeling and Channel Estimation

In this section, we delve into the technical aspects of our wireless channel modeling and the channel estimation method that is used from the APs.

Wireless communication channels are affected by various phenomena that introduce fluctuations in the signal characteristics. By modeling these phenomena, we can estimate the channel's properties. We describe those signal variations with the term "fading". There are two main fading categories: Large-scale fading and small-scale fading [11].

Large-scale fading refers to the long-term changes in signal strength due to the distance between the transmitter and the receiver or obstructions in the large environment. One key factor to consider in large-scale fading is path loss. As a signal travels over a distance, it spreads out and loses energy, resulting in attenuation in signal strength. The path loss is influenced by the distance between the transmitter and receiver and the frequency of the signal. The second key factor is shadow fading, which refers to the signal being "shadowed" by obstacles such as buildings, trees, and hills. These obstructions cause the received signal strength to vary slowly over relatively large distances. Shadow fading is often modeled as a random process, with the received signal strength varying according to a log-normal distribution. Our investigation incorporates the 3GPP Urban Microcell model, described in [12, Table B.1.2.1-1] addressing path loss, shadowing, and obstacles in urban environments. The model is given in eq. 3.1.

$$\beta_{kl}[dB] = -30.5 - 36.7\log_{10}\left(\frac{d_{kl}}{1m}\right) + F_{kl} \tag{3.1}$$

, where d_{kl} is the 3D distance between a transmitter k and a receiver l in meters and $F_{kl} \sim \mathcal{N}(0, 4^2)$ is used to describe the shadow fading.

Small scale fading refers to rapid fluctuations in signal strength over short distances or time durations. This type of fading is caused by the constructive and destructive interference of multiple signal paths, which is called multipath propagation. Two typed of small-scale fading are relevant to our project: Uncorrelated Rayleigh and Correlated Rayleigh Fading.

The uncorrelated Rayleigh fading model describes the case when there is no dominant signal path (like a direct line-of-sight between transmitter and receiver), and instead, the received signal is the sum of multiple reflected, scattered, and diffracted signals. When numerous signals, each with random phases, add up, they produce a statistical phenomenon that follows a Rayleigh distribution.

The correlated Rayleigh fading model is similar to uncorrelated Rayleigh fading, but the multiple paths between the transmitter and receiver aren't all independent. This could be due to the movement of the receiver, causing some paths to change in correlated ways. More importantly, the correlated Rayleigh model can be utilized to capture the characteristics of spatial correlation, a phenomenon that more accurately describes practical systems [2]. Spatial correlation arises in such systems from two factors:

1. Transmission direction towards the UE is affected by spatial directions.
2. The geometry of the antenna array (i.e., shape and antenna spacing) impacts the decision for a suitable transmission direction.

By using the correlated Rayleigh fading model, we can describe the channel between AP l and UE k by using eq. 3.2:

$$\mathbf{h}_{kl} \sim \mathcal{N}_{\mathbb{C}}\left(\mathbf{0}_N, \mathbf{R}_{kl}\right) \tag{3.2}$$

, where $\mathbf{R}_{kl} \in \mathbb{C}^{N\times N}$ is the spatial correlation matrix between AP l and UE k.

We then define the large-scale fading coefficient as the average value given in eq. 3.3. This represents the average channel gain between an antenna at AP l and UE k.

$$\beta_{kl} = \frac{1}{N} tr\left(\mathbf{R}_{kl}\right) \tag{3.3}$$

The last term we need to mathematically express is the spatial correlation matrix that is dependent on the array geometry and the angular distribution of the multipath components. We consider that the APs are using a Uniform Linear Array (ULA) where the N antennas are equally spaced on a horizontal line. We compute the (m,l)th element of a generic spatial correlation matrix $\mathbf{R}$ in the eq. 3.4.

$$[\mathbf{R}]_{m\ell} = \beta e^{\mathrm{j}\pi(m-\ell)\sin(\bar{\varphi})\cos(\bar{\theta})} f(\bar{\varphi},\bar{\theta}) d\bar{\varphi} d\bar{\theta} \tag{3.4}$$

, where β is the common large-scale fading coefficient, $\bar{\varphi}$ is the azimuth angle and $\bar{\theta}$ is the elevation angle of a multipath component computed with respect to the broadside of the ULA and $f(\bar{\varphi},\bar{\theta})$ is the joint probability density function (PDF) of $\bar{\varphi}$ and $\bar{\theta}$ [1]. For the PDF, we utilize the joint Gaussian distribution given in 3.5.

$$f(\bar{\varphi},\bar{\theta}) = \frac{1}{2\pi\sigma_{\varphi}\sigma_{\theta}} e^{-\frac{(\bar{\varphi}-\varphi)^2}{2\sigma_{\varphi}^2}} e^{-\frac{(\bar{\theta}-\theta)^2}{2\sigma_{\theta}^2}} \tag{3.5}$$

3.1.2.1 Minimum Mean Squared Error (MMSE)

In User-Centric Cell-Free Massive MIMO systems, the MMSE (Minimum Mean Squared Error) estimator is the top choice for channel estimation. It excels at minimizing errors in channel estimates when you have complete statistical knowledge of the channel. This knowledge includes details about the correlation matrix, noise characteristics, the use of orthogonal pilots, and pilot index information.

The implementation of the MMSE estimator is demonstrated through the provided MATLAB code. The code generates uncorrelated Rayleigh fading channel realizations, applies spatial correlation matrices, and performs channel estimation. The following key components of MMSE estimation are addressed in the code:

- Uncovering uncorrelated channel realizations.
- Applying correlation matrices to uncorrelated channel realizations.
- Utilizing noise statistics.

By optimally weighing these statistical elements and mitigating noise and interference, along with the linear estimation technique the MMSE estimator provides accurate channel estimates.

3.1.2.2 Channel Estimation Methodology

The process of channel estimation in User-Centric Cell-Free Massive MIMO systems consists of a few steps:

- **Uplink Pilot Transmission:** The system model during uplink pilot transmission is defined, wherein User Equipment's (UEs) transmit pilot sequences within designated coherence blocks. These pilots are used to estimate the channels. Each UE is assigned a random τ_p-length pilot from a set of τ_p mutually orthogonal pilots utilized by the APs. Let $t_k \in \{1, \ldots, \tau_p\}$ denote the index of the pilot assigned to UE k.

  After correlating the received signal at AP l with pilot t_k for the estimation of the UE k channel, the signal $y^p_{t_k l} \in \mathbb{C}^{N \times 1}$ is obtained as:

  $$y^p_{t_k l} = \sum_{i=1, t_i = t_k}^{K} \sqrt{\tau_p p_i} h_{il} + n_{t_k l} \tag{3.6}$$

  ,where:

  p_i is the transmit power of UE i, and

  $n_{t_k l} \sim \mathcal{N}_{\mathbb{C}}(0, \sigma^2 I_N)$ represents the additive Gaussian noise vector at AP l.

- **MMSE Estimation:** The MMSE estimator is applied to remove interference and noise from the received pilot signals. The minimum mean-squared error (MMSE) estimation of h_{kl} is the vector $\hat{h}_{kl}$ that minimizes the MSE

$\mathbb{E}\{\|h_{kl} - \hat{h}_{kl}\|^2\}$. Here we are trying to obtain an approximate value of h_{il} by observing $y^p_{t_k l}$, hence the estimated channel $\hat{h}_{kl}$ for UEs using pilot indexed t_k is given by:

$$\hat{h}_{kl} = \sqrt{\tau_p p_k} R_{kl} \left(\sum_{i=1, t_i = t_k}^{K} \tau_p p_i R_{il} + \sigma^2 I_N \right)^{-1} y^p_{t_k l} \tag{3.7}$$
$$\sim \mathcal{N}_{\mathbb{C}}\left(0, \tau_p p_k R_{kl} \Psi_{kl}^{-1} R_{kl}\right)$$

,where:
$\Psi_{kl} = \sum_{i=1, t_i = t_k}^{K} \tau_p p_i R_{il} + \sigma^2 I_N$ represents the correlation matrix of the received pilot signal.

Channel estimation in User-Centric Cell-Free Massive MIMO systems faces challenges, including interference from UEs sharing pilots and spatial correlation which leads to Pilot Contamination, Handling Imperfect Statistical Knowledge. These challenges need to be addressed for accurate estimation.

3.1.2.3 Impact of Spatial Correlation on Channel Estimation

Spatial correlation has a profound impact on the quality of channel estimates. When channels exhibit spatial correlation, estimation error variances are reduced, leading to improved estimation quality. Higher Signal-to-Noise Ratios (SNRs) make the estimation of strong eigen directions more accurate.

3.1.2.4 Impact of Pilot Contamination on Channel Estimation

Pilot contamination, resulting from the shared use of pilot sequences by multiple UEs, can introduce correlated channel estimates and reduce estimation quality. This effect is dependent on UE angles and signal strengths. However, spatial channel correlation can mitigate these effects, enhancing system efficiency.

3.1.3 Precoding

We use precoders to boost the signal to interference and noise ratio (SINR), which raises the data rates and enhances the performance of our communication systems. In this project, we choose to employ the Maximum Ratio (MR) and Regularized Zero Forcing (RZF) precoders as explained below.

- Maximum Ratio Combining (MR): MR combing is a linear transformation of data symbols, which involves the multiplication of data symbols by a set of precoding vectors, which are derived from the channel state information

(CSI). The coding scheme has advantages in suppressing noise but not actively suppressing the interferences. The normalized MR precoding vector is defined as:

$$W_{il} = \frac{\hat{h}_{kl}}{\|\hat{h}_{kl}\|} \tag{3.8}$$

,where $\hat{h}_{kl}$ is the estimated channel.

- Regularized Zero Forcing (RZF): RZF precoding combines zero forcing with regularization to mitigate multi-user interference. RZF thereby enhances system performance by optimizing the trade-off between interference cancellation and noise amplification, particularly in our scenarios with cell free massive MIMO communication system. We define the normalized RZF precoding vector as:

$$W_{il} = \frac{\left(\sum_{i=1}^{k} p_i \hat{h}_{il} \hat{h}_{il}^H + \sigma^2 I_N\right)^{-1} p_k \hat{h}_{kl}}{\left\|\left(\sum_{i=1}^{k} p_i \hat{h}_{il} \hat{h}_{il}^H + \sigma^2 I_N\right)^{-1} p_k \hat{h}_{kl}\right\|} \tag{3.9}$$

,where $\rho_i \geq 0$ is the transmit power allocated to UE i.

Note: In both precoders, the considered precoding vectors $\{W_{il}\}$ satisfy short-term power constraints, meaning that $\|W_{il}\|^2 = 1$ in each coherence block, and not on average.

3.1.4 Power Allocation

Heuristic power allocation is a method used in wireless communication systems to efficiently distribute downlink power to User Equipments (UEs) connected to Access Points (APs). It leverages channel gain information and a scaling parameter to allocate power proportionally among UEs, ensuring fairness and optimizing spectral efficiency.

`Heuristic_Power_Allocation` allocates downlink power (ρ_{pp}) to each user (UE) connected to a specific access point (AP). It ensures fairness in power distribution among UEs based on their channel gains (b_{kl}) and a scaling parameter (v). The power allocated to each UE is determined as a fraction of the total available power (P_{max}) and is calculated using the formula given in eq. 3.10:

$$\rho_{pp} = P_{max} \cdot \left(\frac{b_{kl}^{v}}{\sum_{k} b_{kl}^{v}} \right) \tag{3.10}$$

This formula calculates the power allocated to a UE by taking the ratio of its channel gain to the sum of channel gains for all UEs connected to the same AP, scaled by the maximum available power.

The MATLAB function `Heuristic_Power_Allocation` embodies this allocation strategy, efficiently distributing downlink power among UEs connected to APs. It prioritizes fairness by ensuring a balanced distribution of available power based on channel gain information and the defined scaling parameter.

3.1.4.1 Algorithm Overview

The heuristic power allocation algorithm operates as follows:

- **Input Preparation**: The algorithm processes Large-Scale Fading (LSF) coefficients using a heuristic scheme, enhancing dynamic range and distribution for better performance during power allocation.
- **Power Allocation**: It computes power coefficients using the heuristic formula given in eq. 3.10 that balances power among UEs connected to specific APs based on their LSF coefficients and a scaling parameter.
- **Standardization**: The dataset undergoes standardization to reduce outliers, improving the algorithm's robustness.

Heuristic power allocation plays a pivotal role in optimizing downlink power distribution in wireless communication systems. By considering channel gains and a scaling parameter, it ensures fair power allocation among UEs connected to APs, contributing to enhanced spectral efficiency and efficient resource utilization.

3.1.5 Spectral Efficiency Calculation

Spectral Efficiency (SE) serves as a fundamental metric in wireless communication systems, characterizing the efficiency of data transmission over a given bandwidth. It quantifies the rate at which information can be reliably transmitted through the available frequency spectrum. In essence, SE represents the capacity to convey information reliably and efficiently within a specific channel or bandwidth allocation. A crucial aspect influencing spectral efficiency is the ability to mitigate interference and noise while maximizing the desired signal strength. The spectral efficiency metric becomes paramount in evaluating the system's ability to deliver higher data rates reliably within the available frequency bands. Mathematically, SE is determined by the Shannon-Hartley theorem and is typically expressed as in eq. 3.11:

$$\mathrm{SE}_k^{(\mathrm{dl,d})} = \frac{\tau_d}{\tau_c} \log_2\left(1 + \mathrm{SINR}_k^{(\mathrm{dl,d})}\right) \quad \mathrm{bit/s/Hz} \tag{3.11}$$

, where $\frac{\tau_d}{\tau_c}$ is the prelog factor, which represents the efficiency of updating channel conditions, considering data transmission time τ_d and channel coherence time τ_c . As discussed later, a prelog factor of 0.9 has been set for our project.

Additionally, the $\mathrm{SINR}_k^{(\mathrm{dl,d})}$ is the Signal-to-Interference-plus-Noise Ratio (SINR) and is calculated using 3.12:

$$\mathrm{SINR}_k^{(\mathrm{dl,d})} = \frac{\left|\sum_{l=1}^{L} \mathbb{E}\left\{\mathbf{h}_{kl}^{\mathrm{H}}\mathbf{D}_{kl}\mathbf{w}_{kl}\right\}\right|^2}{\sum_{i=1}^{K} \mathbb{E}\left\{\left|\sum_{l=1}^{L} \mathbf{h}_{kl}^{\mathrm{H}}\mathbf{D}_{il}\mathbf{w}_{il}\right|^2\right\} - \left|\sum_{l=1}^{L} \mathbb{E}\left\{\mathbf{h}_{kl}^{\mathrm{H}}\mathbf{D}_{kl}\mathbf{w}_{kl}\right\}\right|^2 + \sigma_{\mathrm{dl}}^2} \tag{3.12}$$

, where the term $|\sum_{l=1}^{L} \mathbb{E}\left\{\mathbf{h}_{kl}^{\mathrm{H}}\mathbf{D}_{kl}\mathbf{w}_{kl}\right\}|^2$ represents the power of the desired signal received at UE k, $\sum_{i=1}^{K} \mathbb{E}\left\{\left|\sum_{l=1}^{L} \mathbf{h}_{kl}^{\mathrm{H}}\mathbf{D}_{il}\mathbf{w}_{il}\right|^2\right\}$ represents the cumulative interference power affecting UE k and σ_{dl}^2 is the noise- power in the downlink.

MATLAB Function Overview:
The function `SE_calculation` computes the spectral efficiency for each user. It takes into account parameters such as the precoding vector, power allocation, channel conditions, and number of realizations. The function iterates through UEs and calculates the SINR and SE for each. **Algorithm Overview:**

- **Step 1 - Initialize Parameters:** Gather necessary input parameters.
- **Step 2 - Construct D Matrix:** Create a D matrix based on the serving APs for each UE.
- **Step 3 - Compute Signal and Interference:** Calculate the power of the desired signal received at each UE and compute interference from other UEs.
- **Step 4 - Calculate SINR and Spectral Efficiency:** Use the computed values to derive SINR and then compute Spectral Efficiency for each UE.

This function aids in determining the spectral efficiency for individual UEs in the system, considering their signal strengths, interference levels, and channel conditions, crucial in optimizing the overall system performance.

3.2 Gyros Pilot Assignment and Cluster Formation algorithms (GPA)

3.2.1 User-Centric Clustering

The implementation of clustering is based on the clustering described in reference 6, however it has been extensively modified, so it can better fit our experimental process and produce better results. In reference 6, clustering is done based on the distances between each UE and the APs. While this approach might be easier to implement, because the distance calculation can be obtained in a straightforward manner, we assume that it is not realistic for us to know the actual distances between each UE and the APs, thus we used the average channel gains (between each UE and AP).

By using the average channel gains as the metric, based on which clusters are formed (in this context a cluster is synonymous to the AP-UE assocations), we can make a more informed decision regarding the formation of the AP-UE associations, because we take into account the signal strength, instead of the distance, which is a more accurate way of determining which are the best APs to serve each UE. As a result, each UE forms a cluster with the APs for which this value is the highest. In our attempt to refine this process, we do not set a specific value for the size of the cluster, but instead we determine a maximum size of 7 APs in a cluster and then eliminate the APs, whose channel gains are less than 5% of the total channel gains (provided by all APs that are serving one particular UE). This approach is scalable and realistic, as it does not make unrealistic assumptions about knowing the exact location of each UE, while at the same time remaining user-centric, meaning that it is not constrained as the network grows bigger (because each UE selects which APs are going to serve them). The clusters formed here are also utilized in the pilot assignment process, as described in section 3.2.3.

3.2.2 Network-Centric (fixed) Clustering

Another clustering method that was implemented was a Network-Centric, or fixed, clustering method. The clusters formed using this method are independent of the UEs and they are simply created by grouping a certain number of APs together. Since the layout of our experiments follows a grid pattern that has a slightly random nature (in the sense that the APs are not completely aligned with each other, to form a perfect square, but they have a slight randomness in their placement), the decision was made to group each row of APs into a cluster, thus forming a total of number of clusters equal to $\sqrt{APs}$.

The UEs are choosing to be served by the group of APs that has the highest total sum of channel gains to that particular UE (we get this metric by adding all of the channel gains from each AP in that group to that particular UE). This method is not user-centric, but centralized by nature, thus it might be easier to implement (and scale up), but the expected performance is low, due to its static nature, that does not take into consideration the specific channel conditions of each UE, or whether an AP is beneficial to the service received by a UE or not (as UEs are essentially forced to be served by a specific group of APs and cannot choose to omit any of them, or add any to that group).

3.2.3 Pilot Assignment based on additional interference

Pilot assignment is implemented based on the scalable pilot assignment algorithm described in [6], which is then combined and integrated with the clustering process described above. On a high-level analysis, the algorithm tries to reuse pilot sequences by primarily avoiding to share a pilot between two UEs that are served by at least one common AP and secondarily, by taking into account an interference metric that is calculated based on the following formula (This interference metric is only calculated between each UE and its serving APs):

However, the final decision regarding the pilot that is assigned to each UE follows a more complicated process than just picking the pilot that has the best metric calculation. In a nutshell: if there is only one pilot sequence that can guarantee the degree of separation that we set (no common APs serving the two UEs that are sharing this pilot), then we pick this pilot sequence. Else, if there are more than one pilots that satisfy this degree of separation, then we pick the one among them that has the lowest score in the additional interference metric, specified above. Finally, if pilot sharing between two UEs that are served by at least one common AP is inevitable (meaning we cannot guarantee the degree of separation that we set), then the algorithm selects the pilot, which has the lowest score in the interference metric. We have improved the algorithm slightly, so it can randomly choose a pilot in any of the above cases, if the metric is not a tiebreaker and that way, we are not assigning the initial pilots more times than the ones close to the end of the array (thus ensuring a more uniform pilot distribution).

The process that has been implemented involves more steps than just calculating the interference metric and making a decision based on that, which is what is proposed in [6].

3.3 Novel Pilot Assignment and Cluster Formation Algorithms

3.3.1 K-means user clustering

K-means clustering is an unsupervised machine learning algorithm used for partitioning a dataset into K distinct, non-overlapping clusters. It's widely employed in various fields, including data mining, pattern recognition, and image analysis. The primary goal of K-means is to group similar data points together and discover intrinsic patterns within the data.

The following algorithm overview demonstrates the implementation of the K-means clustering algorithm using MATLAB on UE coordinates. It initiates by setting up the required parameters and randomly selecting initial centroids from the data points. The iterative process begins by assigning UEs to the nearest centroids, updating centroids based on assigned UEs, and monitoring centroid changes to check for convergence.

Once the centroids converge, it proceeds to assign UEs to clusters based on their proximity to centroids while considering the available space within each cluster. Finally, it displays the indices of UEs within each cluster and employs a simple algorithm for pilot assignment using the obtained clusters.

This algorithm aims to group similar UEs together and assign them into clusters based on proximity, facilitating further analysis or actions, such as pilot assignment in a communication network setup.

3.3.1.1 Algorithm Overview

Step 1 - Initialize Data:

- Gather coordinates of User Equipment (UE) data.
- Determine the number of UEs, clusters, and necessary parameters.

Step 2 - Initial Centroids:

- Select random data points as initial cluster centroids.

Step 3 - K-means Iteration:

- Calculate distances of UEs from centroids and assign each UE to the nearest centroid.
- Update centroids by calculating means of UEs within each cluster.
- Repeat until centroids stop changing significantly (convergence).

Step 4 - Cluster UE Assignment:

- For each UE, find the cluster with the closest centroid and available space for assignment.
- Assign the UE to the chosen cluster.

Step 5 - Cluster Display:

- Display the indices of UEs within each cluster.

Step 6 - Pilot Assignment:

- Execute the pilot assignment algorithm in section 3.3.2 using the obtained clusters.

3.3.2 Pilot Assignment based on the K-means clusters

After distributing the UEs to their corresponding k-means clusters, what is left is to distribute the pilot signals among them, while avoiding pilot reuse within the same k-means cluster. This is done initially, by randomly distributing the pilots within the clusters and then running a modified version of our algorithm from 3.2.2, based on which we attempt to redistribute them more efficiently. To avoid pilot reuse (within the same k-means cluster), we swap the pilot signals of two users with each other after comparing the theoretical reduction in interference. By randomly distributing the pilots first, we assign more weight to the k-means cluster decision of pilot distribution and not the metric based on which we redistribute them. This is the case, because our simulation cannot infinitely keep trying to reassign the pilots to achieve a final, optimized, converged state because: a) this state might not even exist (because of the randomness in all of the parameters) and b) this is not practical, as it would require a much longer simulation run time and would not give us a realistic final result.

Chapter 4

Analysis

In this section, the simulation results alongside with the analysis and interpretation are presented. We first present our system setup and then the results for the existing pilot assignment and clustering formation algorithms as well as for the novel algorithm that we developed are jointly presented.

Our main focus is to analyze the Cumulative Distribution Function (CDF) of the downlink spectral efficiency per UE. The CDF can indicate the performance of the system and give prominence to worse and best served UE cases. In addition, we inspect the average spectral efficiency as a function of the number of UEs to see how the latter affects the performance of the different algorithms. In order to quantify the fairness gains, and hence gain more thorough insight on the effect of the algorithms on user experience, we also consider the 90%-, 50%- and 10%-likely SE value. For instance, the 90%-likely SE value, i.e., where the CDF is 0.1, represents the SE that can be provided to 90% of all UEs.

Furthermore, we aim to compare the two precoding schemes that were investigated in our project, namely MR and RZF precoding. For that reason, we present cases where the main focus is the performance difference between the two.

4.1 Simulation Setup

Our simulation setup is given in Table. 4.1. We consider a network of an area size of 1000m x 1000m where the APs and UEs are uniformly and randomly distributed. The APs are equipped with $N = 4$ antennas each and the antennas are equally spaced by half-wavelength. We model the large scale coefficients using the 3GPP Urban Microcell Model described in (3.1.2) and hence consider that the APs and UEs are 10m and 1m above the ground respectively. As for the small scale fading we consider the correlated Rayleigh fading model. We assume a coherence block with $\tau_c = 200$ transmission symbols, where $\tau_p = 10$ are symbols for the uplink pilots or the length of each pilot sequence and $\tau_d = 190$ are the symbols for the downlink data. Since we are only investigating the downlink, we assume that the symbols for the uplink data are $\tau_u = 0$. For the pilot transmission, the UEs are using a maximum transmit power of $p_{max} = 100mW$. We also adjust our system in regard to the number of UEs K to test its scalability. To perform the simulations we used MATLAB R2023b for developing the code and acquiring the required data.

Area of interest	1000 m x 1000 m
Bandwidth	20 MHz
Number of APs	L = 36
Number of UEs	K = 20, 30, 40
Number of antennas per AP	N = 4
Pathloss exponent	$\alpha = 3.76$
Per-AP maximum DL transmit power	1 W
UL transmit power	0.1 W
UL/DL noise power	-94 dBm
Pilot sequence length	10
Monte-Carlo Simulations	1000
Channel Realizations	100
Antenna Spacing	$\lambda/2$
Angular Standard Deviation	10°

Table 4.1: Simulation Setup

4.2 Simulation Scenarios and Results

We perform simulations for the following scenarios:

- We consider that the pilot assignment is done either randomly or by using the pilot assignment algorithm, labeled as “Gyros Pilot Assignment”, or “GPA”, described in (3.2.3) or by using the k-means algorithm described in (3.3).
- We consider three cases for the AP-UE association, meaning the way that the UEs are served by the APs:
  - All of the APs are serving all of the UEs, labeled as “all”.
  - We define a predetermined cluster of APs that are serving the UEs. The cluster size in that case is $\sqrt{L}$. The method is referred to as “predetermined”.

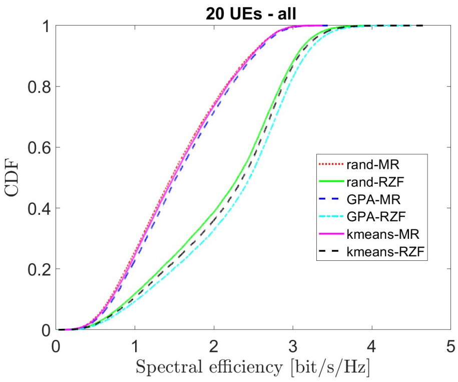
- We create clusters of size 7 with the 7 APs that have the best, i.e., highest channel gains to the UEs. We then choose to optimize the group of the APs that are serving one UE by removing the APs that contribute less than 5% in the sum of the channel gains for that particular UE. The method is labeled as "cluster".

4.2.1 Analysis of the CDF graphs

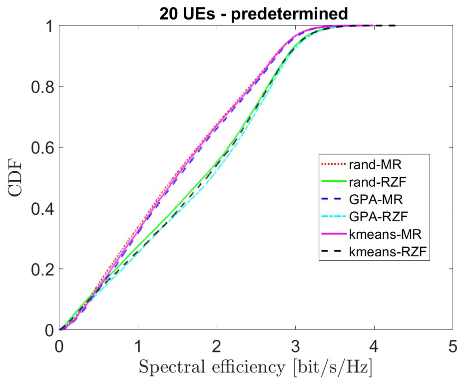
Based on the aforementioned scenarios there are in total 6 simulation scenarios to run for each case of number of UEs in our network. We first want to verify that the RZF precoding performs better than the MR precoding in all of scenarios, as suggested by the theory. After confirming it, our focus shifts to finding out in which combination scenario of pilot assignment algorithm and AP-UE association method, our system performs better.

We begin by presenting the joint results for different pilot assignment for the three UE-AP association cases. In Fig. 4.1, the downlink spectral efficiency per UE is presented for the three combination cases for 20 UEs. It is evident that in all of the three cases, the RZF precoding performs better, since the RZF curve is always more shifted to the right, i.e., towards the higher SE values, which means that for the same percentile of users the RZF precoding provides higher SE per UE than the MR precoding. The difference between the two precoding schemes appears to be more prominent in the case where all of the APs are serving all of the UEs. The reason why RZF performs better than MR in all of the cases is because it suppresses interference, hence providing improved signal quality compared to the one that MR precoding does.

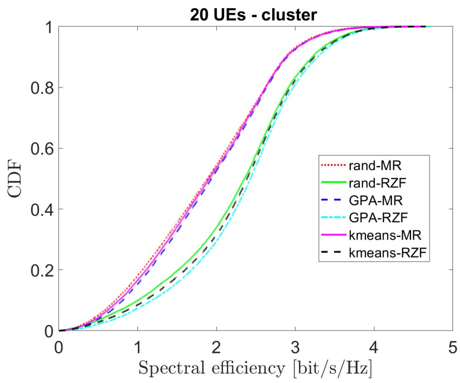
Similar conclusions can be assumed by inspecting Fig. 4.2 for 30 UEs and Fig. 4.3 for 40 UEs. One of the main reasons this happens is because the pilot sequence length does not change when the number of users is increased, and hence there is higher probability that pilot contamination will occur. As expected, it is observed that the SE provided to the UEs drops when increasing the number of UEs. In addition, it seems that the performance of the precoding techniques are not affected by the number of the UEs and that the clustering methods exhibit similar properties regardless of the number of UEs. The GPA algorithm combined with RZF precoding and the "cluster" UE-AP association seems to be the one that performs the best for all of the number of UEs and the random-"cluster" combination comes second in performance. This observation is more clear in the figures of the average spectral efficiency presented in subsection 4.2.2.



(a) “all”



(b) “predetermined”



(c) “cluster”

Figure 4.1: Spectral efficiency per UE with $K = 20$ for k-means and simple algorithm based pilot assignment.

4.2.2 Analysis of the Average SE graphs

To obtain more apparent conclusions on the performance of the pilot assignment algorithms, we study the average SE per UE in Fig. 4.4. In this figure we compare all the different combinations of pilot assignment and clustering methods. If we first compare the random one with the GPA one, the superiority of RZF precoding is even more evident. As for the performance of the pilot assignment algorithms, we can observe that the GPA cluster method provides the highest average SE per UE for both precoding schemes. However, the random pilot assignment combined with the case where all of the APs are serving all of the UEs, seems to perform better with RZF precoding than with MR. A following notable observation is that the worst performing AP-UE association for RZF is the predetermined one while for MR precoding it is the “all” case. This suggests that the predetermined method may lead to sub-optimal AP-UE

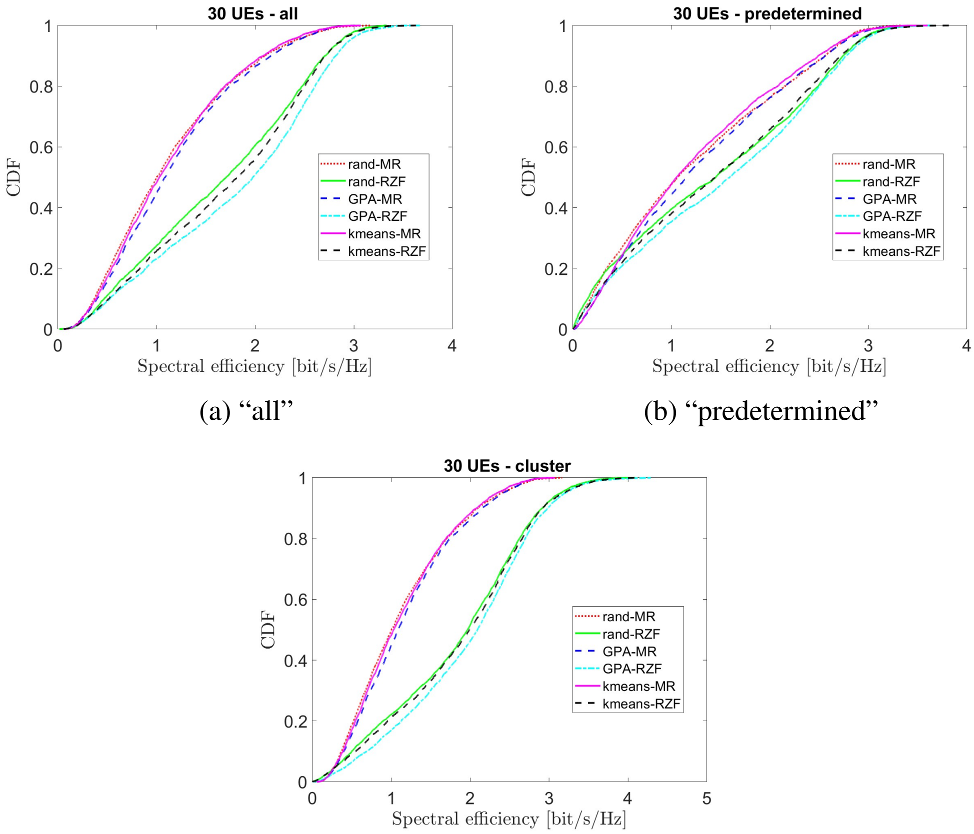
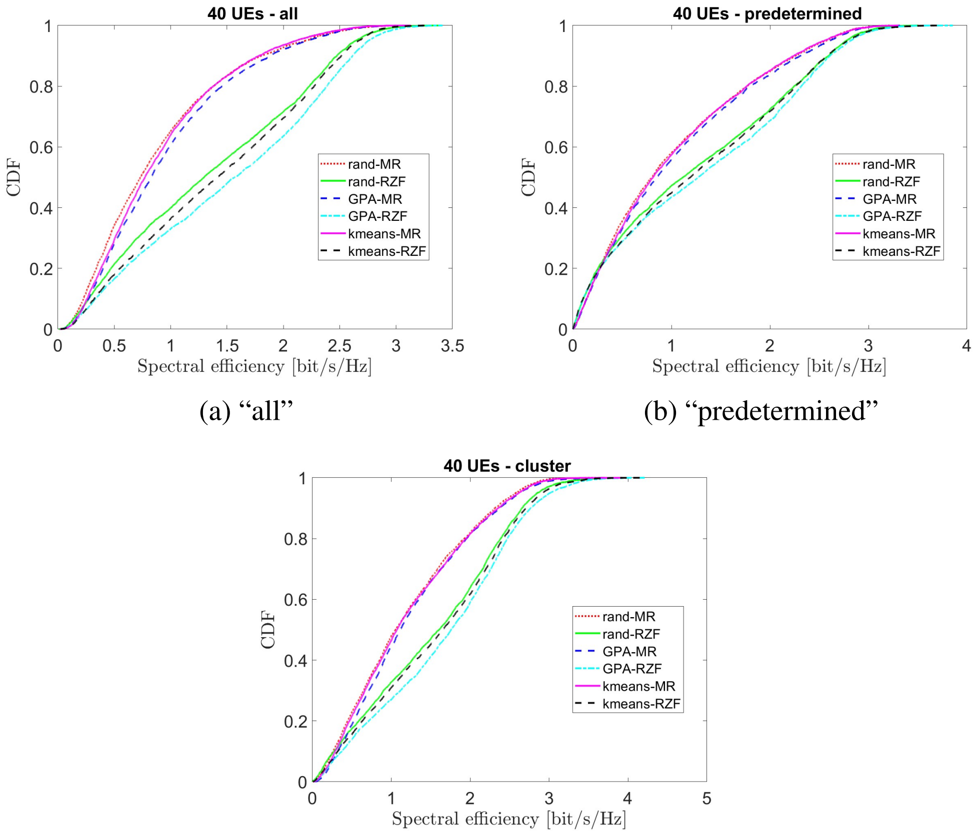


Figure 4.2: Spectral efficiency per UE with $K = 30$ for k-means and simple algorithm based pilot assignment.

associations in certain situations, meaning that the highest channel gains might not always guarantee the best performance.

Regarding the results for the k-means algorithm, as observed, the k-means algorithm performs similarly to the random pilot assignment, which means we can compare directly the k-means algorithm with the GPA one. From Fig. 4.4 is harder to compare the two, so we create a bar plot shown in Fig. 4.5 where we compare the best performing scenarios only for RZF precoding, since it accentuates the differences between the algorithms. Disregarding the random algorithm, the best performing combinations for RZF precoding are: GPA-cluster, k-means cluster, GPA-all and kmeans- all.

When comparing the algorithms for the same AP-UE assocations, with GPA there is a 3.67%, 5.85%, 7.04% increase in comparison to the k-means algorithm for 20, 30 and 40 UEs respectively and when all of the APs are serving the UEs. When the "cluster" method is used the corresponding percentages are 2.19%,



(a) “all” (b) “predetermined”

(c) “cluster”

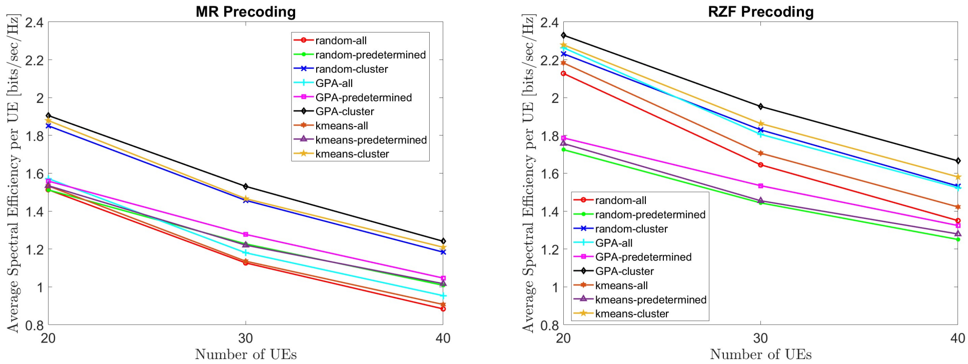
Figure 4.3: Spectral efficiency per UE with $K = 40$ for k-means and simple algorithm based pilot assignment.

4.84% and 5.67%.

For the GPA algorithm, using the “cluster” method provides a 3.1%, 7.73% and 9.86% increase in the average spectral efficiency per UE respectively. For the k-means algorithm the same percentages are 4.59%, 8.77% and 11.27%.

4.2.3 Analysis of the 90% likelihood graphs

In Fig. 4.6 we plot the 10th percentile of the CDF of SE per user vs the number of users for all combinations of pilot assignment and clustering methods. We notice that predetermined clusters perform the worst in the 90% likelihood scenario for both RZF and MR. This is the case, because the APs that are serving each UE are not optimal, as the predetermined clusters include APs that are not necessarily close to the UE that selects that particular cluster. Because the process of selection in the predetermined cluster case is done by finding the maximum of



(a) MR Precoding (b) RZF Precoding

Figure 4.4: Average Spectral efficiency per UE vs Number of Users for the two precoding schemes.

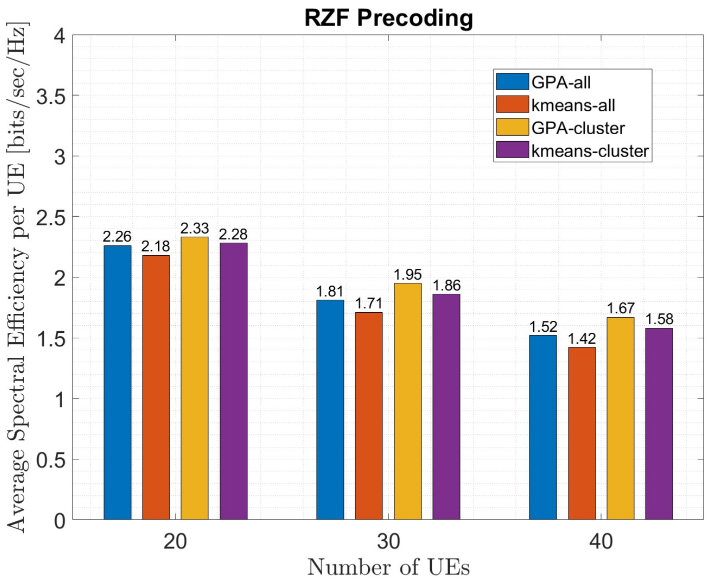
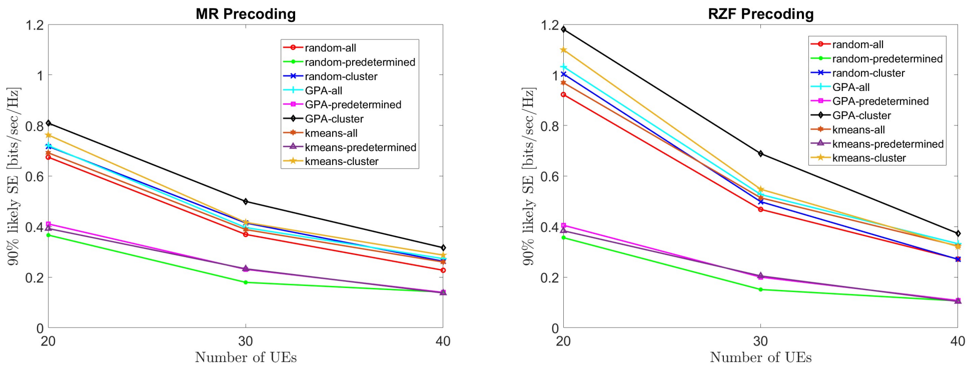


Figure 4.5: Bar Plot of Average Spectral efficiency per UE vs Number of Users for the RZF precoding scheme.

the sums, of the channel gains from a particular UE to all the APs that belong to a particular cluster, that does not guarantee that some APs that are chosen to serve that UE are actually not providing any benefit to the UE in terms of performance (or that the benefit is minimal). In reality, any signals received from the UE from the APs with poor channel gains, are actually contributing more to the interference that a UE receives and are downgrading the performance of that UE. Additionally, the performance increases when all APs serve all UEs, as the

APs that have better channel gains to a particular UE are serving them. This can be explained as following: in the case of predetermined clusters, some of the APs with the best channel gains to a particular UE, are not necessarily serving that UE, because the clusters are predetermined and the UE just decides for a whole group of APs to serve them, that are, in almost no case, the best (but they might contain some of the best APs, that are however not enough to outweigh the loss of performance from the "bad" APs and the loss of performance from losing the "best" APs). All of the above are especially true in the 90% likelihood case, because this refers to the performance of the worst performing UEs in the simulation, so the effects are more pronounced in this case.



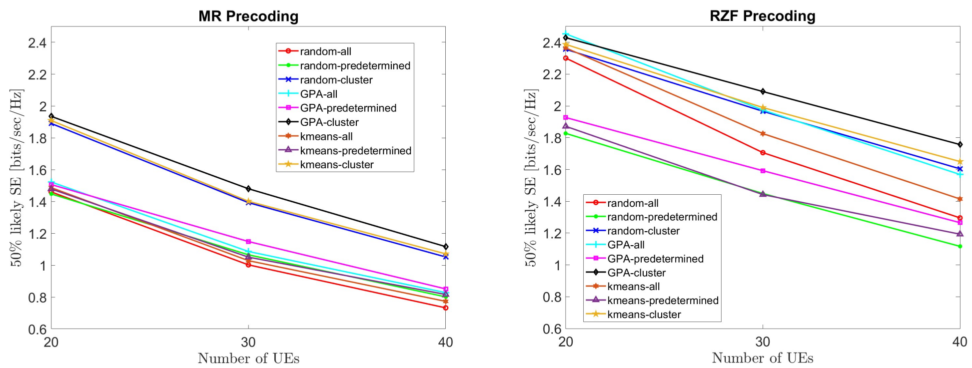
(a) MR Precoding (b) RZF Precoding

Figure 4.6: 90% likely Spectral efficiency per UE vs Number of Users for the two precoding schemes.

4.2.4 Analysis of the median graphs

By observing the median or 50%-likely SE per UE, we can understand how our system performs for the middle case users, i.e., for an average or typical user. The first observation from Fig. 4.7 is that again RZF provides higher spectral efficiency in all cases and for all number of users, meaning that its performance is independent both of the scale of the system and of the type of users. Another interesting insight from these graphs is that while for MR precoding the top three combinations are always GPA-cluster, k-means-cluster and random-cluster, for RZF that is not the case. For RZF we have GPA-all performing the best for 20 users, and GPA-cluster for 30 and 40 users, and then GPA-cluster for 20 users and k-means-cluster for 30 and 40 users at second place. K-means cluster comes third for 20 users, while for 30 users GPA-all is in the third place and random-cluster is the third best for 40 users. Despite the performance being quite close to be able to deduce influential conclusions, we have an indication that for RZF precoding the average user experience is different depending on the overall number of users as well as on the choice of clustering. For example, for lower total number of users, it seems that letting all of the APs serve the users is more efficient than selectively letting the best ones serve the users. This could derive from the interference suppression that RZF performs. For lower number of users the effect of having all of the APs serve all users is borderline better.

Looking towards the higher number of users, we can see that implementing random pilot assignment but clustering in a smarter way is borderline better than assigning pilots smarter but serving all users. One reason why this may happen is spatial diversity. Random pilot assignment could lead to a more diverse set of users being served, providing spatial diversity. In contrast, assigning pilots based on interference metrics might prioritize certain users, potentially leading to less diversity. This is not the same for lower number of users because with a smaller number of users, the spatial diversity is inherently limited. The benefits of random pilot assignment in terms of spatial diversity may not be as pronounced since there are fewer users to distribute pilots among.



(a) MR Precoding (b) RZF Precoding

Figure 4.7: 50% likely Spectral efficiency per UE vs Number of Users for the two precoding schemes.

4.2.5 Analysis of the 10% likelihood graphs

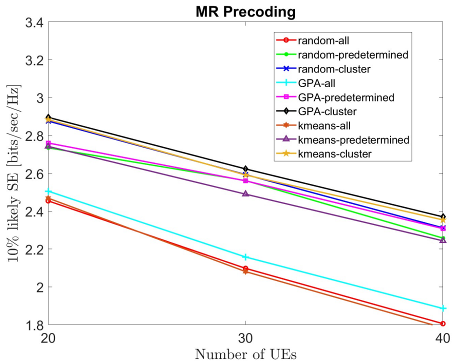
In Fig. 4.8, we plot the 90th percentile of the CDF of SE per user for the two precoding schemes. The 10% SE focuses on the upper tail of the SE distribution. By observing the 10%-likely SE per UE, we can understand how our system performs for the elite case users. In contrast to the median case, here the performance seems to follow a certain fashion regarding the top three best ones both for MR and for RZF. In this case, it also interesting to highlight that for MR the worst performing combinations are the ones where no clustering is employed. One reason for this is that clustering can facilitate more effective resource allocation, ensuring that the top users receive optimal pilot assignments and precoding weights. This targeted optimization could be more challenging when serving all users without clustering, especially in scenarios with a high number of users. This is also noticeable on the graph, since the gap in provided SE between the combinations that use "all" and the rest is larger for 40 users.

When comparing the MR case with the RZF case, it is evident that while SE decreases with the increase of the number of users, it does not decrease in a similar rate as in MR precoding. On the contrary, we could say that RZF manages to provide higher SE for the elite users regardless of the scale of the system. This is of course related to the inherent characteristic of RZF of being able to mitigate interference. In addition, RZF may exhibit a more saturated SE performance with an increased number of users, since the regularization in RZF prevents excessive interference amplification, leading to a smoother decrease in SE as user density increases. The difference in the elite user cases is more pronounced because elite users are associated with higher channel gains and hence the performance degradation stemming from interference management is more prominent.

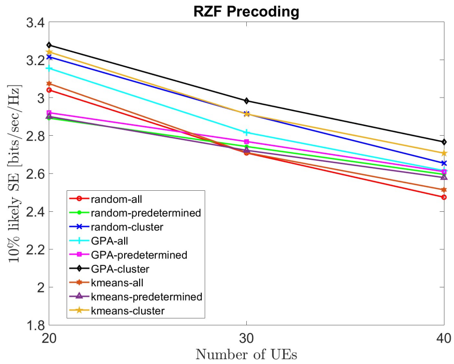
To give a more qualitative assessment, we have isolated the best performing algorithms in the majority of the cases exclusively for RZF precoding in Fig. 4.9. From these bar plots it is clear that the GPA-cluster combination provides the highest spectral efficiency in every case and for all types of users except for borderline performing worse that the GPA-all combination for 20 users in the median SE, which was mentioned earlier. In general the trend is that smart clustering of the APs that are serving each UE provides better performance than all APs serving each UE and that the GPA algorithm outperforms k-means.

When comparing the GPA-cluster with k-means cluster, we can see that the GPA algorithm greatly improves the performance of the users represented in the lower tail of the CDF curve, meaning the majority (90%) of the users, since it provides around 7%, 25% and 15% increase for 20, 30 and 40 users correspondingly, compared to k-means.

When we examine the performance of the GPA-all versus the GPA-cluster case, the GPA-cluster method provides an increase of 14%, 30% and 12% for 20, 30 and 40 users respectively when looking at the majority of the users (90%-likely SE) and -0.82%, 6% and 12% 2for for 20, 30 and 40 users correspondingly when looking at the median, or the average user cases. For the elite case users, it seems like it is providing greater enhancement for the larger number of users with a 5.6% and 6.1% increase for 30 and 40 users.

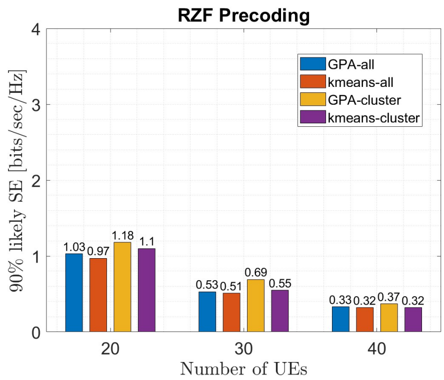


(a) MR Precoding

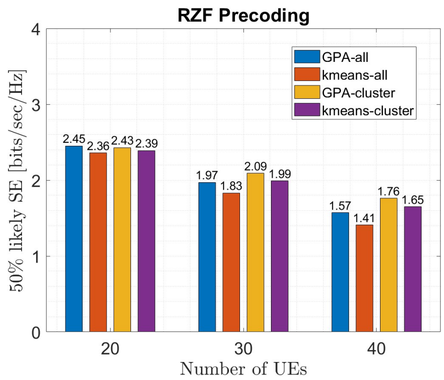


(b) RZF Precoding

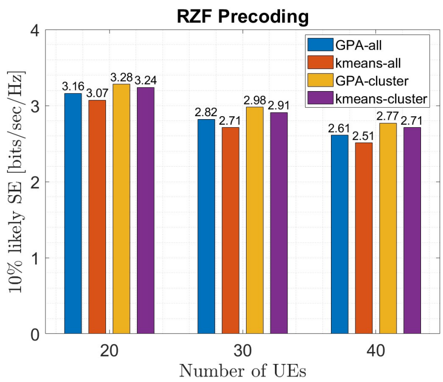
Figure 4.8: 10% likely Spectral efficiency per UE vs Number of Users for the two precoding schemes.



(a) 90%-likely



(b) 50%-likely



(c) 10%-likely

Figure 4.9: 90-50-10%-likely SE per UE for RZF precoding.

Chapter 5

Conclusions

5.1 Conclusion

In this project, we investigated the upcoming technologies introduced by cell-free massive MIMO networks in the field of wireless communications. Our work was focused on developing algorithms for efficient pilot assignment and clustering formation, along with evaluating the system's performance in regard to the spectral efficiency provided to each user. This study also dealt with the system modeling of such networks, with great focus on channel modelling and channel estimation methods, as well as precoding schemes. The investigation delved into the efficiency of MR versus RZF precoding, revealing the superiority of RZF in suppressing interference later on in the study.

At the heart of our project were pilot assignment and clustering algorithms and especially the development of the GPA algorithm paired with RZF precoding that showcased resilience and efficiency across different scenarios. Beyond precoding, we analyse how clustering influences resource distribution and user experience, concluding that the GPA algorithm with RZF precoding is superior for average spectral efficiency.

5.2 Limitations

In general, while cell-free massive MIMO networks are gaining attention from the researching community, the current work, including ours, seems to remain only in simulation level and not to align with real-life implementation and integration in modern cellular networks. For example, in our project we have considered that a ULA is used by the APs, while in reality there exist other types of antennas, such as compact planar arrays [2] that have been deployed in modern 5G networks. With that being said, it is conceivable, that our project and our results are confined by the same limitations that simulation projects are subject to.

In addition, to be more specific to our implementation, while the project addresses pilot contamination, the dynamic nature of user mobility and evolving interference patterns in practical scenarios may present challenges not fully accounted for in the simulation setup. The same principle applies for the neglected mobility of users, since our setup is investigating a static environment.

Finally, since we are working in a simulation environment, it is evident that our results are affected by the accuracy of our simulator. However, the accuracy itself partly depends on the number of simulations and hence the hardware capabilities of our computer systems, as well as time availability. With enhanced hardware and less time constraints, we could increase the number of simulations and channel realizations even more and have a more accurate channel estimation as well as more accurate results from our Monte-Carlo simulations.

5.3 Future work

The k-means based, pilot assignment algorithm implementation could be improved more, so it can achieve better performance when implemented on a larger scale. Because we set a UE threshold on a per-cluster basis that is equal to the amount of total pilot sequences that are available to us, this has created some problems as byproducts. One such example is user redistribution if at least one of the clusters exceeds its determined capacity. Our approach is currently lacking a reliable mechanism, that can efficiently reassign users when a k-means cluster is full. As a result, we cannot be sure whether we can avoid some very bad cases when the implementation scenarios are scaled up and users will need to (almost inevitably) need to be redistributed.

One method that was proposed to combat this issue, was to take the over-capacity clusters and kick out the excess UEs that are located the furthest away from the cluster's centroid position. After that, we reassign these UEs to the cluster whose centroid is the 2nd closest to them and then we recalculate and update all of the centroid positions in our simulation. This process is repeated until all clusters are not exceeding capacity. Upon first inspection, this approach seems reasonable and effective, however this method does not guarantee that it will ever reach a converged state, so it might create a never-ending loop where we are constantly performing the same steps without progression. At this point, any further investigation of this algorithm was deemed unnecessary due to: the deadline for this project being very close and because the k-means algorithm is not easy to scale by design, because it is centralized and expects the network to have knowledge of the location of the UEs (to calculate the distances from the centroids).

5.4 Reflections

In examining the implications of our work, it becomes apparent that the optimization of cell-free massive MIMO networks holds multifaceted impacts. From an economic standpoint, the increased spectral efficiency and enhanced user experience contribute to the economic sustainability of wireless communication systems. By maximizing resource utilization, our findings provide a pathway for mobile operators to deliver high data rates efficiently, potentially reducing infrastructure costs.

On a social level, the improved performance of these networks ensures equitable access to robust communication services. This is especially crucial in a world where connectivity is integral to education, healthcare, and social interactions.

From a sustainability point of view, optimized network performance translates to reduced energy consumption, aligning with the broader industry trend towards eco-friendly communication technologies.

In summary, our investigation into cell-free massive MIMO networks goes beyond technical innovations, encompassing economic sustainability, social inclusivity and environmental awareness. Through these considerations, our efforts aim to play a role in fostering a wireless communication future that is both sustainable and equitable.

Bibliography

[1] E. Björnson, J. Hoydis, and L. Sanguinetti, *Massive MIMO Networks: Spectral, Energy, and Hardware Efficiency*, 2017.

[2] T. Demir, E. Björnson, and L. Sanguinetti, *Foundations of User-Centric Cell-Free Massive MIMO*, 2021.

[3] A. Håkansson, "Portal of research methods and methodologies for research projects and degree projects," in *The 2013 World Congress in Computer Science, Computer Engineering, and Applied Computing WORLDCOMP 2013; Las Vegas, Nevada, USA, 22-25 July*. CSREA Press USA, 2013, pp. 67–73.

[4] [Online]. Available: https://ma-mimo.ellintech.se/2018/11/02/when-are-downlink-pilots-needed

[5] R. Sabbagh, C. Pan, and J. Wang, "Pilot allocation and sum-rate analysis in cell-free massive mimo systems," in *2018 IEEE International Conference on Communications (ICC)*, 2018. doi: 10.1109/ICC.2018.8422575 pp. 1–6.

[6] Z. Ren, A. Doufexi, and M. A. Beach, "Scalable pilot assignment for user-centric cell-free massive mimo networks," in *ICC 2022 - IEEE International Conference on Communications*, 2022. doi: 10.1109/ICC45855.2022.9838896 pp. 2555–2560.

[7] W. Li, X. Sun, and D. Chen, "Pilot assignment based on weighted-count for cell-free massive mimo systems," in *2021 Asia-Pacific Conference on Communications Technology and Computer Science (ACCTCS)*, 2021. doi: 10.1109/ACCTCS52002.2021.00058 pp. 258–261.

[8] C. F. Mendoza, S. Schwarz, and M. Rupp, "Cluster formation in scalable cell-free massive mimo networks," in *2020 16th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, 2020. doi: 10.1109/WiMob50308.2020.9253391 pp. 62–67.

[9] B. Zhong, X. Zhu, and E. G. Lim, "Clustering-based pilot assignment for user-centric cell-free mmwave massive mimo systems," in *2022*

IEEE 96th Vehicular Technology Conference (VTC2022-Fall), 2022. doi: 10.1109/VTC2022-Fall57202.2022.10012778 pp. 1–5.

[10] S. Chen, J. Zhang, E. Björnson, J. Zhang, and B. Ai, "Structured massive access for scalable cell-free massive mimo systems," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 4, pp. 1086–1100, 2021. doi: 10.1109/JSAC.2020.3018836

[11] L. Ahlin, J. Zander, and B. Slimane, *Principles of wireless communications*, andra upplagan ed. Lund: Studentlitteratur, 2018. ISBN 9789144126531

[12] 3GPP, "Further advancements for e-utra physical layer aspects (release 9)," 3rd Generation Partnership Project (3GPP), Technical Specification TS 36.814, 2017.