ANR POSEIDON Deliverable D3.2 - Cell-free MIMO: Channel estimation, acquisition and MU precoding and detection - preliminary version

Chaima Beldi, Didier Le Ruyet, Hmaied Shaiek, Rafik Zayani, Jean-Baptiste Doré

anr agence nationale de la recherche

Grant ANR-22-CE25-0015

Deliverable D3.2
Cell-free MIMO: Channel estimation, acquisition and MU precoding and detection - preliminary version

Delivery date	12/10/2024
Version	3.2
Editor	CNAM
Authors	Chaima Beldi (CNAM), Hmaied Shaiek (CNAM), Didier Le Ruyet (CNAM), Rafik Zayani (CEA), Jean-Baptiste Doré (CEA)
Dissemination	Public
Keywords	Cell-free massive MIMO, Hardware impairment, CSI, CSI compression

History

Version	Date	Modification	Authors
1.0	12/10/2024	First version	Chaima Beldi (CNAM), Hmaied Shaiek (CNAM), Didier Le Ruyet (CNAM), Rafik Zayani (CEA), Jean-Baptiste Doré (CEA)

Executive summary

The POSEIDON project aims to define solutions for scalable CF-mMIMO operating in the sub-7 GHz frequency bands (where the available spectral resources are scarce) to overcome various challenges (coverage, capacity, environmental sustainability. . .). In particular, scalable CF-mMIMO architectures must be able to handle i) the dramatic increase in wireless traffic demand, which is caused by the exponential growth of connected wireless devices, and ii) the emerging services/applications requiring huge data traffic (e.g., high-quality video calls, holographic communications and Internet of Things/mMTC). Moreover, as an additional target, POSEIDON will propose architectures that may contribute in the roadmap for the necessary transition to greener solutions and infrastructures. With the ICT sector's power consumption increasing exponentially through the different generations of radio mobile networks, a tenfold increase of the power consumption for the wireless access is expected over the next decade. Thus, power consumption is among the critical key performance indicators (KPIs) to be optimized in 6G networks. To this end, POSEIDON will provide solutions to satisfy the expected 6G's requirements with ever-increasingly ubiquitous and reliable wireless connectivity while at the same time steadily addressing the crucial reduction of the ecological impact of cellular infrastructures. The project will focus on lower layers, physical and access, of CF-mMIMO where the consequences of these aforementioned objectives are direct and challenging. WP3 is dedicated to the development of signal processing solutions that enhance energy efficiency, spectral efficiency, and scalability in CF-mMIMO systems. These solutions will be based on realistic propagation and hardware impairment models from WP1, ensuring that the outcomes are practical. Special attention is given to distributed and scalable aspects of these architectures. This first version of this deliverable explores several areas of CF-mMIMO system development:

- **Channel Estimation Errors and Non-linearity:** Addressing the inaccuracy of channel estimation and its effects on system efficiency. A detailed analysis of CSI imperfections, such as channel estimation errors and the influence of hardware impairments like non-linear power amplifiers, is provided. Additionally, precoding schemes for downlink data transmission are analyzed to improve performance in the presence of limited feedback.
- **State-of-the-Art CSI Compression:** Different approaches for CSI compression are explored, such as quantization based, compressed sensing, dimensionality reduction techniques, and deep learning-based methods. These techniques are crucial for enhancing the feedback efficiency and computational scalability of CF-mMIMO systems, particularly in large-scale deployments. In the next version of this deliverable, we will implement one of these approaches specifically within the cell-free massive MIMO context, providing practical insights into their real-world application.
- **Energy-Efficient Precoding in CF-mMIMO-OFDM:** A focus on local precoding techniques, specifically tailored to deal with hardware imperfections and to improve energy efficiency in OFDM-based CF-mMIMO systems. Two key precoding strategies, full-pilot zero-forcing (FZF) and regularized zero-forcing (RZF), are considered to enhance performance under hardware impairments (HWI), particularly power amplifier distortions.

Table of content

1 INTRODUCTION 5
1.1 General introduction 5
1.2 Notations 5
2 PERFORMANCE ANALYSIS OF CF-MMIMO SYSTEMS WITH IMPERFECT CHANNEL CSI 7
2.1 Introduction 7
2.2 System Model 7
2.3 Downlink data transmission 8
2.3.1 Precoding vectors for DL data transmission 8
2.4 Effect of imperfect CSI on spectral efficiency 8
2.4.1 Channel Estimation Error Model 9
2.4.2 Spectral efficiency in DL transmission considering CEE 9
2.5 Impact of non-linear power amplifier distortion on spectral efficiency 9
2.5.1 Modeling the power amplifier non linearity 9
2.5.2 Signal model in the presence of non-linear power amplifier distortion 10
2.5.3 Analytical calculation of α_m and σ^2_{dis} 11
2.5.4 Spectral efficiency expression considering the non-linear power amplifier distortion 11
2.6 Impact of limited feedback on spectral efficiency 11
2.6.1 Random Vector Quantization of the channel direction information 12
2.6.2 Optimized vector quantization of the channel direction information 12
2.6.3 Channel reconstruction at APs 13
2.7 Simulation Results 13
3 CSI COMPRESSION IN THE CF-MMIMO CONTEXT 18
3.1 State of the Art on Channel State Information Compression in Massive MIMO Context . 18
3.1.1 Quantization-Based Approaches 18
3.1.2 Compressed Sensing (CS) Techniques 18
3.1.3 Dimensionality Reduction Techniques 18
3.1.4 Deep Learning-Based Approaches 18
4 LOCAL PAPR-AWARE PRECODING FOR ENERGY-EFFICIENT CELL-FREE MASSIVE MIMO-OFDM SYSTEMS 20
4.1 Introduction 20
4.2 System, Channel and hardware impairment models 20
4.2.1 System Model 20
4.2.2 Local MMSE Channel estimation 20
4.2.3 Hardware Impairment Model 23

4.3 Downlink CF-mMIMO-OFDM under HWI. 25
4.3.1 Downlink CF-mMIMO-OFDM Data Transmission 25
4.3.2 Downlink Spectral Efficiency 26
4.4 Simulation results 27
4.4.1 Simulation Scenario 27
4.4.2 Performance Evaluation. 28
5 CONCLUSIONS 31
REFERENCES 32

1 Introduction

1.1 General introduction

The demand for reliable and high-speed wireless connectivity has surged in both personal and professional environments, placing significant pressure on existing communication networks. This growing need is largely fueled by data-intensive applications such as high-definition video streaming, virtual and augmented reality, and the proliferation of the Internet of Things (IoT). These applications require communication systems that not only offer seamless connectivity but also demonstrate high efficiency and reliability. Addressing these evolving requirements is crucial for future wireless networks, particularly in the context of the 5G and beyond ecosystems.

In response to this need, the POSEIDON project represents a pivotal initiative aimed at advancing cell-free massive multiple-input multiple-output (CF-mMIMO) technology in the sub-7 GHz frequency bands. CF-mMIMO systems offer a revolutionary approach to wireless communication by utilizing a network of distributed antennas that work cooperatively to enhance system performance. Unlike conventional cellular networks that rely on fixed base stations and are limited by the boundaries of individual cells, CF-mMIMO systems eliminate cell boundaries, enabling continuous and widespread coverage. This architecture minimizes issues such as inter-cell interference and variations in signal quality, resulting in more efficient and robust communication.

However, the implementation of CF-mMIMO systems presents significant challenges. Key among these are the issues related to imperfect Channel State Information (CSI), hardware impairments, and the limitations of feedback systems. These factors can introduce inefficiencies and degrade overall system performance if not adequately addressed. The complexities of managing distributed antennas, ensuring accurate CSI, and overcoming hardware limitations require advanced signal processing solutions.

In this context, WP3 of the POSEIDON project focuses on developing signal processing techniques to enable energy-efficient and spectrally efficient CF-mMIMO systems. This first version of this deliverable marks an exploration of such solutions.

In the first chapter, we examine how inaccuracies in channel estimation can impact system efficiency and provide a comprehensive analysis of the effects of these imperfections. Additionally, we investigate the challenges posed by non-linear power amplifiers and constrained feedback mechanisms, both of which can negatively influence the performance of CF-mMIMO systems.

In the second chapter, we provide a thorough review of the state-of-the-art approaches proposed in the literature, focusing on those addressing the problem of CSI compression in the context of massive MIMO. While many of these approaches have not yet been implemented specifically in cell-free systems, they offer valuable insights that can inspire solutions for CF-mMIMO. By exploring these methods, we aim to draw inspiration from the MIMO context and adapt on of these strategies to address the unique challenges posed by cell-free architectures.

Finally, we present an energy-efficient downlink transmission scheme for CF-mMIMO-OFDM based system. Two precoding techniques are considered : local full-pilot zero-forcing and local regularized zero-forcing. In this work we will give a description of the system model and provide a first analysis as well as simulations results.

1.2 Notations

The notations used in this deliverable are listed as follows:

- Lowercase boldface letters (e.g. $\mathbf{x}$) stand for column vectors,
- Bold lowercase letters with a superscript $(.)^t$ (e.g. $\mathbf{x}^t$) denotes row vectors,
- Bold uppercase letters (e.g. $\mathbf{X}$) denotes matrices,
- The $\check{}$ marks time-domain variables over the paper,
- We denote by matrix transpose, matrix conjugate transpose, matrix pseudo-inverse and trace of a matrix by $\mathbf{X}^T$,$\mathbf{X}^H$, $\mathbf{X}^\dagger$ and $tr(\mathbf{X})$, respectively,
- For a $M \times N$-dimensional matrix $\mathbf{X} = \{x_{mn}\}$,

- We use $\mathbf{x}_n$ to designate the n-th column, and $\mathbf{x}_m^t$ to designate the m-th row,
- The $N \times N$ identity matrix and the $M \times N$ all-zeros matrix are denoted by $\mathbf{I}_N$ and $\mathbf{0}_{M \times N}$, respectively,
- We use $\|\mathbf{x}\|_2$ and $\|\mathbf{x}\|_\infty$ to denote l_2-norm and l_∞-norm of vector $\mathbf{x}$, respectively,
- $\mathbf{X} = diag\{x_1, ..., x_K\}$ denotes a diagonal matrix with elements $\{x_i\}$,
- the cardinality and complement of set Ξ is $|\Xi|$ and Ξ^c, respectively,
- $\mathrm{E}[.]$ stands for the expectation operator and j denotes $\sqrt{-1}$.

2 Performance Analysis of CF-mMIMO systems with imperfect channel CSI

2.1 Introduction

The CF-mMIMO system has the potential to significantly enhance wireless connectivity by using numerous distributed antennas and advanced signal processing to improve efficiency and performance. However, real-world implementations face challenges such as imperfect Channel State Information (CSI), non-linear hardware issues, and limited feedback mechanisms, all of which can affect performance. In this initial version of the deliverable, we closely examine these challenges. We start by explaining the system model to provide a basic understanding of CF-mMIMO. Next, we discuss the precoding schemes used for downlink transmissions and describe two different cooperation strategies among the Access Points (APs). We then look at how errors in channel estimation impact efficiency and provide a detailed analysis of these imperfections. Finally, we explore the effects of non-linear power amplifier issues and limited feedback on system performance.

2.2 System Model

We consider a CF-mMIMO network consisting of P single antenna UEs and M APs, each equipped with K antennas as shown in Fig.1. The total number of antennas in the coverage area is denoted as N, where $N = MK$.
The CPU controls all the APs over fronthaul links, providing the signaling needed for synchronization and coherent joint reception and transmission to the UEs. The channel from the m-th AP to the p-th UE is denoted by $\mathbf{h}_{p,m} \in \mathbb{C}^{K\times 1}$. We use the correlated channel Rayleigh fading model to describe the spatial correlations of each channel with $\mathbf{h}_{p,m} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}_K, \mathbf{R}_{p,m})$ where $\mathbf{R}_{p,m} \in \mathbb{C}^{K\times K}$ is the spatial correlation matrix. The complex Gaussian distribution represents the small-scale fading of objects within the propagation environment, whereas the positive semi-definite correlation matrix $\mathbf{R}_{p,m}$ accounts for the large-scale effects.

Since the geographical coverage area of the CF-mMIMO network can be large, it must be designed to be in a way that the computational capability and fronthaul capacity of existing APs remain sufficient as more UEs are being introduced to the network. We consider so an implementation of CF-mMIMO that is inspired by the guidelines for the Dynamic Cooperation Clustering (DCC) framework [2].

With this implementation, the p-th UE is exclusively served by the APs whose indices belong to the ordered set $\mathcal{M}_p = \{m_{p,1}, m_{p,2} \ldots, m_{p,Q_p}\}$ where $Q_p = |\mathcal{M}_p|$.

For the DL, the received signal at the p-th UE can be written as:

$$r_p = \sum_{m=1}^{M} \mathbf{h}_{p,m}^H \mathbf{x}_m + b_p, \tag{1}$$

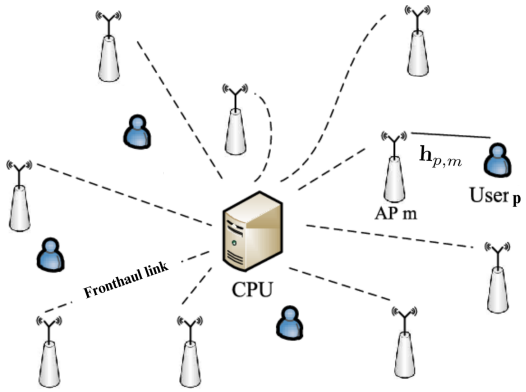

Figure 1: CF-mMIMO network with numerous distributed APs linked to CPU [12].

where $b_p \sim \mathcal{N}_{\mathbb{C}}(0, \sigma_{dl}^2)$ is the receiver noise with noise variance σ_{dl}^2 and $\mathbf{x}_m \in \mathbb{C}^{K\times 1}$ denotes the signal sent by the m-th AP which consists of a precoded superposition of the signals designed for the different UEs belonging to $\mathscr{D}_m$, where $\mathscr{D}_m$ denotes the set of UEs served by the m-th AP. This signal can be expressed as:

$$\mathbf{x}_m = \sum_{i\in\mathscr{D}_m} \mathbf{w}_{i,m}\chi_i, \tag{2}$$

where χ_i is the complex transmit symbol for the i-th UE, and $\mathbf{w}_{im} \in \mathbb{C}^{K\times 1}$ is the precoding vector that the m-th AP assigns to i-th UE. Finally, by substituting the transmitted signal in (2) into (1), the received DL signal at the p-th UE becomes:

$$r_p = \sum_{m=1}^{M} \mathbf{h}_{p,m}^H \left(\sum_{i\in\mathscr{D}_m} \mathbf{w}_{i,m}\chi_i \right) + b_p. \tag{3}$$

2.3 Downlink data transmission

In this section, we introduce two precoding strategies designed for both centralized and distributed scenarios.

2.3.1 Precoding vectors for DL data transmission

We need first to compute for the p-th UE, the designated precoding vector $\mathbf{w}_p = \frac{\bar{\mathbf{w}}_p}{\sqrt{\mathbb{E}\{\|\bar{\mathbf{w}}_p\|^2\}}}$ [2], where $\bar{\mathbf{w}}_p \in \mathbb{C}^{N_p}$ is the precoding vector that describes the spatial directivity of the DL transmission.

Centralized precoding In this scenario, the CPU utilizes the combined channel estimates from the APs serving the p-th UE whose indices are part of the set $\mathscr{M}_p$. These estimates are denoted as $\hat{\mathbf{h}}_p = [\hat{\mathbf{h}}_{p,m_{p,1}}^T \dots \hat{\mathbf{h}}_{p,m_{p,Q_p}}^T]^T \in \mathbb{C}^{N_p\times 1}$, where $N_p = Q_p \times K$ represents the total number of antennas dedicated to serving the p-th UE. The Partially-Minimum Mean Square Error (P-MMSE) precoding vector can be calculated using the following expression:

$$\begin{aligned} \bar{\mathbf{w}}_p^{P-MMSE} &\triangleq [\bar{\mathbf{w}}_{p,m_{p,1}}^T \dots \bar{\mathbf{w}}_{p,m_{p,Q_p}}^T]^T \\ &= p_p^{dl} \left(\sum_{i\in\mathscr{P}_p} p_i^{dl} \hat{\mathbf{h}}_i \hat{\mathbf{h}}_i^H + \sigma_{dl}^2 \mathbf{I}_{N_p} \right)^{-1} \hat{\mathbf{h}}_p, \end{aligned} \tag{4}$$

where $\mathscr{P}_p$ represents the set of UEs that are partially served by the same APs as the p-th UE and p_p^{dl} is the total transmit power allocated to the p-th UE.

The vector $\bar{\mathbf{w}}_p^{P-MMSE}$ is divided into Q_p vectors that are subsequently used in the computation of the received signal given by equation (3).

Distributed precoding Under this processing, the selected AP m computes locally the precoding vectors associated to the user $p \in \mathscr{D}_m$ on the basis of their local channel estimates $\{\, \hat{\mathbf{h}}_{i,m}, i \in \mathscr{D}_m \,\}$.

We propose in this work to implement a Local Precoding - Minimum Mean Square (LP-MMSE) scheme. This precoding vector associated to the p-th user and m-th AP can be expressed in the following manner:

$$\bar{\mathbf{w}}_{p,m}^{LP-MMSE} = p_p^{dl} \left(\sum_{i\in\mathscr{D}_m} p_i^{dl} \hat{\mathbf{h}}_{i,m} \hat{\mathbf{h}}_{i,m}^H + \sigma_{dl}^2 \mathbf{I}_K \right)^{-1} \hat{\mathbf{h}}_{p,m}. \tag{5}$$

It's important to highlight that for each UE p, we calculate only $\bar{\mathbf{w}}_{p,m}^{LP-MMSE}$ at the APs $m \in \mathscr{M}_p$.

2.4 Effect of imperfect CSI on spectral efficiency

In this section, we analyze the channel estimation erro (CEE) impact on the spectral efficiency (SE) of our DL transmission scenario.

2.4.1 Channel Estimation Error Model

In practice, channel estimation is imperfect and cannot perfectly capture the actual channel conditions [5, 13]. The estimated channel $\hat{\mathbf{h}}_{p,m}$ between UE p and AP m, for $p = 1,\ldots,P$ and $m = 1,\ldots,M$, differs from the real channel $\mathbf{h}_{p,m}$ due to estimation errors.

To account for these imperfections, we model the reconstructed channel at each AP as:

$$\hat{\mathbf{h}}_{p,m} = \mathbf{h}_{p,m} + \varepsilon, \tag{6}$$

where $\varepsilon \sim \mathcal{N}_{\mathbb{C}}(0, \sigma_e^2 \mathbf{I}_K)$ is the channel estimation error, assumed to follow a complex Gaussian distribution with zero mean and variance σ_e^2 [13]. Here, σ_e represents the standard deviation of the channel estimation error.

Next, the precoding vectors are computed based on the reconstructed channel $\hat{\mathbf{h}}_{p,m}$. Since the channel estimates are imperfect, the resulting precoding vectors differ from those calculated using perfect channel information.

As a result, both the central and distributed precoding schemes must adapt to these imperfections, as the computed precoding vectors directly depend on the instantaneous channel estimates $\hat{\mathbf{h}}_{p,m}$ in each coherence block.

2.4.2 Spectral efficiency in DL transmission considering CEE

To evaluate the effects of CEE, we derive the SE expression. The DL SE depends on the normalized precoding vectors of all UEs [2], which in fact are being computed using imperfect channel estimates $\hat{\mathbf{h}}_{p,m} = \mathbf{h}_{p,m} + \varepsilon$, and as such the SE is given by:

$$SE_p^{(CEE)} = \mathbb{E}\left\{\log_2(1 + SINR_p^{(CEE)})\right\}, \tag{7}$$

where $\text{SINR}_p^{(CEE)}$ is given by

$$SINR_p^{(CEE)} = \frac{p_p^{dl} |\mathbf{h}_p^H \mathbf{w}_p|^2}{\sum_{i \in \mathcal{D}_m, i \neq p} p_i^{dl} |\mathbf{h}_p^H \mathbf{w}_i|^2 + \sigma_{dl}^2}. \tag{8}$$

2.5 Impact of non-linear power amplifier distortion on spectral efficiency

In this section, we investigate how the distortion introduced by non-linear power amplifiers PA affects the SE in the DL transmission scenario.

2.5.1 Modeling the power amplifier non linearity

Non-linear PAs introduce distortions to the signals they amplify. This distortion can be especially problematic when operating the PAs close to saturation level, which is recommended to improve the energy efficiency of the overall system [30]. Therefore, our aim is to model this distortion and investigate its impact on the system performance. To begin let's write the transmitted signal from the m^{th} AP,

$$\mathbf{u}_m = f(\mathbf{x}_m), \tag{9}$$

where $\mathbf{x}_m$ is the precoded signal and $f(.)$ represents the non-linear characteristic function the PA used at the m^{th} AP's RF chains. A PA is commonly modelled by its input/output characteristics. These characteristics are expressed as Amplitude to Amplitude (AM/AM) and Amplitude to Phase (AM/PM) conversions. They describe, respectively, the relationship between the magnitude and phase of the output signal as function of the input signal amplitude. The signal at the output of the non-linear circuit is given by:

$$\mathbf{u}_m = F_a\left(|\mathbf{x}_m|\right) \exp\left(j\left(\arg\left(\mathbf{x}_m\right) + F_p\left(|\,\mathbf{x}_m\,|\right)\right)\right), \tag{10}$$

where $\arg(\mathbf{x}_m)$ is the phase of the complex signal $\mathbf{x}_m$, and $F_a(.)$, $F_p(.)$ represent respectively the AM/AM and AM/PM conversion characteristics of the PA. It's worth mentioning that the previous equation assumes identical conversion characteristics over all the PAs used in the m^{th} AP radio frequency (RF) chains. Various mathematical models

are utilized in the literature to describe the conversion characteristics of the PAs [34]. In this work, we considered a polynomial model from [33]. This polynomial model, exhibiting both AM/AM and AM/PM distortion, has been established by measurements on a real $4GHz$ PA with $1W$ maximal output power.

In practice, to avoid or reduce the effects of non-linearity, the PA is operated at a certain input power level, characterized by its Input Back-Off (IBO) from its $1dB$ compression point. In this work, we will adopt the following definition of IBO:

$$\mathrm{IBO} = 10\log_{10}\left(\frac{P_{1\ \mathrm{dB}}}{P_{\mathrm{moy}}}\right), \tag{11}$$

where P_{1dB} represents the input power at the $1dB$ compression point, and $P_{\mathrm{moy}} = \mathbb{E}\{|\mathbf{x}_m|^2\}$ is the average power of the input signal.

2.5.2 Signal model in the presence of non-linear power amplifier distortion

In this subsection, we model the non-linearity of the power amplifier using the Bussgang theorem [4]. According to this theorem, if the input signal $\mathbf{i}_m$ of a memory-less non-linear function follows a Gaussian distribution, the resulting output of the amplification device can be expressed as a scaled version of the input signal along with an additional distortion component. Crucially, this distortion component is statistically independent from the input signal. This condition, on the input signal, is satisfied for multi-carrier-based transmissions. Therefore, at the output of the m-th AP RF chains, we have:

$$\mathbf{u}_m = \alpha_m \mathbf{i}_m + \eta_{\mathrm{dis}}^{\mathrm{AP}(m)}, \tag{12}$$

where:

- $\alpha_m = diag(\alpha_m(1), \ldots, \alpha_m(K)) \in \mathbb{C}^{K\times K}$ is the collective vector of the complex gain of the K antennas of the m-th AP.
- $\eta_{\mathrm{dis}}^{\mathrm{AP}(m)} \sim \mathcal{N}_{\mathbb{C}}(0_K, \sigma_{dis}^2 \mathbf{I}_K)$ is a centered noise, uncorrelated with $\mathbf{i}_m$, with variance equal to σ_{dis}^2.

We remind that for the sake of simplicity, and without loss of generality, we have assumed that the RF chains of any AP exhibit identical PAs conversion characteristics. Therefore the complex gain $\alpha_m(k)$ and the corresponding noise $\eta_{\mathrm{dis}}^{\mathrm{AP}(m)}(k)$ are identical ove the K RF chains of the m^{th} AP. To achieve a specified IBO value in dB [34], the input PA signal, $\mathbf{x}_m$, is weighted by a gain δ and becomes:

$$\mathbf{i}_m = \delta \mathbf{x}_m = \delta \sum_{i\in\mathscr{D}_m} \mathbf{w}_{i,m}\chi_i. \tag{13}$$

where $\delta = \sqrt{\frac{P_{1\ \mathrm{dB}}}{10^{\frac{\mathrm{IBO}}{10}} P_{\mathrm{moy}}}}$. We remind that for a multi-carrier-base transmission, we can approximate the distribution of $\mathbf{x}_m$ as Gaussian random process. Consequently, we can apply the Bussgang theorem [34] to the signals at the PA outputs and write:

$$\mathbf{u}_m = \delta\alpha_m \sum_{i\in\mathscr{D}_m} \mathbf{w}_{i,m}\chi_i + \eta_{\mathrm{dis}}^{\mathrm{AP}(m)}. \tag{14}$$

So, we can express the received signal at a UE p with both centralized and distributed operation modes as follows:

$$y_p = \sum_{m=1}^{M} \mathbf{h}_{p,m}^H \eta_{\mathrm{dis}}^{\mathrm{AP}(m)} \tag{15}$$

$$+ \delta \sum_{m=1}^{M} \mathbf{h}_{p,m}^H \left(\alpha_m \sum_{i\in\mathscr{D}_m, i\neq p} \mathbf{w}_{i,m}\chi_i\right) + b_p. \tag{16}$$

2.5.3 Analytical calculation of α_m and σ^2_{dis}

Analytically, the complex gain of the k-th antenna of the m-th AP $\alpha_m(k)$ can be computed by the following equation [3]:

$$\alpha_m(k) = \frac{1}{2}\mathbb{E}\left[\frac{\partial S(r_k)}{\partial r_k} + \frac{S(r_k)}{r_k}\right], \tag{17}$$

where $r_k = |\mathbf{u}_m(k)|$, for $k = 1 \dots K$ is the magnitude of the signal to be amplified, $S(r_k) = F_a(r_k)\exp(jF_p(r_k))$ represents the complex envelope of the amplified signal. The variance $\sigma^2_{dis}(k)$ of the non-linear distortion can be also computed analytically by [3]:

$$\sigma^2_{dis}(k) = \mathbb{E}\left\{|S(r_k)|^2\right\} - |\alpha_m(k)|^2\mathbb{E}\{r_k^2\}. \tag{18}$$

The analytical computation of the non-linear distortion (NLD) parameters $\alpha_m(k)$ and $\sigma^2_{dis}(k)$ depends on the complexity of the expression of $S(.)$. A generalized approach for computing these parameters for any measured of modelled PA have been proposed in [3]. This approach, based on a polynomial approximation of $S(.)$, have been applied to compute the NLD parameters corresponding to the 3GPP PA model considered in our work.

2.5.4 Spectral efficiency expression considering the non-linear power amplifier distortion

Assuming perfect channel estimation, to gain insight into the effect of the non-linear PA distortions on the performance of the CF-mMIMO systems, we derive the expression of DL SE, given by :

$$\mathrm{SE}_p^{(NLPA)} = \mathbb{E}\left\{\log_2(1 + SINR_p^{(NLPA)})\right\}. \tag{19}$$

where $\mathrm{SINR}_p^{(NLPA)}$ is given by

$$SINR_p^{(NLPA)} = \frac{p_p^{dl}|\delta\mathbf{h}_p^H\alpha_p\mathbf{w}_p|^2}{\sum_{i\in\mathscr{D}_m, i\neq p} p_i^{dl}|\delta\mathbf{h}_p^H\alpha_p\mathbf{w}_i|^2 + \mathbb{E}\left\{\left|\mathbf{h}_p^H\eta_{dist}\right|^2\right\} + \sigma^2_{dl}}. \tag{20}$$

with $\eta_{\text{dist}} = \left[\eta_{\text{dist}}^{AP(m_{p,1})^T} \dots \eta_{\text{dist}}^{AP(m_{p,Q_p})^T}\right]^T \in \mathbb{C}^{N_p\times 1}$ is the collective vector of distortions originating from the amplifiers of all the APs served the UE p and $\alpha_p = diag(\alpha_{m_{p,1}}, \dots, \alpha_{m_{p,Q_p}}) \in \mathbb{C}^{N_p\times N_p}$ is the collective vector of the complex gain of all the APs serving the UE p.

2.6 Impact of limited feedback on spectral efficiency

When channel reciprocity cannot be fully utilized, whether in TDD or FDD systems, CSI must be estimated at the UE level and then transmitted to the APs via feedback links. CSI can be represented as the product of Channel Quality Information (CQI) and Channel Direction Information (CDI). In [38], the authors evaluated the performance of a non-scalable CF-mMIMO system with limited feedback. They designed codebooks for both CDI and CQI and assessed the downlink performance of the system using Maximum Ratio Transmission (MRT) precoding. In this work, we assume that CQI is perfectly known at the APs, while only the CDI is quantized using two approaches: Random Vector Quantization (RVQ) [29] and optimized Vector Quantization (VQ). The RVQ method generates vectors independently from a uniform distribution on the complex unit sphere, whereas the optimized VQ codebooks are constructed using the Lloyd algorithm [23, 25]. Unlike [38], we will examine a scalable CF-mMIMO system and evaluate the spectral efficiency (SE) of downlink transmissions with both centralized and distributed precoding strategies under limited feedback conditions. We will carefully consider how the number of feedback bits affects system performance.

2.6.1 Random Vector Quantization of the channel direction information

RVQ is a simple approach to evaluate the effects of quantization of the channel direction information (CDI), $\mathbf{g}_{pm}$. In RVQ, the codebook is composed of $Q = 2^B$ quantization vectors that are independently chosen from the isotropic distribution on the K dimensional complex unit sphere [38]. Here, B is the number of feedback bits.

We decompose the channel vector $\mathbf{h}_{pm}$ into two parts namely the channel quality information CQI $||\mathbf{h}_{pm}||$ and the normalized channel vector or Channel Direction Vector (CDI) $\mathbf{g}$ as follows:

$$\mathbf{g}_{pm} = \frac{\mathbf{h}_{pm}}{||\mathbf{h}_{pm}||} \tag{21}$$

The quantized CDI $\hat{\mathbf{g}}_{pm}$ is determined as the one minimizing the chordal distance between the quantization vectors of the codebook and the CDI $\mathbf{g}_{pm}$.

Let us define the square chordal distance Z, between $\mathbf{g}_{pm}$ and $\hat{\mathbf{g}}_{pm}$, as follows:

$$\begin{aligned} Z &= d^2(\mathbf{g}_{pm}, \hat{\mathbf{g}}_{pm}) \\ &= 1 - |\mathbf{g}_{pm}^H \hat{\mathbf{g}}_{pm}|^2 \end{aligned} \tag{22}$$

Z is a random variable within the interval $[0,1]$. Let us denote $p(z)$ and $F(z)$ as the probability density function and the cumulative distribution function of Z. In [25], the authors have introduced an upper bound on $F(z)$ for quantized feedback scheme assuming that the regions associated to each codeword do not overlap:

$$F(z) \leq \tilde{F}(z) = \begin{cases} Qz^{K-1} & \text{if } 0 \leq z < Q^{-\frac{1}{K-1}} \\ 1 & \text{if } z \geq Q^{-\frac{1}{K-1}} \end{cases} \tag{23}$$

The associated power density function (pdf) $\tilde{p}(z)$ is given by:

$$\tilde{p}(z) = \begin{cases} Q(K-1)z^{K-2} & \text{if } \quad 0 \leq z < Q^{-\frac{1}{K-1}} \\ 0 & \text{if } \quad z \geq Q^{-\frac{1}{K-1}} \end{cases} \tag{24}$$

This pdf can be obtained from a Beta distribution $Beta(x, \alpha, \beta)$ with $\alpha = K-1, \beta = 1$ and performing a change of support. A simple approach to generate $\hat{\mathbf{g}}_{pm}$ from $\mathbf{g}_{pm}$ for a given realization Z is to first generate a vector $\mathbf{s}$ such as $d^2(\mathbf{s}, \mathbf{o}) = Z$ where $\mathbf{o} = [1\, 0 \ldots 0]^T$. The vector $\mathbf{s}$ can be obtained as follows [11]:

$$\mathbf{s} = \begin{bmatrix} \sqrt{1-Z} \\ \tilde{\mathbf{s}} \end{bmatrix} \tag{25}$$

where $\tilde{\mathbf{s}}$ is a $(K-1) \times 1$ random unit norm vector.

Then we can obtain the vector $\hat{\mathbf{g}}_{pm}$ by multiplying $\mathbf{s}$ by the rotation matrix $\mathbf{U_{rot}}$ as follows:

$$\hat{\mathbf{g}}_{pm} = \mathbf{U_{rot}} \mathbf{s} \tag{26}$$

where $\mathbf{U_{rot}}$ is the complex Householder matrix given by [6]:

$$\mathbf{U_{rot}} = \mathbf{I_K} - \frac{1}{\mathbf{u^H o}} \mathbf{u u^H} \tag{27}$$

with $\mathbf{u} = \mathbf{o} - \mathbf{g}_{pm}$

2.6.2 Optimized vector quantization of the channel direction information

In the quantized feedback scheme, the vector $\mathbf{f}_{opt}$ is taken from a set or codebook of the precoding vectors $\mathscr{F} = \{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_Q\}$ where $Q = 2^B$ and B is the number of feedback bits. The challenge is to design the codebook $\mathscr{F}$ in order to maximize a performance criterion such as the bit error rate or the system capacity. To measure the

average distortion due to quantization, we define the distortion function. For the Multiple Input Single Output (MISO) transmission, we have:

$$G(\mathscr{F}) = \mathbb{E}_{\mathbf{h}}\left[1 - |\mathbf{g}_{pm}^{H}\mathbf{f}_{opt}|^2\right] \tag{28}$$

The codebook $\mathscr{F}$ can be seen as a collection of lines in the Euclidean space $\mathbb{C}^K$. $\mathbf{f}_i$ can be seen as the coordinates of a point situated at the surface of an hypersphere with a unit radius centered at the origin. Mathematically, $\mathbf{f}_i$ is a point in the Grassman manifold or Grassmanian space $G(K,1)$. For i.i.d. channels, the construction of the set $\mathscr{F}$ should maximize the angular separation between the two closest lines [25] [23]. This problem is called the Grassmannian line packing problem. The codebook design is formulated as follows:

$$\min_{\mathscr{F}} G(\mathscr{F}) \quad \text{subject to} \quad ||\mathbf{f}_i||^2 = 1 \quad \forall i \tag{29}$$

This problem can be seen as a Vector Quantization (VQ) problem [26] with source input $\mathbf{g} \sim \text{uniform}(\mathbb{O}_K)$, codebook $\mathscr{F}$ and distortion metric $d(\mathbf{g},\mathbf{f}_i)$. This VQ problem is also known as a spherical vector quantization problem and can be efficiently solved using the Lloyd algorithm [26].

2.6.3 Channel reconstruction at APs

We assume that the APs have perfect CQI. On the other hand, the APs and UEs are using the same CDI codebook. After perfect channel estimation each user sends the index of the codebook to the AP that is able to retrieve the corresponding channel from the codebook based on this index.

$$\hat{\mathbf{g}}_{pm} = \arg\min_{\mathbf{f}_i \in \mathscr{F}} d^2(\mathbf{g}_{pm}, \mathbf{f}_i) \tag{30}$$

The reconstructed CSI obtained through the limited rate feedback is consequently given as follows:

$$\hat{\mathbf{h}}_{pm} = ||\mathbf{h}_{pm}||\hat{\mathbf{g}}_{pm} \tag{31}$$

2.7 Simulation Results

First, we assess the performance of the proposed precoding techniques and examine the impact of CEE and the distortions caused by the non-linearity of PA on the SE under various scenarios. Drawing inspiration from [2], we construct a simulation environment where 100 access points (APs), each equipped with 4 antennas, and 100 single-antenna user equipments (UEs) are uniformly distributed across a 2×2 square kilometer area. By employing the wrap-around technique, we approximate a large network characterized by a density of 100 antennas per square kilometer and 25 UEs per square kilometer. Other simulation parameters include a bandwidth of 20 MHz, $\sigma_{dl}^2 = -94$ dBm and $p_p^{dl} = 1$ W for downlink transmission power. The curves presented are the result of independently sorted 1000 channel realizations conducted independently for each configuration.

To evaluate the CEE, we vary the variance of the CEE $\sigma_e^2 \in \{0, 10^{-2}, 10^{-1}\}$. Fig. 2 depicts the achievable downlink spectral efficiency (DL SE) of the CF-mMIMO system with imperfect channel estimation, where P-MMSE and LP-MMSE are adopted for DL transmission, respectively.

A primary observation is that the LP-MMSE precoding provides smaller SE compared to the centralized scheme, P-MMSE. Further observations will demonstrate that the former is more robust against estimation error. It is clear that the presence of CEE degrades the performance of the system with both schemes. A small performance degradation in the SE is noted when σ_e^2 is very small compared to performance with perfect channel estimation (i.e., $\sigma_e^2 = 0$). However, the performance gap is larger for P-MMSE and shrinks for LP-MMSE. When increasing the severity of the errors from $\sigma_e^2 = 10^{-2}$ to $\sigma_e^2 = 10^{-1}$, the performance is severely degraded by 2.748 bit/s/Hz for P-MMSE and 1.22 bit/s/Hz for LP-MMSE for only 50% of the UEs.

Furthermore, the degradation of the average DL SE is depicted in Fig. 3. It is observed that LP-MMSE precoding is more robust to estimation noise than P-MMSE. This is because LP-MMSE precoding can take advantage of the spatial diversity available in a cell-free network. By using multiple APs to precode the signal, it can better adapt to variations and noise in the channel between the APs and the UEs. Moreover, LP-MMSE precoding operates on

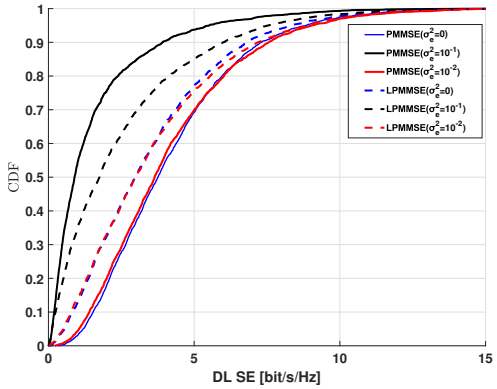

Figure 2: DL SE per UE with P-MMSE, LP-MMSE precoding with perfect and imperfect channel estimation when σ_e^2=0, 0.01, 0.1).

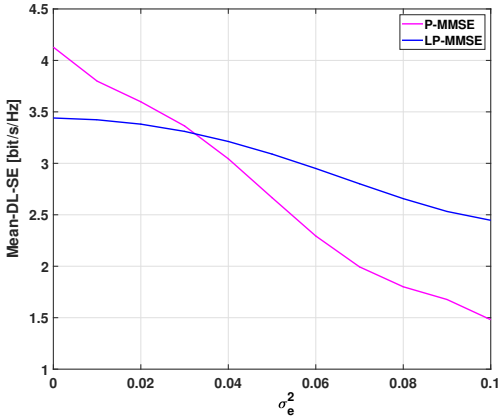

Figure 3: Average SE for DL transmission versus different values of σ_e^2.

the entire transmit signal, which includes multiple data streams. In contrast, P-MMSE precoding considers only the transmit signal for each UE and applies power constraints to the transmit signal at each AP independently.

Next, we investigate the impact of distortions induced by the PA non-linearities. As explained in Section 2.2, the PA model used in our work is derived from polynomial modeling based on real measurements of a 1 W PA operating in a 4 GHz frequency band. This model exhibits both AM/AM and AM/PM distortion, as shown in Fig. 4.

Fig. 5 illustrates the Cumulative Distribution Function (CDF) of the achievable per-user SE with LP-MMSE and P-MMSE precoders for various values of input back-off (IBO).

Notably, lower IBO values, such as $IBO = 2$ dB, correspond to curves shifted towards the left, indicating a higher probability of observing lower SE values due to increased non-linear distortion. Conversely, higher IBO values result in curves shifting towards the right, corresponding to a higher probability of achieving better SE performance. Comparing these results with those in Fig. 2, we can conclude that for an IBO value of 8 dB, we achieve performance comparable to the perfect CSI case ($\sigma_e^2 = 0$).

Fig. 6 illustrates the average DL SE across varying IBO values, ranging from 4 to 8 dB. An interesting pattern emerges from $IBO = 7$ dB, where we see a floor of ~ 3.3 bit/s/Hz and ~ 4.2 bit/s/Hz for LP-MMSE and P-MMSE precoding structures, respectively, corresponding to the performances achieved in the perfect CSI case ($\sigma_e^2 = 0$). This behavior is linked to the power amplifier working point. As we increase IBO, we move away from the amplifier's 1 dB compression point and work within the linear region. This amplification region helps reducing the distortion

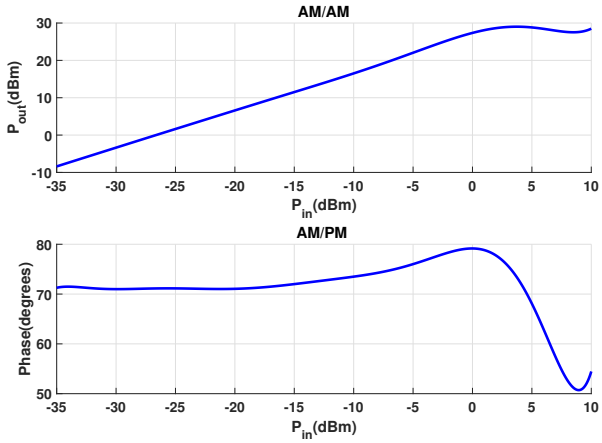

Figure 4: AM/AM and AM/PM characteristics of the 3GPP PA model [33].

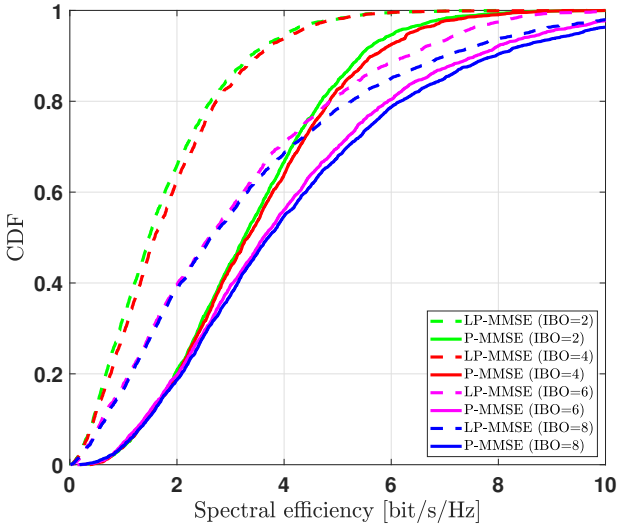

Figure 5: DL SE per UE with P-MMSE, LP-MMSE precoding considering non-linear PA operating at $IBO = 2, 4, 6, 8dB$.

caused by the amplifier while sacrificing the energy efficiency of the CF-mMIMO.

Additionally, when comparing the two different precoding schemes, LP-MMSE and P-MMSE, we observe a difference in how they handle amplifier distortion. LP-MMSE suffers more from distortion, with a noticeable drop in SE from 3.3 bit/s/Hz to 1.9 bit/s/Hz as IBO decreases from 8 to 4 dB. In contrast, P-MMSE maintains a consistently higher SE across different IBO values.

In summary, the dynamics observed in Fig. 6 underscore not only the crucial role of IBO settings in SE optimization but also highlight the differential impact of power amplifier non-linearity across various precoding schemes.

Next, we evaluate the performance of the proposed precoding schemes and observe the effects of the number of feedback bits on the DL SE. Fig. 7 shows the CDF of the achievable per-user SE with LP-MMSE and P-MMSE precoders considering RVQ for CSI with $B \in \{4, 8, 10, 16\}$ bits and perfect CSI. The solid lines represent outcomes achieved through the application of distributed LP-MMSE precoding. In contrast, the dashed lines illustrate results obtained in a centralized scenario, utilizing the P-MMSE precoding technique.

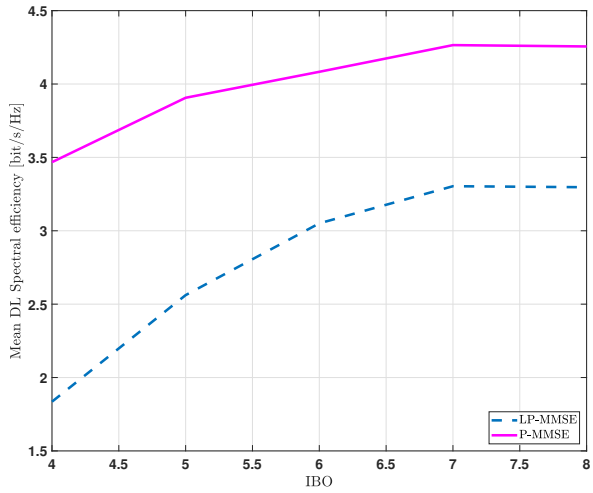

Figure 6: Average SE of DL transmission versus different values of IBO.

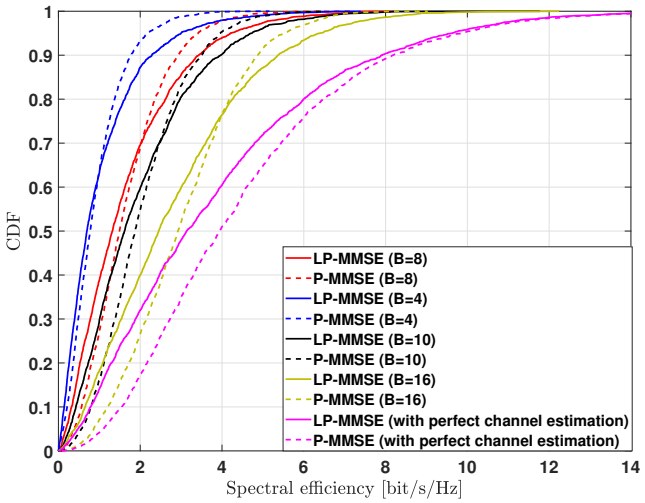

Figure 7: CDF of the DL SE per UE using P-MMSE and LP-MMSE precoders, when varying the number of feedback bits.

It can be seen from Fig. 7 that the system performance improves as the number of feedback bits increases. It is also worth noting that the gap between the two precoders grows with the number of bits. When $B = 16$, the system SE reaches 2.38 bit/s/Hz with P-MMSE and 2.06 bit/s/Hz with LP-MMSE, for 60% of UEs. Additionally, the results gradually converge towards the perfect scenario as B increases.

Fig. 8 illustrates the evolution of average DL SE as a function of the number of quantization bits B using RVQ codebooks. The number of bits B considered in the analysis ranges from 4 to 32. As seen in this figure, the DL SE, when using the LP-MMSE precoding scheme, outperforms that of the P-MMSE scheme for values of B between 4 and 8 bits. However, beyond $B = 8$, this trend reverses with higher values of B, where P-MMSE precoding exhibits superior performance as expected. For perfect CSI, average SEs of 3.9 bit/s/Hz and 4.5 bit/s/Hz can be achieved with LP-MMSE and P-MMSE precoders, respectively. Thus, it can be inferred that when $B = 32$, the achieved

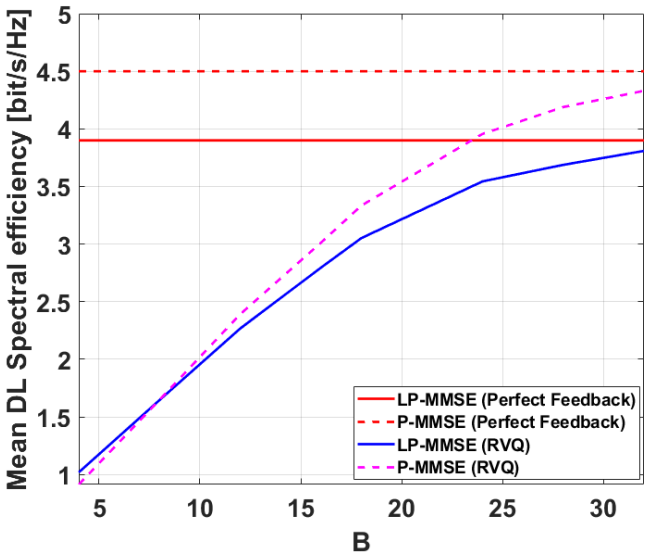

Figure 8: Average of DL SE with quantized feedback for B ranging from 4 up to 32 bits.

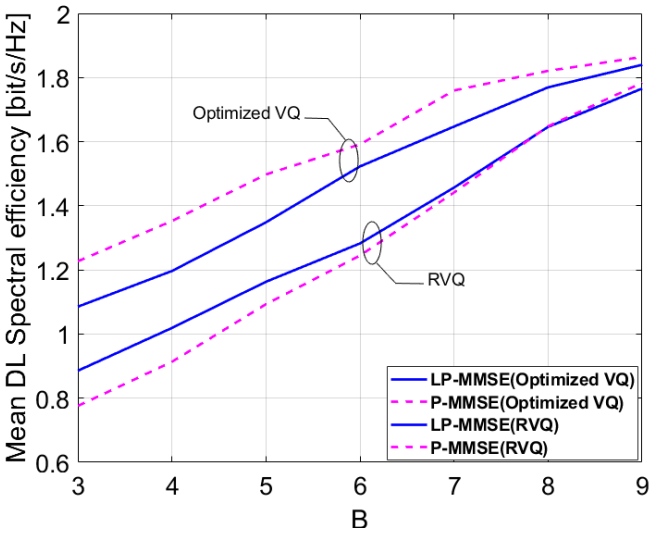

Figure 9: Performance comparison of the mean DL SE using both the RVQ and the optimized VQ codebooks for $B \in [3,4,5,6,7,8,9]$.

.

average SEs are close to those achieved in the perfect CSI cases.

Fig. 9 provides a performance comparison of the mean DL SE using both RVQ and optimized VQ codebooks for B ranging from 3 to 9 bits. We observe that the use of optimized codebooks improves the performance of both centralized and distributed precoding schemes. For instance, with $B = 4$, there is an increase of 0.44 bit/s/Hz for P-MMSE and 0.18 bit/s/Hz for LP-MMSE. Furthermore, it is clear that when using optimized VQ, P-MMSE outperforms LP-MMSE, which is not the case with RVQ. This is because with the RVQ approach, the codebooks are randomly generated, whereas in the optimized approach, the code words are designed to maximize system capacity.

3 CSI compression in the CF-mMIMO context

3.1 State of the Art on Channel State Information Compression in Massive MIMO Context

In the context of Massive MIMO and Cell-Free systems, CSI plays a crucial role in enabling high SE and reliable communication. However, the massive amount of CSI generated by large-scale antenna arrays introduces significant challenges in terms of feedback overhead and computational complexity. Efficient compression of CSI is, therefore, essential to ensure system scalability and performance. Over the years, various approaches have been proposed in the literature, each addressing the CSI compression problem from different perspectives.

3.1.1 Quantization-Based Approaches

Quantization remains one of the fundamental methods for CSI compression, aiming to reduce the precision of CSI data while minimizing the loss of information. Early methods such as *Scalar Quantization (SQ)* focus on individually quantizing each element of the CSI matrix, which, while simple, may lead to substantial quantization errors. As a refinement, *Vector Quantization (VQ)* considers the entire CSI vector for quantization, leading to more efficient compression [35]. Specifically, *Lloyd-Max Quantization* has been widely used, where an iterative algorithm is employed to optimize the quantization levels according to the distribution of the CSI data. This method is beneficial for reducing quantization noise and improving the quality of the reconstructed CSI [31]. Moreover, *Non-Uniform Quantization* techniques have been introduced to adapt the quantization steps to the variance of the CSI data, which is particularly effective in scenarios with heterogeneous channel conditions [40].

3.1.2 Compressed Sensing (CS) Techniques

Compressed Sensing (CS) has gained significant traction as a method for CSI compression, particularly due to its ability to exploit the inherent sparsity in high-dimensional CSI data. CS-based methods reduce the number of required measurements for accurate CSI reconstruction, thereby alleviating the feedback burden in massive MIMO systems. In [17], a *CS-based CSI Feedback Scheme* was proposed that leverages the sparsity of CSI in the angular domain, significantly reducing the amount of data that needs to be transmitted back to the base station. Another notable approach is the *Distributed Compressed Sensing (DCS)*, which extends the traditional CS framework to exploit both intra- and inter-antenna correlations in distributed antenna systems. DCS offers improved compression efficiency and reconstruction accuracy by considering the collective sparsity of the CSI across multiple antennas [10].

3.1.3 Dimensionality Reduction Techniques

To further reduce the dimensionality of the CSI data, techniques such as *Principal Component Analysis (PCA)* and *Singular Value Decomposition (SVD)* have been employed. *PCA* works by identifying the principal components of the CSI matrix, which represent the directions of maximum variance, and projecting the CSI onto these components. This reduces the size of the CSI matrix, thereby lowering the feedback overhead [27]. However, the challenge lies in selecting an appropriate number of principal components to avoid losing critical information. *SVD* decomposes the CSI matrix into a set of singular values and corresponding vectors, and by truncating the smaller singular values, the CSI can be compressed [27]. SVD is particularly effective in reducing the dimensionality of CSI while preserving most of the channel's energy, making it a valuable tool in MIMO systems where channel prediction and estimation are critical.

3.1.4 Deep Learning-Based Approaches

The advent of deep learning has opened new avenues for CSI compression. *Autoencoders*, a type of neural network designed for unsupervised learning, have been applied to compress CSI by encoding it into a low-dimensional latent space and then decoding it back to reconstruct the original CSI [37]. This approach leverages the ability of neural

networks to capture complex patterns in the data, leading to efficient compression. *Recurrent Neural Networks (RNNs)*, known for their ability to handle sequential data, have also been used for CSI compression in time-varying channels [24]. RNNs can model temporal correlations in the CSI data, enabling them to predict future CSI based on past observations. This reduces the need for frequent CSI feedback, making them highly effective in dynamic wireless environments. **Autoencoders** and other advanced neural network models, such as *Convolutional Neural Networks (CNNs)* [21], have also been employed to capture complex spatial correlations in the CSI matrix. This approach takes advantage of the spatial structure in the CSI, leading to more efficient compression compared to traditional methods. Furthermore, hybrid models combining *autoencoders* with techniques like *compressed sensing* or *quantization* have been proposed to further improve the efficiency and accuracy of CSI compression in Massive MIMO systems [9].

4 Local PAPR-Aware Precoding for Energy-Efficient Cell-Free Massive MIMO-OFDM Systems

4.1 Introduction

In this work, we present an energy-efficient downlink transmission scheme adequate to the emerging CF-mMIMO-OFDM technology. In this first deliverable we give a description of the system model, first analysis as well as simulations results.

4.2 System, Channel and hardware impairment models

4.2.1 System Model

In this section, we study an OFDM based cell-free massive MIMO downlink system operating in time division duplex (TDD) mode. Let there be L APs, each equipped with M antennas, serving coherently K single-antenna users, where $LM \gg K$. We assume that all the users and the APs are randomly located in a geographic area, as shown in Fig. 10. Frequency-selective channels are considered and the K UEs are arbitrary distributed in the coverage area. An OFDM modulator is implemented at each antenna branch where the total number of subcarrier is M_{FFT}, which is the IFFT/FFT size, while the number of used subcarrier is $N \leq M_{FFT}$.

The TDD mode separates the downlink (DL) and uplink (UL) transmissions (see Fig. 11) with the assumption of perfect channel reciprocity which can be assured by accurate calibration methods [36]. Moreover, the transmission of a TDD frame in OFDM based CF-mMIMO is performed within the coherence interval and the physical resource block (RB) width is smaller than the coherence bandwidth.

To be compliant with the 5G NR standard, let us consider a radio frame whose time-frequency resource is divided into N_{rb} resource blocks. Each RB comprises $N_{sc} = N/N_{rb}$ consecutive subcarriers. We note by $(t,n)_{l,m}$ the resource unit (RU), which represents the smallest time-frequency resource of the n-th subcarrier of the t-th OFDM symbol corresponding to the m-th antenna of the l-th AP.

Here, the TDD frame contains N_c OFDM symbols that fits the shortest coherence interval of all the users. It corresponds to the transmission of $\tau_c = N_{sc}N_c$ RUs per RB, where τ_p of them are used as pilots which are distributed among the UL payload transmission, as shown by Fig. 11. These pilot data will serve to estimate MK frequency-domain channels, per RB, at each AP. Consequently, we leave $N_D = N_{sc}N_c - \tau_p$ for payload data, in samples per RB, that will be split between DL and UL transmissions as ξN_D and $(1-\xi)N_D$, respectively, where $0 < \xi < 1$.

One of the major drawbacks of OFDM is related to the fact that its time-domain signal has very high amplitude fluctuations, making it power-hungry as well as very sensitive to hardware impairments, like PA saturation. These fluctuations are characterized by the peak-to-average power ratio (PAPR) [19] [20], which is defined as the ratio of the highest signal peak power and its average power value. Then, the PAPR, corresponding to the signal at antenna m of AP l, can be given by

$$PAPR(\mathsf{a}_{l,m}^t) = \frac{\underset{0\leq t\leq ON-1}{max}\left[|\mathsf{a}_{l,m}(t)|^2\right]}{E\left\{|\mathsf{a}_{l,m}(t)|^2\right\}}, \tag{32}$$

where $\mathsf{a}_{l,m}^t = \left[\mathsf{a}_{l,m}(0),...,\mathsf{a}_{l,m}(ON-1)\right]$ and O denotes the oversampling factor.

4.2.2 Local MMSE Channel estimation

We denote by $\mathbf{h}_{l,k,n} \in \mathbb{C}^{M\times 1}$ the frequency-domain channel response between the AP l and the user k on the n-th subcarrier, where $n = 0,...,N-1$. The channels are modeled using independent Rayleigh fading, i.e., $\mathbf{h}_{l,k,n} \sim \mathscr{CN}(\mathbf{0}, \beta_{l,k}\mathbf{I}_M)$, where $\beta_{l,k}$ is the large-scale fading coefficient between AP l and UE k, which is antennas and subcarriers independent.

Note that we consider block-fading channels, which are constant during a time-frequency interval, known as the *coherence interval*, and varies independently between coherence intervals. Besides, we assume that the large-

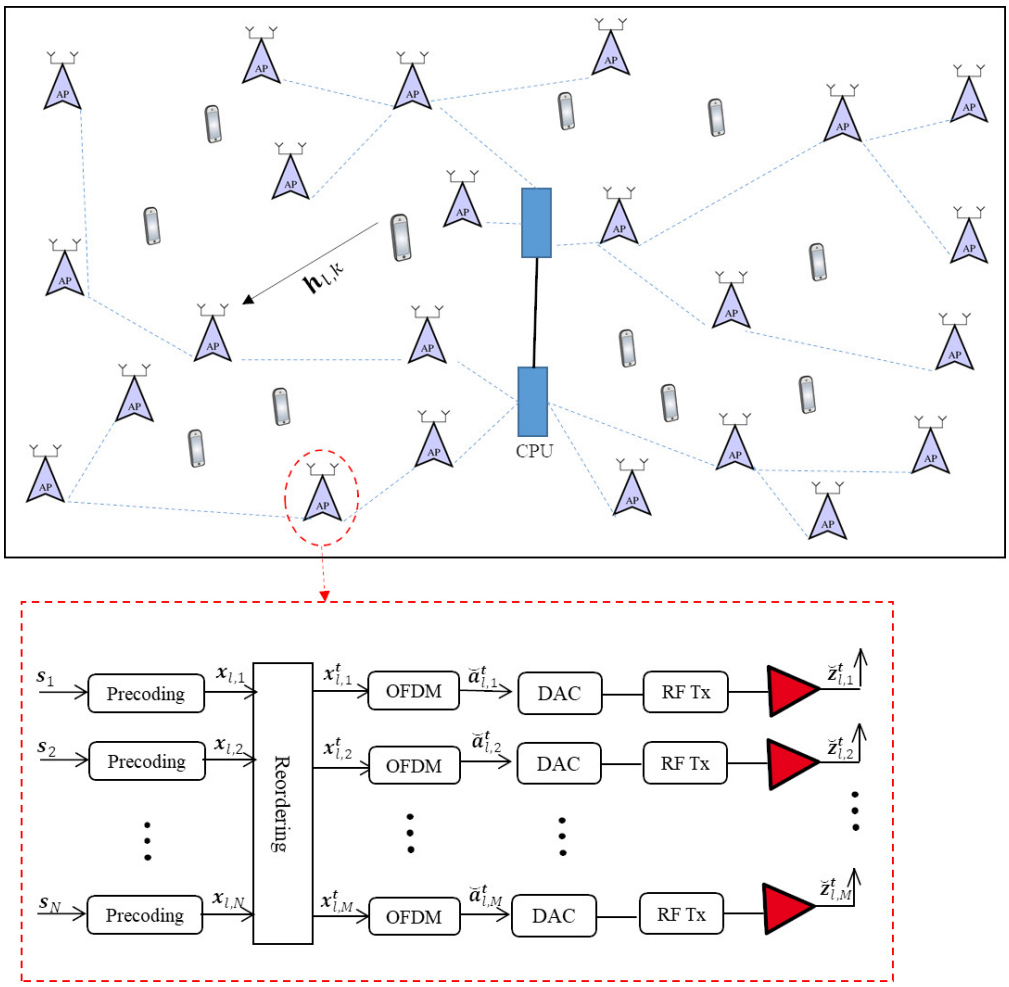

Figure 10: System model of the OFDM based CF-mMIMO with segmented fronthaul: M transmit antennas at the AP, K independent single-antenna terminals. The serial fronthaul, which interconnects the APs, is highlighted in blue.

scale fading coefficients vary slowly, in range of several coherence intervals. It is therefore obvious to consider that the channel gains are known a-priori at each AP and are used to estimate the current channel responses.

The channel estimation is performed locally at each AP to preserve the network scalability. In this investigation, we adopt the *minimum mean square error* (MMSE) channel estimation method and the *pilot-to-precoder mapping* strategy inspired from the work in [15]. Hence, the frequency-domain channel estimate between the l-th AP and the k-th UE, corresponding to the $n\prime$-th RB can be given by

$$\hat{\mathbf{h}}_{n\prime,l,k} = c_{l,k}\bar{\mathbf{H}}_{n\prime,l}\mathbf{e}_{\mathbf{i_k}}, \tag{33}$$

where

- $c_{l,k}$ is a frequency independent scalar, which is defined as [18]

$$c_{l,k} \triangleq \frac{\sqrt{\eta^u_k}\beta_{l,k}}{\tau_p \sum_{t\in\mathscr{P}_k}\eta^u_t\beta_{l,t}+1} \tag{34}$$

- $\bar{\mathbf{H}}_{n\prime,l} = \mathbf{Y}_{n\prime,l}\blacksquare \in \mathbb{C}^{M\times\tau_p}$ denotes the corresponding full-rank matrix of the frequency-domain channel estimates, where $\mathbf{Y}_{n\prime,l}$ represents the frequency-domain pilot signal received at the l-th AP on the $n\prime$-th RB.

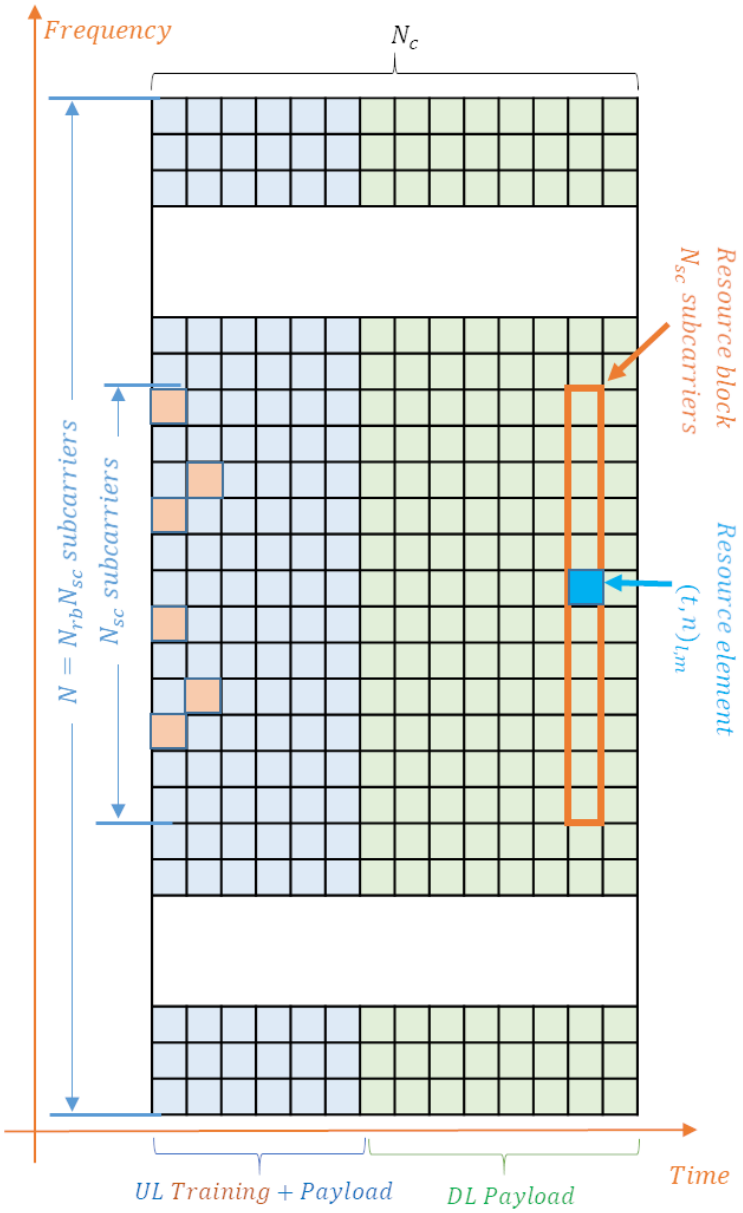

Figure 11: A radio frame with N_c time-domain OFDM symbols and N subcarriers in frequency-domain that are grouped into N_{rb} RBs. The time-frequency resource grid of a single RB with N_{sb} subcarriers is shown. The UL training, UL payload and DL payload are also shown.

- $\mathbf{e}_{\mathbf{i_k}}$ is the i_k-th column of $\mathbf{I}_{\tau_p}$,

Note that the channel estimates and estimation errors, corresponding to the $n\prime$-th RB, which are denoted by $\hat{\mathbf{h}}_{n\prime,l,k}$ and $\tilde{\mathbf{h}}_{n\prime,l,k} = \mathbf{h}_{n\prime,l,k} - \hat{\mathbf{h}}_{n\prime,l,k}$, are independent and distributed as $\hat{\mathbf{h}}_{n\prime,l,k} \sim \mathcal{CN}\left(\mathbf{0}, \gamma_{l,k}\mathbf{I}_M\right)$, $\tilde{\mathbf{h}}_{n\prime,l,k} \sim \mathcal{CN}\left(\mathbf{0}, \left(\beta_{l,k} - \gamma_{l,k}\right)\mathbf{I}_M\right)$, where

$$\gamma_{l,k} \triangleq \frac{\eta_k^u \tau_p \beta_{l,k}^2}{\tau_p \sum_{t \in \mathscr{P}_k} \eta_t^u \beta_{l,t} + 1} \tag{35}$$

Note that $\gamma_{l,k}$, which is the mean-square of the channel estimate between the AP l and the UE k, is antenna and

frequency independent.

Remark 1 *(Per-RB to per-subcarrier channel mapping)* The channels at AP l are estimated per-RB and are assumed to be identical for RUs within one RB. Then, each AP can effectively construct τ_p channel coefficients per-subcarrier n as

$$\hat{\mathbf{h}}_{n,l,k} = \hat{\mathbf{h}}_{n\prime,l,k} \text{ for } n \in \text{RB}_{n\prime}, \tag{36}$$

we can also write

$$\bar{\mathbf{H}}_{n,l} = \bar{\mathbf{H}}_{n\prime,l} \text{ for } n \in \text{RB}_{n\prime}. \tag{37}$$

4.2.3 Hardware Impairment Model

In the case of OFDM based CF-mMIMO, the modulated signals $\{\mathbf{a}^t_{l,m}, \forall m\}$ are fed, towards a given antenna, through a transmit radio-frequency (RF) chain. This RF chain is commonly nonlinear due to hardware impairments, which are dominated by nonlinear power amplifiers (PAs). These latter can be characterized by amplitude-to-amplitude (AM/AM) and amplitude-to-phase (AM/PM) conversions. According to the modified Rapp model [28] which has the advantage of exhibiting greater simplicity and accuracy than other models, AM/AM and AM/PM conversions of the solid state power amplifier (SSPA) can be represented as follow:

$$F_A(\rho) = \frac{G\rho}{\left(1 + \left|\frac{G\rho}{V_{sat}}\right|^{2p}\right)^{\frac{1}{2p}}}, \qquad F_P(\rho) = \frac{A\rho^q}{\left(1 + \left(\frac{\rho}{B}\right)^q\right)}, \tag{38}$$

Where ρ is a modulus of the input signal, G is PA gain in linear region, V_{sat} is PA saturation at the output, p is the smoothness factor that controls the nonlinearity and A, B and q are fitting parameters. The output of the PA can be expressed as :

$$z = F_A(\rho)\exp^{j(\phi + F_P(\rho))}, \tag{39}$$

where ϕ is the phase of the input signal. Figure 12 depicts amplitude and phase distortions caused by PA models as modified Rapp or simple clipping for $G = 16, Vsat = 1.9, p = 1.1, A = -345, B = 0.17, q = 4$. Note that in the clipping model there is no phase distortion i.e. no AM/PM distortion is modelled. In modified Rapp, which resembles more closely to realistic PAs, both amplitude (AM/AM) and phase (AM/PM) distortions are modelled.

Let us consider the resulting transmit signal at a given transmit antenna (UE/AP side) can be written as

$$\mathbf{z} = \mathbf{f}(\mathbf{a}) \tag{40}$$

where $\mathbf{f}(.)$ represents the nonlinear operation which is assumed to be memoryless and identical over all antennas of all APs/UEs, for the sake of simplicity. Specifically, the nonlinear behavior of the PA can be modeled by the following full-rank polynomial

$$z(t) = \sum_{i=1}^{I} \lambda_i a(t)|a(t)|^{i-1}, \tag{41}$$

where $\lambda_1, ..., \lambda_I$ are complex-valued model parameters that capture both amplitude-to-amplitude (AM/AM) and amplitude-to-phase (AM/PM) conversions.

According to the Bussgang theorem [32], the time-domain OFDM signal at the output of the nonlinear function can be expressed by the unique decomposition as

$$\mathbf{z} = K_0\mathbf{a} + \mathbf{d}, \tag{42}$$

where K_0 is a complex gain which is frequency independent and $\mathbf{d}$ is a zero mean noise and with σ_d^2 variance, which is uncorrelated with $\mathbf{a}$. This latter is not Gaussian, but at the receiver side, after the OFDM demodulation, it becomes Gaussian [8].

Note that these hardware impairments (HWI) parameters, K_0 and σ_d^2, can be computed analytically, letting us to be able to develop closed-form expressions to evaluate the in-band performance of CF-mMIMO-OFDM systems [7]. The closed-form expressions for the NLD parameters in terms of the model PA parameters $\{\lambda_i\}$ are given in equations (43) and (44), where σ is the standard deviation of the PA input signal. Readers interested in details about computing these equations can refer to [7].

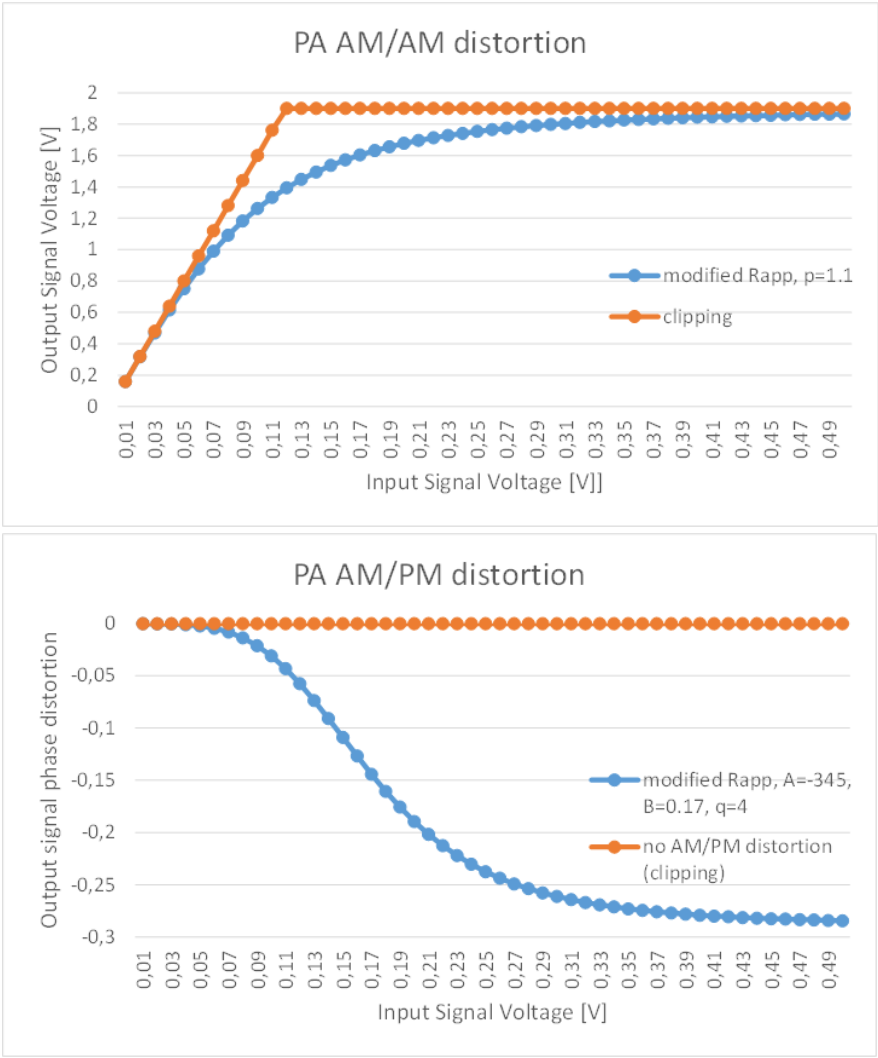

Figure 12: Power Amplifier AM/AM and AM/PM distortions for clipping (only AM/AM distortion) and modified Rapp model (including both AM/AM and AM/PM distortions) [28].

$$K_0 = \lambda_1 + \sqrt{\frac{\pi}{8}} \sum_{i=2,\ i\ even}^{I} (i+1)\lambda_i \sigma^{i-1} \prod_{i\prime=0}^{\frac{i-2}{2}} (2i\prime+1) + \frac{1}{2} \sum_{i=3,\ i\ odd}^{I} (i+1)\lambda_i (\sqrt{2}\sigma)^{i-1} (\frac{i-1}{2})! \tag{43}$$

$$\begin{aligned} \sigma_d^2 = &\sum_{i=1}^{I} |\lambda_i|^2 2^i \sigma^{2i} i! - 2|K_0|^2 \sigma^2 + \sqrt{\frac{4\pi}{2}} \sum_{i,l=1, i\neq l,\ (i+l)\ odd}^{I} \Re[\lambda_i \lambda_l^*] \sigma^{i+l} \prod_{i\prime=0}^{\frac{i+l-1}{2}} (2i\prime+1) \\ &+ 2 \sum_{i,l=1, i\neq l,\ (i+l)\ even}^{I} \Re[\lambda_i \lambda_l^*] (\sqrt{2}\sigma)^{i+l} (\frac{i+l}{2})!, \end{aligned} \tag{44}$$

4.3 Downlink CF-mMIMO-OFDM under HWI

4.3.1 Downlink CF-mMIMO-OFDM Data Transmission

Local Precoding schemes Precoding is needed to be performed at each AP in order to remove the multi-user interference (MUI) at receivers. In CF-mMIMO, the local nature of the implemented precoders is crucial to preserve system scalability. To stress this aspect, we use the terminology *local* precoding.

The signal vectors $\mathbf{s}_n, \forall n$, can be linearly precoded, at the l-th AP, as

$$\mathbf{x}_{l,n} = \mathbf{W}_{l,n}\mathbf{P}_l\mathbf{s}_n, \tag{45}$$

where:

- $\mathbf{x}_{l,n} \in \mathbb{C}^{M\times 1}$ denotes the precoded vector that contains samples to be transmitted over the the n-th subcarrier through the M antennas of AP l,
- $\mathbf{W}_{l,n} \in \mathbb{C}^{M\times K}$ is the precoding matrix corresponding to the n-th subcarrier at the l-th AP,
- $\mathbf{P}_l \in \mathbb{C}^{K\times K}$, which is frequency independent, represents a diagonal matrix whose elements $\sqrt{\eta_{l,k}}, k = 1, ..., K$ are the normalized transmit powers allocated to the K UEs.

To be compliant with the existing literature, we consider max-min fairness power control that consists in maximizing the lowest user's downlink SE. Then, the *normalized* transmit power (normalized by the noise power σ_b^2) at each AP l, denoted by $\|\mathbf{x}_{l,n}\|^2 = \sum_{k=1}^{K} \eta_{l,k}, \forall n$ has to be constrained as $\sum_{k=1}^{K} \eta_{l,k} \leq \eta_l^{max}$, where η_l^{max} represents the maximum sum power at the input of the different PAs at each AP. It can be written as

$$\eta_l^{max} = \frac{1}{\sigma_b^2}\frac{M(V_{sat}/G)^2}{\text{IBO}}, \tag{46}$$

where the IBO denotes the input back-off which represents the PA input power level relative to input saturation power, V_{sat} is PA amplitude saturation level, G is the PA linear gain and $\frac{(V_{sat}/G)^2}{\text{IBO}}$ represents the maximum power at the input of each PA [22] [39].

The power control optimization problem can be formulated as follows [15] [14]

$$\begin{aligned} &\underset{\{\eta_{l,k}\geq 0\}}{maximize}\ \min_{k}\ \text{SINR}_{k,n} \\ &s.t.\ \sum_{k=1}^{K}\eta_{l,k} \leq \eta_l^{max},\ \forall l \end{aligned} \tag{47}$$

where $\text{SINR}_{k,n}$ denotes the signal-to-interference plus noise ratio (SINR) for user k at subcarrier n.

For the sake of self-containment, we consider, in this paper, the distributed power control policy where the power control coefficients are given in [15] as

$$\eta_{l,k} = \frac{\gamma_{l,k}}{\sum_{i=1}^{K}\gamma_{l,i}}\eta_l^{max},\ \forall l, \forall k \tag{48}$$

Inspired by [15] [14] that study *local* precoding schemes in the downlink, we will consider the following linear precoders in frequency-domain of our OFDM based CF-mMIMO: (i) *local* full-pilot zero-forcing (FZF) and (ii) *local* regularized zero-forcing (RZF). Here, the precoding vector employed by the l-th AP towards the k-th UE on subcarrier n, denoted by $\mathbf{w}_{l,n,i_k} \in \mathbb{C}^{M\times 1}$, is given by

$$\mathbf{w}_{l,n,i_k} = \begin{cases} \dfrac{\bar{\mathbf{H}}_{l,n}\left(\bar{\mathbf{H}}_{l,n}^H\bar{\mathbf{H}}_{l,n}\right)^{-1}\mathbf{e}_{i_k}}{\sqrt{E\left\{\|\bar{\mathbf{H}}_{l,n}\left(\bar{\mathbf{H}}_{l,n}^H\bar{\mathbf{H}}_{l,n}\right)^{-1}\mathbf{e}_{i_k}\|^2\right\}}} = \sqrt{(M-\tau_p)\rho_{l,k}}\bar{\mathbf{H}}_{l,n}\left(\bar{\mathbf{H}}_{l,n}^H\bar{\mathbf{H}}_{l,n}\right)^{-1}\mathbf{e}_{i_k}, & \text{for } \textit{local} \text{ FZF} \\ \dfrac{\hat{\mathbf{H}}_{l,n}\left(\hat{\mathbf{H}}_{l,n}^H\hat{\mathbf{H}}_{l,n}+P_l^{-1}\right)^{-1}\mathbf{e}_k}{\sqrt{E\left\{\|\hat{\mathbf{H}}_{l,n}\left(\hat{\mathbf{H}}_{l,n}^H\hat{\mathbf{H}}_{l,n}+P_l^{-1}\right)^{-1}\mathbf{e}_k\|^2\right\}}}, & \text{for } \textit{local} \text{ RZF} \end{cases}$$

where:

- $\rho_{l,k}$, which is the mean square of $\bar{\mathbf{H}}_{l,n}\mathbf{e}_{i_k}$, can be defined as $\rho_{l,k} = \gamma_{l,k}/c_{l,k}^2$ [15],
- $\hat{\mathbf{H}}_{l,n} \in \mathbb{C}^{M\times K}$ is the matrix of the channel estimates between the AP l and the K users on the subcarrier n, which can be collected as $\hat{\mathbf{H}}_{l,n} = \left[\hat{\mathbf{h}}_{l,n,1}, ..., \hat{\mathbf{h}}_{l,n,K}\right]$,
- $\mathbf{P}_l \in \mathbb{R}^{K\times K}$ is a diagonal regularization matrix,
- $\mathbf{e}_{i_k}$ is the k-th column of $\mathbf{I}_K$, i.e, taking the k-th column of $\hat{\mathbf{H}}_{l,n}\left(\hat{\mathbf{H}}_{l,n}^H\hat{\mathbf{H}}_{l,n}\right)^{-1}$.

***Remark 2.1**:* It is worth mentioning that any AP can design the different aforementioned precoders by only using its local CSI. Exceptionally for the *local* FZF, it works only when $M \geq \tau_p + 1$ as explained in [15] and [14].

Received signal The frequency-domain received signal at UE k and subcarrier n can be given by

$$\begin{aligned} y_{k,n} &= \sum_{l=1}^{L} \mathbf{h}_{l,k,n}^H \mathbf{z}_{l,n} + b_{k,n} \\ &= \sum_{l=1}^{L} \mathbf{h}_{l,k,n}^H \mathbf{K}_0 \mathbf{x}_{l,n} + \sum_{l=1}^{L} \mathbf{h}_{l,k,n}^H \mathbf{d}_{l,n} + b_{k,n}, \end{aligned} \tag{49}$$

where $\mathbf{z}_{l,n} \in \mathbb{C}^{M\times 1}$ denotes the frequency-domain amplified signal transmitted by AP l on subcarrier n, $\mathbf{d}_{l,n} \in \mathbb{C}^{M\times 1}$ is the frequency-domain version of hardware-related noise, $\mathbf{K}_0$ is the $M \times M$ diagonal matrix whose elements are equal to K_0 and $b_{k,n} \sim \mathscr{CN}(0,1)$ is an i.i.d. Gaussian noise.

By plugging (45) in (49), $y_{k,n}$ can be expanded as

$$\begin{aligned} y_{k,n} &= \sum_{l=1}^{L} \mathbf{h}_{l,k,n}^H \mathbf{K}_0 \mathbf{W}_{l,n} \mathbf{P}_l \mathbf{s}_n + \sum_{l=1}^{L} \mathbf{h}_{l,k,n}^H \mathbf{d}_{l,n} + b_{k,n} \\ &= \underbrace{\sum_{l=1}^{L} \sqrt{\eta_{l,k}} \mathbf{h}_{l,k,n}^H \mathbf{K}_0 \mathbf{w}_{l,n,k} s_{k,n}}_{\text{Desired signal}} + \underbrace{\sum_{l=1}^{L}\sum_{t\neq k}^{K} \sqrt{\eta_{l,t}} \mathbf{h}_{l,k,n}^H \mathbf{K}_0 \mathbf{w}_{l,n,i_t} s_{t,n}}_{\text{Multi-user interference}} + \underbrace{\sum_{l=1}^{L} \mathbf{h}_{l,k,n}^H \mathbf{d}_{l,n}}_{\text{HWI}} + \underbrace{b_{k,n}}_{\text{Noise}} \end{aligned} \tag{50}$$

4.3.2 Downlink Spectral Efficiency

The per-RB and per-user SE in the downlink can be computed using [16] as

$$\text{SE}_k = \xi \left(1 - \frac{\tau_p}{N_{sc}N_c}\right) N_{sc}\Delta f \log_2\left(1 + \text{SINR}_{k,n}\right), \tag{51}$$

with the effective signal-to-interference-plus-noise ratio (SINR) of UE k at subcarrier n is given by

$$\text{SINR}_{k,n} = \frac{\left|\sum_{l=1}^{L} \sqrt{\eta_{l,k}} \text{E}\left\{\mathbf{h}_{l,k,n}^H \mathbf{K}_0 \mathbf{w}_{l,n,k}\right\}\right|^2}{\sum_{t=1}^{K} \text{E}\left\{\left|\sum_{l=1}^{L} \sqrt{\eta_{l,t}} \mathbf{h}_{l,k,n}^H \mathbf{K}_0 \mathbf{w}_{l,n,i_t}\right|^2\right\} - \left|\sum_{l=1}^{L} \sqrt{\eta_{l,k}} \text{E}\left\{\mathbf{h}_{l,k,n}^H \mathbf{K}_0 \mathbf{w}_{l,n,k}\right\}\right|^2 + \text{E}\left\{\left|\sum_{l=1}^{L} \mathbf{h}_{l,k,n}^H \mathbf{d}_{l,n}\right|^2\right\} + 1} \tag{52}$$

Note that this SE expression is valid for all precoding schemes presented in (4.3.1) and analytical expressions concerning the FZF scheme can be written as stated in **Corollary 2.1**. Note that a closed-form expression for the *local* RZF is intractable and only results, performed by Monte-Carlo (MC) simulations, are presented regarding this latter precoding scheme.

Corollary 2.1*:* The per-user downlink SE (bit/s) is obtained for FZF precoding as in equation (51), where the $\text{SINR}_{k,n}$ can be expressed as given by

$$\text{SINR}_{k,n} = \frac{(M-\tau_p)|K_0|^2 \left(\sum_{l=1}^{L} \sqrt{\eta_{l,k}\gamma_{l,k}}\right)^2}{(M-\tau_p)|K_0|^2 \sum_{t\in\mathscr{P}_k\setminus\{k\}} \left(\sum_{l=1}^{L} \sqrt{\eta_{l,t}\gamma_{l,k}}\right)^2 + |K_0|^2 \sum_{l=1}^{L}\sum_{t=1}^{K} \eta_{l,t}(\beta_{l,k}-\gamma_{l,k}) + \Psi_{k,n} + 1} \tag{53}$$

where $\Psi_{k,n}$ denotes the NLD variance that can be expressed as

$$\begin{aligned}\Psi_{k,n} &= \text{E}\left\{\left|\sum_{l=1}^{L} \mathbf{h}_{l,k,n}^{H} \mathbf{d}_{l,n}\right|^2\right\} \\ &= \sum_{l=1}^{L}\sum_{l\prime=1}^{L}\sum_{m=1}^{M}\sum_{m\prime=1}^{M} \text{E}\left\{h_{l,k,n,m}^{*} d_{l,n,m} d_{l\prime,n,m\prime}^{*} h_{l\prime,k,n,m\prime}\right\}\end{aligned} \tag{54}$$

Note that $\Psi_{k,n}$ depends on both intra-user and inter-user nonlinear distortions, which can not be eliminated by using conventional transmit precoding techniques. Moreover, the behavior of this NLD depends on many parameters, such as number of users, number of pilots, number of APs, number antennas per-AP, precoding scheme, channel estimation, pilot contamination and essentially the adopted PA and its operating point.

In order to derive a closed-from expression for the SINR of the k-th user at the n-th subcarrier, we have to compute the terms in (54). Note that due to the possibility of NLD correlation, the latter are not necessary null even when $l \neq l\prime$ and $m \neq m\prime$. But, it is still possible to have closed-form expressions, for $\Psi_{k,n}$, for some scenarios as stated in ***Lemma 2.1***, corresponding to FZF precoding.

Lemma 2.1 *(NL OFDM based CF-mMIMO with local FZF precoding):* With FZF precoding, when $\tau_p \to M$, Ψ_k can be expressed as

$$\Psi_k^{FZF} = \frac{M}{\sigma_b^2} \sum_{l=1}^{L} \beta_{l,k} \sigma_d^2, \tag{55}$$

and achievable SE for the k-th UE can be defined as stated in ***Theorem 2.1***.

Theorem 2.1*:* In OFDM based CF-mMIMO adopting nonlinear PA and FZF precoding ($\tau_p \to M$), the downlink capacity of the k-th UE is lower bounded by the expression given by equation (56),

$$\text{SE}_k = \chi N_{sc} \Delta f \log_2\left(1 + \frac{(M-\tau_p)|K_0|^2 \left(\sum_{l=1}^{L} \sqrt{\eta_{l,k}\gamma_{l,k}}\right)^2}{(M-\tau_p)|K_0|^2 \sum_{t\in\mathscr{P}_k\setminus\{k\}} \left(\sum_{l=1}^{L} \sqrt{\eta_{l,t}\gamma_{l,k}}\right)^2 + |K_0|^2 \sum_{l=1}^{L}\sum_{t=1}^{K} \eta_{l,t}(\beta_{l,k}-\gamma_{l,k}) + \Psi_k^{FZF} + 1}\right) \tag{56}$$

where $\chi = \xi\left(1 - \frac{\tau_p}{N_{sc}N_c}\right)$, which denotes the percentage of DL payload data. We recall that closed-form expression for σ_d^2 is given in (44).

4.4 Simulation results

This section provides numerical results to evaluate the performance of OFDM based cell-free massive MIMO in terms of DL SE, in presence of nonlinear power amplifiers.

4.4.1 Simulation Scenario

Let us introduce the considered CF-mMIMO-OFDM system setup. We assume that the L APs and the K UEs are independently and uniformly distributed within an area of size $D \times D$ squared meters, which is wrapped around at the edges to avoid boundary effects.

We consider an uncoded OFDM with IFFT/FFT of $N = 256$ and use a spectral map Ξ, in which $N_a = |\Xi| = 248$ subcarriers are used for data transmission and $N_{gb} = |\Xi^c| = 4$ are used as guard-band at each side. Thus $N = N_a + 2N_{gb} = 256$. It is worth mentioning that a 4-oversampling factor is considered in the time-domain to evaluate the PAPR accurately. Note that we consider an OFDM radio frame of duration 1 ms (corresponding to the channel

Table 1: Simulation settings.

Description	Value	Description	Value
D (simulation area)	$1000 \times 1000 m^2$	AP/UE distribution	ind. unif. rand.
Carrier frequency	2 GHz	AP/UE antenna height	$10/1.5$ m
Sbc bandwidth (Δf)	15 kHz	Coherence bandwidth	180 kHz
Coherence time	1 ms	τ_d (RUs)	168
ξ	0.5	σ_{sh}	4 dB
σ_b^2	-92 dBm	Bandwidth	15.36 MHz
$\eta_k^u \; \forall \; k$	100 mW	P_c	0.1 W
P_0	0.1 W	P_{bt}	0.25 W/(Gbits/s)

coherence time), which consists of 14 OFDM symbols. From the frequency-domain perspective, each RB contains $N_{rb} = 12$ subcarriers with 15 kHz of bandwidth each.

The large-scale fading coefficients $\{\beta_{l,k}\}$ are modeled as [15]

$$\beta_{l,k} = \text{PL}_{l,k}.10^{\frac{\sigma_{sh} z_{l,k}}{10}}, \tag{57}$$

where $\text{PL}_{l,k}$ denotes the path-loss and $10^{\frac{\sigma_{sh} z_{l,k}}{10}}$ models log-normal shadow fading with standard deviation σ_{sh} and $z_{l,k} \sim \mathcal{N}(0,1)$. Note that, in this investigation, we consider the 3GPP Urban Microcell path-loss model, which is given by equation (58) when assuming a 2 GHz carrier frequency [15] [1].

$$\text{PL}_{l,k}\,[\text{dB}] = -30.5 - 36.7 \log_{10}\left(\frac{d_{l,k}}{1\ m}\right), \tag{58}$$

where $d_{l,k}$ denotes the distance between the l-th AP and the k-th UE including AP and UE's heights.

The simulation settings are reported in Table 1. Note that the PA characteristics are modeled by the memoryless modified Rapp model (see chapter 2), defined by the 3GPP in [28], with parameters $G = 16$, $V_{sat} = 1.9$, $p = 1.1$, $A = -345$, $B = 0.17$ and $q = 4$. Its corresponding complex-valued polynomial model parameters, which capture the AM/AM and AM/PM conversions, can be computed as in [7].

Finally, random pilot assignment is adopted for simplicity. So, each UE randomly select a pilot sequence from a predefined set of $\tau_p \leq K$ orthogonal pilot sequences of length τ_p samples.

4.4.2 Performance Evaluation

To evaluate the performance of the considered CF-mMIMO-OFDM system, we adopt the cumulative density function (CDF) of the SE, which corresponds to the SE values collected over different random realizations of the AP/UE locations for 500 network snapshots.

CF-mMIMO-OFDM Downlink under PA nonlinearity Fig. 13(a) shows the CDFs of the SEs achieved by the CF-mMIMO-OFDM adopting the *local* FZF precoding scheme and running under PA non-linearity. Both results achieved with ideal PA (black curve) and non-linear PA (with different values of IBO) are presented.

First, one can observe that the results obtained in closed-form (solid curves) and by Monte-Carlo simulations (markers) are in good match, which numerically validates our derived closed-form expression in equation (56) for three IBO values. Second, we can see the significant impact of the PA non-linearity on the SE performance, especially for high percentiles. This can be explained by the fact that the PA distortion is beam-formed towards UEs with good channel conditions, since the control power scheme boosts the energy towards these UEs.

In Fig. 13(b), the CDFs of SEs achieved by the *local* RZF are given. One can clearly note the higher sensitivity of the *local* RZF precoding scheme to the PA non-linearity compared with the FZF one. It is clearer now that only UEs with good channels (i.e., upper SE percentiles) are impacted.

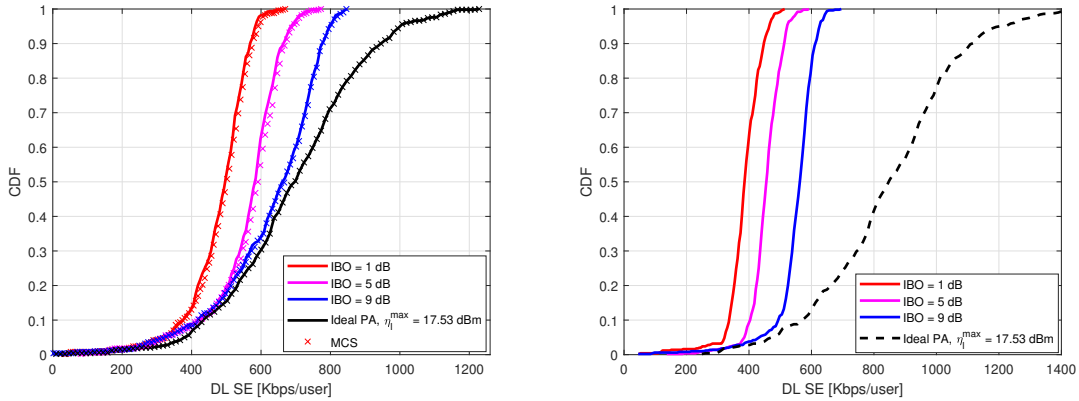

Figure 13: CDFs of DL per-RB SE achieved by *local* FZF and RZF precoding schemes: $L = 200$, $M = 16$, $K = 20$, $\tau_p = 15$, power control is based on (48). Solid curves denote analytical results obtained in closed-form (equation (56)) while markers indicate the results obtained by MC Simulations.

We also observe the following findings when comparing the two precoding schemes: in linear case (Ideal PA), we see that that the RZF performs better than FZF, for low and high percentiles. The reason is that RZF can support the interference better by taking into account channel estimation error and pilot contamination, thanks to the regularization matrix. Moreover, it is clear that RZF is more sensitive than FZF against the nonlinear distortions, especially at high SE percentiles. This is related to the fact that the RZF boosts power towards UEs with good channel conditions, leading to high NLD power towards this UEs. Moreover, FZF still has worse performance at low SE percentiles because it uses all the DoFs to suppress interference towards all the available orthogonal directions, resulting in a small array gain $M - \tau_p$. Besides, it would be better to cancel only interference towards UEs ($< \tau_p$) with good channel gains and increase the array gain $M - \tau_p$.

It shall be noted that RZF can be improved by taking into account the NLD power in computing the regularization matrices $\{\mathbf{P}_l\}$, which is out of scope of this paper. The reason is because we aim to show the relevance of our proposed PAPR reduction algorithm in improving the performance of this traditional RZF.

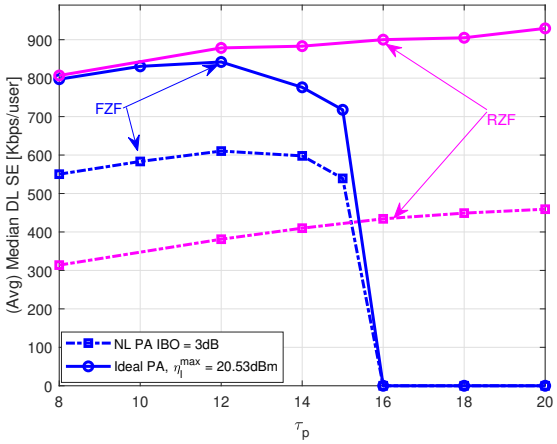

Figure 14: Median DL per-RB SE, achieved by FZF and RZF precoders, averaged over several large-scale fading realizations, $L = 200$, $M = 16$, $K = 20$, IBO = 3dB. Only results obtained by MC simulations are provided.

Effect of the number of pilots (τ_p) The implementation versatility of RZF versus the limitation of FZF is shown in Fig. 14. When M is fixed, RZF precoding can be performed by any number of pilots τ_p, while FZF constraints the number of orthogonal pilots, such that $\tau_p < M$ has to be verified (else the FZF inverse peudo-matrix is not defined).

For instance, Fig. 14 shows the ability of RZF to achieve much higher SEs in the operation regime in which FZF cannot be implemented. Moreover, from these results, we observe a loss in the performance of the FZF when τ_p goes from 12 to 15. In linear case, it gets lower from 842 to 718 kbps/user. Such degradation is caused by decreasing the array gain $M - \tau_p$ when τ_p tends to M. Therefore, it is not necessary to cancel interference towards all the available orthogonal directions for high array gain purpose. In nonlinear case, FZF performs better than RZF in its implementable operation regime, else it does not. For $\tau_p = 12$, FZF offers a SE gain of 57% for an IBO of 3dB.

5 Conclusions

The deliverable explores various key challenges and solutions related to CF-mMIMO systems, particularly focusing on aspects of channel state information (CSI), energy efficiency (EE), and hardware impairments (HWI) across multiple chapters.

The first chapter analyzes the impact of imperfect CSI on the performance of CF-mMIMO systems, particularly in downlink data transmission. The study reveals that channel estimation errors (CEE), non-linear power amplifier distortions, and limited feedback mechanisms degrade spectral efficiency (SE). Simulations show that local precoding (LP-MMSE) is more robust against CSI imperfections compared to centralized precoding (P-MMSE), though it has slightly lower SE under perfect conditions. As channel estimation errors increase, both precoding schemes suffer, but P-MMSE is more sensitive to these errors than LP-MMSE. In this chapter, we examined also the impact of limited feedback on system performance, particularly with quantized feedback bits. The results show that increasing the number of feedback bits improves SE, but the performance gain diminishes beyond a certain number of bits. Optimized vector quantization (VQ) outperforms random vector quantization (RVQ) in terms of SE, especially when the feedback bit budget is constrained. The findings emphasize the importance of designing efficient feedback mechanisms to balance overhead and performance in CF-mMIMO systems.

The second chapter reviews state-of-the-art approaches to CSI compression in Massive MIMO systems. Efficient CSI compression is crucial for reducing feedback overhead and computational complexity, which are significant challenges in large-scale antenna systems. The chapter examines various techniques and notes that while many compression methods have been developed for traditional Massive MIMO systems, adapting these methods to cell-free architectures presents a major challenge. The findings emphasize the need to balance compression efficiency while maintaining high performance levels, particularly as CF-mMIMO systems evolve. Future work will likely focus on refining these techniques to better address the distributed nature of cell-free networks.

In the last chapter, an energy-efficient downlink transmission scheme for CF-mMIMO-OFDM based system has been proposed and studied. We have derived the spectral efficiency of downlink CF-mMIMO-OFDM under hardware impairment considering local full-pilot zero-forcing and local regularized zero-forcing. As preliminary results, we have shown that in the case of ideal power amplifier, the local regularized zero-forcing performs better than the local full-pilot zero-forcing. In presence of power amplifier with non-linearity,local RZF is more sensitive than local FZF against the nonlinear distortions at high spectral efficiency percentiles. Regarding the effect of the number of pilots, RZF precoding can be performed using any number of pilots while FZF requires that the pilots remain orthogonal.

References

[1] Further Advancements for E-UTRA Physical Layer Aspects (Release 9). *3GPP, document TS 36.814, Mar. 2017*, 2017.

[2] Emil Björnson and Luca Sanguinetti. Scalable cell-free massive mimo systems. *IEEE Transactions on Communications*, 68(7):4247–4261, 2020.

[3] Hanen Bouhadda, Hmaied Shaiek, Daniel Roviras, Rafik Zayani, Yahia Medjahdi, and Ridha Bouallegue. Theoretical analysis of ber performance of nonlinearly amplified fbmc/oqam and ofdm signals. *EURASIP Journal on Advances in Signal Processing*, 2014(1):1–16, 2014.

[4] Julian J. Bussgang. Crosscorrelation functions of amplitude-distorted gaussian signals. 1952.

[5] Snehith T. C., Anil K. K., Raju A. K., and Ramanathan R. Impact of channel estimation errors on lattice reduction gains in mimo systems. In *2015 IEEE International Advance Computing Conference (IACC)*, pages 594–597, 2015.

[6] Kuo-Liang Chung and Wen-Ming Yan. The complex Householder transform. *IEEE Transactions on Signal Processing*, 45(9):2374–2376, 1997.

[7] Maha C. Dakhli, Rafik Zayani, Oussema B. Belkacem, and Ridha Bouallegue. Theoretical analysis and compensation for the joint effects of HPA nonlinearity and RF crosstalk in VBLAST MIMO-OFDM systems over Rayleigh fading channel. *Eurasip Journal on Wireless Communications and Networking*, 2014(61), 214.

[8] D. Dardari, V. Tralli, and A. Vaccari. A theoretical characterization of nonlinear distortion effects in ofdm systems. *IEEE Transactions on Communications*, 48(10):1755–1764, 2000.

[9] Jiajia Guo, Chao-Kai Wen, Shi Jin, and Geoffrey Ye Li. Convolutional neural network-based multiple-rate compressive sensing for massive mimo csi feedback: Design, simulation, and analysis. *IEEE Transactions on Wireless Communications*, 19(4):2827–2840, 2020.

[10] Khaled S. Hassan, Martin Kurras, and Lars Thiele. Performance of distributed compressive sensing channel feedback in multi-user massive mimo. In *2015 IEEE 11th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pages 430–436, 2015.

[11] R. W. Heath, T. Wu, and A. C. Soong. Progressive refinement of beamforming vectors for high-resolution limited feedback. *EURASIP Journal on Advances in Signal Processing*, pages 1–13, 2009.

[12] Xiaoling Hu, Caijun Zhong, Xiaoming Chen, Weiqiang Xu, Hai Lin, and Zhaoyang Zhang. Cell-free massive mimo systems with low resolution adcs. *IEEE Transactions on Communications*, 67(10):6844–6857, 2019.

[13] Jinliang Huang and Svante Signell. Impact of channel estimation error on performance of adaptive mimo systems. In *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 2865–2868, 2008.

[14] Giovanni Interdonato, Marcus Karlsson, Emil Bjornson, and Erik G. Larsson. Downlink Spectral Efficiency of Cell-Free Massive MIMO with Full-Pilot Zero-Forcing. In *2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 2018.

[15] Giovanni Interdonato, Marcus Karlsson, Emil Björnson, and Erik G. Larsson. Local Partial Zero-Forcing Precoding for Cell-Free Massive MIMO. *IEEE Transactions on Wireless Communications*, 19(7):4758–4774, 2020.

[16] Wei Jiang and Hans Dieter Schotten. Cell-Free Massive MIMO-OFDM Transmission Over Frequency-Selective Fading Channels. *IEEE Communications Letters*, 25(8):2718–2722, 2021.

[17] Zheng Jiang, Bin Han, Liang Lin, Peng Chen, Fengyi Yang, and Qi Bi. On compressive csi feedback beamforming scheme for fdd massive mimo. In *2016 IEEE/CIC International Conference on Communications in China (ICCC Workshops)*, pages 1–5, 2016.

[18] Steven M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, Inc., USA, 1993.

[19] Mounira Laabidi, Rafik Zayani, and Ridha Bouallegue. A new tone reservation scheme for papr reduction in fbmc/oqam systems. In *2015 International Wireless Communications and Mobile Computing Conference (IWCMC)*, pages 862–867, 2015.

[20] Mounira Laabidi, Rafik Zayani, and Ridha Bouallegue. A novel multi-block selective mapping scheme for papr reduction in fbmc/oqam systems. In *2015 World Congress on Information Technology and Computer Applications (WCITCA)*, pages 1–5, 2015.

[21] Jialing Li, Qi Zhang, Xiangjun Xin, Ying Tao, Qinghua Tian, Feng Tian, Dong Chen, Yufei Shen, Guixing Cao, Zihe Gao, and Jinxi Qian. Deep learning-based massive mimo csi feedback. In *2019 18th International Conference on Optical Communications and Networks (ICOCN)*, pages 1–3, 2019.

[22] Yves Louet, Daniel Roviras, Amor Nafkha, Hmaied Shaiek, and Rafik Zayani. Global power amplifier efficiency evaluation with papr reduction method for post-ofdm waveforms. In *2018 15th International Symposium on Wireless Communication Systems (ISWCS)*, pages 1–5, 2018.

[23] D.J. Love, R.W. Heath, and T. Strohmer. Grassmannian beamforming for multiple-input multiple-output wireless systems. volume 4, pages 2618–2622 vol.4, 2003.

[24] Chao Lu, Wei Xu, Hong Shen, Jun Zhu, and Kezhi Wang. Mimo channel information feedback using deep recurrent network. *IEEE Communications Letters*, 23(1):188–191, 2019.

[25] K.K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang. On beamforming with finite rate feedback in multiple-antenna systems. *IEEE Transactions on Information Theory*, 49(10):2562–2579, 2003.

[26] Aradhana Narula, Michael J Lopez, Mitchell D Trott, and Gregory W Wornell. Efficient use of side information in multiple-antenna data transmission over fading channels. *IEEE Journal on selected areas in communications*, 16(8):1423–1436, 1998.

[27] Matteo Nerini, Valentina Rizzello, Michael Joham, Wolfgang Utschick, and Bruno Clerckx. Machine learning-based csi feedback with variable length in fdd massive mimo. *IEEE Transactions on Wireless Communications*, 22(5):2886–2900, 2023.

[28] Nokia. Realistic Power Amplifier Model for the New Radio Evaluation. *document R4-163314, 3GPP TSG-RAN WG4 Meeting 79*, 2016.

[29] Berna Özbek and Didier Le Ruyet. *Feedback strategies for wireless communication*. Springer, 2014.

[30] Anastasios Papazafeiropoulos, Emil Björnson, Pandelis Kourtessis, Symeon Chatzinotas, and John M. Senior. Scalable cell-free massive mimo systems: Impact of hardware impairments. *IEEE Transactions on Vehicular Technology*, 70(10):9701–9715, 2021.

[31] Jaeyong Park, Jaewon Kim, Hyung-gil Yoo, and Wonjin Sung. Construction of phase tracking codebooks based on the lloyd-max vector quantization. In *2011 IEEE Wireless Communications and Networking Conference*, pages 1983–1987, 2011.

[32] R. Price. A useful theorem for nonlinear devices having Gaussian inputs. *IRE Transactions on Information Theory*, 4(2):69–72, 1958.

[33] R1-166004(R4-164542). *3GPP TSG-RAN WG1 Meeting #85 R1-166004*, volume May. 2016.

[34] Hmaied Shaiek, Rafik Zayani, Yahia Medjahdi, and Daniel Roviras. Analytical analysis of ser for beyond 5g post-ofdm waveforms in presence of high power amplifiers. *IEEE Access*, 7:29441–29452, 2019.

[35] Jinghan Shen, Xin Liang, Xinyu Gu, and Lin Zhang. Clustering algorithm-based quantization method for massive mimo csi feedback. *IEEE Wireless Communications Letters*, 11(10):2155–2159, 2022.

[36] Joao Vieira, Fredrik Rusek, Ove Edfors, Steffen Malkowsky, Liang Liu, and Fredrik Tufvesson. Reciprocity Calibration for Massive MIMO: Proposal, Modeling, and Validation. *IEEE Transactions on Wireless Communications*, 16(5), 2017.

[37] Chao-Kai Wen, Wei-Ting Shih, and Shi Jin. Deep learning-based CSI feedback approach for time-varying massive MIMO channels. *IEEE Wireless Communications Letters*, 7(5):748–751, 2018.

[38] Xudong Yin, Jianxin Dai, and Jianfeng Shi. Performance Analysis of Cell-Free Massive MIMO Systems Under Limited Feedback. In *2019 IEEE International Conference on Communications Workshops (ICC Workshops)*, pages 1–6, 2019.

[39] Rafik Zayani, Hmaied Shaiek, Xinying Cheng, Xiaotian Fu, Christophe Alexandre, and Daniel Roviras. Experimental testbed of post-ofdm waveforms toward future wireless networks. *IEEE Access*, 6:67665–67680, 2018.

[40] Binggui Zhou, Shaodan Ma, and Guanghua Yang. Transformer-based csi feedback with hybrid learnable non-uniform quantization for massive mimo systems. In *2023 32nd Wireless and Optical Communications Conference (WOCC)*, pages 1–5, 2023.