IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 8 September 2023, accepted 21 September 2023, date of publication 25 September 2023,
date of current version 29 September 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3318882

RESEARCH ARTICLE

Toward Energy-Efficient 6G Networks: Uplink Cell-Free Massive MIMO With NLD Cancellation Technique of Hardware Impairments

ASMA MABROUK AND RAFIK ZAYANI
CEA-Leti, Université Grenoble Alpes, 38054 Grenoble, France

Corresponding author: Asma Mabrouk (asma.mabrouk@cea.fr)

This work was supported in part by the Agence nationale de la recherche (ANR) under the France 2030 Program under Grant NF-PERSEUS: ANR-22-PEFT-0004, and in part by the Framework of the ALEX6 Project through the French CARNOT-ANR under Carnot-SCIENCE-ALEX6.

ABSTRACT This paper focuses on investigating uplink (UL) data transmission in cell-free massive MIMO based on orthogonal frequency division multiplexing (CF-mMIMO-OFDM) systems, taking into account the effects of hardware impairments (HWIs). Specifically, the HWIs arise from nonlinear distortions (NLD) caused by power amplifiers (PAs) at user equipments (UEs). These NLD have a significant impact on both channel estimation and data transmission in UL CF-mMIMO-OFDM. To mitigate NLD while maintaining a good power-efficiency, we propose a successive NLD cancellation approach that is adequate for CF-mMIMO-OFDM. Specifically, a novel frequency-domain channel estimation method is introduced that incorporates NLD cancellation. This method aims to accurately estimate the channel despite the presence of hardware impairments. Additionally, a joint multi-user (MU) combining and NLD cancellation scheme is proposed to mitigate the NLD impact on data detection. Note that three combining schemes are adopted, namely maximum-ratio (MR), zero-forcing (FZF), and partial-FZF (PFZF). Most-importantly, the proposed techniques are designed to be implemented in a distributed and scalable manner, highlighting the advantages of CF-mMIMO-OFDM systems. The performance of the proposed techniques are evaluated with simulations when considering the combining schemes. Results show the capability of our proposed NLD cancellation approach to improve both channel estimation and data detection, especially when leveraging the good features of PFZF combing scheme. For objective comparison purpose, we derived closed-form expressions on UL spectral-efficiency (SE) performance of an UL CF-mMIMO-OFDM system in presence of ideal and nonlinear PA.

INDEX TERMS Cell-free massive MIMO, hardware impairments, spectral efficiency, OFDM, distributed cancellation.

I. INTRODUCTION

Cell-Free massive MIMO (CF-mMIMO) has been shown to be the most promising key technology in the development of beyond 5G networks that are expected to bring new requirements for wireless communication systems, such as higher data rates, ultra-low latency, and massive connectivity [1], [2], [3]. To this end, a new network architecture is needed that can provide a scalable and flexible infrastructure able to support a wide range of applications and services. Purposely, CF-mMIMO systems do not rely on a traditional cellular network architecture, with the system divided into cells [4]. Instead, a large number of distributed access points (APs) connected to a central processing unit (CPU) are used to provide seamless wireless coverage and capacity across the entire coverage area. Each AP serves all user equipments (UEs) in the same time-frequency resource block. As such, UEs can potentially have access to the same amount of resources regardless of their location. With a high density of APs and UEs, the amount of signaling and data exchanged between nodes increases leading to a substantial increase in computational complexity and resource requirements.

The associate editor coordinating the review of this manuscript and approving it for publication was Stefan Schwarz.

Therefore, scalability-related issues should be addressed in CF-mMIMO systems. In particular, user-centric approach enables efficient and dynamic AP selection by UEs to optimize resource allocation and alleviate the challenges posed by a large number of APs and UEs [2], [5]. Based on time-division duplex (TDD) transmission mode, the transmissions in CF-mMIMO are carried out in coherence blocks, each divided into three phases : uplink (UL) training, UL payload transmission and downlink (DL) payload transmission. Obtaining accurate channel estimates is necessary to achieve the full potential of UL/DL transmission in CF-mMIMO by optimizing the design of the combining/precoding schemes at serving APs. Combining schemes are designed to enhance the UL performance efficiency by reducing the inter-user interference and/or enhancing the desired signal of each user. Two levels of cooperation between APs and the CPU are proposed for UL combining in CF-mMIMO: centralized combining and distributed combining [6]. It is shown that, by leveraging distributed combining, CF-mMIMO systems can achieve better scalability [7].

While CF-mMIMO presents promising potential for enhancing wireless communications, its implementation in real-world scenarios poses practical challenges [3]. Notably, the densification feature of CF-mMIMO can indeed result in higher system efficiency at the cost of a substantial increase in energy consumption and hardware costs. This is especially true if both the APs and UEs are equipped with high-quality hardware components. Nevertheless, while low-cost hardware may be attractive due to its affordability, it often comes with the downside of reduced system efficiency. Hence, it is essential to achieve a balance between the cost of the hardware and the system's performance [8]. The aforementioned works rely on simplifying assumptions such as ideal hardware, which is unlikely to hold in practical implementation. In this context, power amplifier (PA) non-linearity (NL) is considered one of the major hardware impairments (HWIs) in the analog transmission chain of modern communication systems. Indeed, non-linear distortion (NLD) of non-linear power amplifier (NL-PA) deteriorates both channel estimates and the useful received data.

A. RELATED WORKS

The impact of potential HWIs on the performance of CF-mMIMO systems has been explored in [9], [10], [11], [12], [13], [14], [15], and [16]. In [9], authors showed that HWIs can significantly degrade the performance of CF-mMIMO systems in terms of spectral efficiency (SE) and energy efficiency (EE). According to the authors, although increasing the number of APs can alleviate the impact of HWIs at APs, the system performance still suffer from HWIs at UEs. The same hardware scaling law was reveled in [10] where closed-form of SE and EE expressions were derived for both UL and DL transmissions. It is also shown that max-min power control algorithm improves the performance of CF-mMIMO systems under HWIs at both APs and UEs. The impact of HWIs on UL CF-mMIMO systems with four low-complexity receiver cooperation levels has been investigated in [11]. The obtained results show that HWIs have a larger impact on the performance of large scale fading decoding (LSFD) combining receiver. In [12], authors studied the effect of HWIs on the UL and DL sum-rate of CF-mMIMO-OFDM systems and revealed that the UL sum-rate is more susceptible to HWIs at UEs than the DL sum-rate. The impact of HWIs, including phase noise, quantization errors, and PA distortion, on the physical layer security of CF-mMIMO systems in the presence of pilot spoofing attack has been presented in [13]. It is shown that the hardware scaling law is almost not applicable to secure CF-mMIMO systems under active attack. The impacts of both phase drifts and distortion noise on CF-mMIMO systems have been investigated in [14]. The achievable rate of CF-mMIMO systems with low resolution analog-to-digital converters at both the APs and users has been investigated in [15]. In [16], authors highlighted the performance of fronthaul-constrained CF-mMIMO under transceiver HWIs.

B. CONTRIBUTIONS

Existing studies have predominantly focused on examining the impact of HWIs at APs and/or UEs on the performance of CF-mMIMO systems, rather than addressing strategies to mitigate this effect. Building upon the findings presented in [10], which indicate that the UL SE is primarily affected by HWIs at the UEs, we investigate the performance of UL OFDM-based CF-mMIMO systems under NL-PA model at UEs. Contrary to the work in [10] and [11], which considered theoretical NLD model, our work specifically considers a representative NL-PA model. It is worth noting that NLD could be compensated either at the UE side or the AP side. The former case may require additional computational complexity, potentially affecting battery life and overall UE performance. This can motivate a receiver cancellation technique which is typically the case in UL transmission where the APs have more resources in terms of power and computational complexity. Therefore, the main contributions of this work are as follows

- We propose a frequency-domain channel estimation approach that incorporates a NLD cancellation technique to improve the efficiency and accuracy of channel estimation by mitigating the impact of nonlinear distortion.
- A new receiver technique, that is able to cancel the NLD caused by NL-PAs, is developed. The proposed technique combines UL detection and NLD cancellation techniques, called successive joint MU combing and NLD cancellation scheme. When applied in a CF-mMIMO-OFDM system, the NLD cancellation technique is evaluated in a representative scenario, which includes channel estimation error and multi-user interference. Simulation results are provided to demonstrate that the proposed techniques improve the

TABLE 1. Notations.

Notation	Definition
$(.)^H$	Hermitian transpose
$\mathbb{E}\{.\}$	Expectation value
$\|.\|$	Euclidean norm
$\|\mathcal{A}\|$	Cardinality of the set $\mathcal{A}$
$\mathbf{I}_N$	$N \times N$ identity matrix
$\text{diag}(x_1, \ldots, x_n)$	Diagonal matrix with the values $x_1, \ldots, x_n$ on the diagonal
$\mathcal{N}_{\mathbb{C}}(\mathbf{0}, \sigma \mathbf{I}_N)$	Zero-mean complex Gaussian random vector with variance σ
	Boldface lower-case and upper-case letters denote vectors and matrices, respectively.

overall performance of the CF-mMIMO-OFDM system by reducing the impact of HWIs.

- We analyze the global SE of a CF-mMIMO-OFDM system when the proposed approaches are implemented. Unlike [10] and [11], instead of considering the MR combining scheme, we investigate the system performance under HWIs based on different combining techniques: MR, full-pilot zero-forcing (FZF) and partial-FZF (PFZF). The achievable SE of the system is derived for the three combining schemes. We show the ability of our NLD cancellation approach especially when leveraging the good features of PFZF combining scheme that further improve the spectral efficiency by providing an adaptable trade-off between NLD cancellation and boosting of the desired signal.
- In addition, we discuss the computational complexity analysis for the proposed schemes. Specifically, we show that the technique converges within a few iterations. Significantly, the added complexity is lower than conventional UL data transmission that includes combining vector design and initial data detection, especially when FZF and PFZF combining schemes are used.
- For objective comparison purpose, we derived closed-form expressions on UL spectral-efficiency (SE) performance of an UL CF-mMIMO-OFDM system in the presence of ideal and nonlinear PA. Unlike [10] and [11], that consider theoretical HWI model, we consider a representative PA model proposed by the 3GPP.

The remainder of this paper is summarized in the following order. Section II describes the considered hardware-constrained CF-mMIMO transmission model and Section III presents the system performance where closed-form of the achievable SE expressions with different combining schemes are derived. The proposed techniques are detailed in Section IV. Section V exhibits simulation results and discussions. At last, Section VI concludes this paper. Notations used in the paper are outlined in Table. 1.

II. SYSTEM MODEL

We consider a UL transmission in a TDD-based CF-mMIMO-OFDM system consisting of L APs and K UEs, which are independently and uniformly distributed within a given area. UEs and APs are equipped with a single transmit antenna and N receive antennas, respectively. We assume that APs are connected to a CPU via a high-capacity and error-free fronthaul network. Operating in TDD mode, the training phase, UL and DL transmissions fit into the channel coherence block of length τ_c, as illustrated in FIGURE 1. We assume a block-fading channel model where channel coefficients remains unchanged during the coherence block. Let $\mathbf{H}_{k,l} = [\mathbf{h}^0_{k,l}, \mathbf{h}^1_{k,l}, \ldots, \mathbf{h}^I_{k,l}] \in \mathbb{C}^{N \times I}$ be the time-domain channel response between AP_l and UE_k that is modeled as a finite impulse response filter with I equally-spaced channel taps. The i-th channel tap is given by

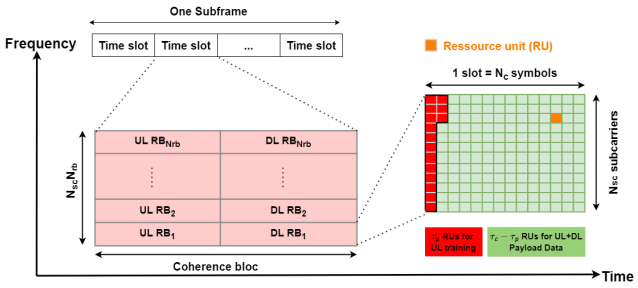


FIGURE 1. OFDM TDD frame structure.

$$\mathbf{h}^i_{k,l} = \sqrt{\beta_{k,l}}\mathbf{g}^i_{k,l} \tag{1}$$

where $\beta_{k,l}$ and $\mathbf{g}^i_{k,l}$ are the large and small scale fading components, respectively. It is assumed that $\mathbf{g}^i_{k,l} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I}_N)$ are independent and identically distributed (i.i.d.). CF-mMIMO architecture aims to offer fully macro-diversity. Hence, UEs should be positioned at varying distances from different APs, including some close, some far, and some very far. Consequently, Rayleigh fading channel might be the appropriate model for characterizing such a scenario. Moreover, we have adopted the Rayleigh channel model for comparison purposes, as it is often used in prior works on CF-mMIMO systems involving hardware impairments [10], [11], [17]. We use OFDM modulation to divide the frequency band into multiple subcarriers that are used by UEs in an orthogonal way to transmit data simultaneously. In the frequency-domain, the channel response on the m-th subcarrier between AP_l and UE_k is denoted by $\mathbf{h}_{k,l,m}, m = 1, \ldots, N_{sub}$ where N_{sub} is the number of OFDM subcarriers.

Following 5G-NR, based on TDD mode, the UL and DL transmissions share the same frequency band but are separated in time. A frame is typically represented as a grid in the time-frequency plane, where the frequency domain is divided into N_{sub} subcarriers spaced by Δ_f and the time domain is divided into N_c OFDM symbols. Moreover, the data transmission is organized into N_{rb} resource blocks (RB), each consisting of N_{sc} contiguous subcarriers and N_c symbols. Therefore, each RB comprises $\tau_c = N_c N_{sc}$ resource units (RUs), representing the smallest time-frequency resource of one subcarrier and one symbol. To ensure accurate channel estimation, pilot symbols are inserted into RUs of each RB. We assume that the number of OFDM symbols, N_c, covers the coherence time of all UEs. In each RB, τ_p RUs are used

for pilot symbols, where the remaining RUs are used for scheduling UL and DL transmissions. Specifically, $\xi(\tau_c-\tau_p)$ and $(1-\xi)(\tau_c-\tau_p)$ RUs are used for scheduling UL and DL transmissions, respectively, where $0<\xi<1$ [18].

A. HARDWARE IMPAIRMENT MODEL

In wireless communication systems, HWIs can arise from various components. However, it is worth noting that NL-PAs are particularly prominent and significant sources of impairments. The non-linearity encountered in a PA can be described through two types of conversion: amplitude to amplitude (AM/AM) and amplitude to phase (AM/PM). These conversions correspond to the changes in both the amplitude and phase of an input signal, r_{in}, passing through the PA resulting in the signal, r_{out}, at the output. The relation between the input, $r_{in}=\rho e^{i\psi}$, and output, r_{out}, signals of a PA is described by its transfer function, $f(.)$, which is given by

$$r_{out}=f(r_{in})=\Omega(\alpha\rho)\exp^{j(\psi+\Psi(\alpha\rho))} \tag{2}$$

where $\Omega(.)$ and $\Psi(.)$ denote the AM/AM and AM/PM conversions, respectively, and α is the multiplicative coefficient that is applied to the input of PA to achieve the desired PA operating point based on the specified input back-off (IBO). The IBO refers to the amount of power reduction at the input of the PA relative to its maximum rated power level. Moreover, ρ and ψ are the magnitude and the phase of the PA's input sample. To accurately model the behavior of NL-PAs, we use the memoryless modified Rapp model based on which the AM/AM and AM/PM conversions are given by [19]

$$\Omega(\rho)=\frac{G\rho}{\left(1+\left|\frac{G\rho}{V_{sat}}\right|^{2p}\right)^{\frac{1}{2p}}},\qquad \Psi(\rho)=\frac{A\rho^q}{\left(1+\left(\frac{\rho}{B}\right)^q\right)} \tag{3}$$

where G is the small signal gain, V_{sat} is the saturation level, p is the smoothness factor and A, B and q are fitting parameters. Note that $\alpha=\frac{V_{sat}}{G\sqrt{p_{in}}}10^{\frac{-IBO[dB]}{10}}$, where p_{in} is the signal average power. Based on the Bussgang's theorem [20], the output of a NL-PA can be expressed in the following form

$$r_{out}=\kappa_0 r_{in}+d \tag{4}$$

where κ_0 denotes a complex gain and d is the added zero-mean distortion noise with variance σ_d^2. Note that, d is uncorrelated with the input signal, r_{in}. It is worth noticing that the Bussgang theorem holds only for Gaussian input signals. This condition is verified in the proposed scheme, since an OFDM modulated signal is amplified at each UE. In addition, it is mentioning that d is not Gaussian at the output of the PA but it becomes Gaussian at the receiver side (AP) after the OFDM demodulation.

B. UPLINK TRAINING PHASE

In wireless CF-mMIMO-OFDM systems, pilot-based channel estimation is a common technique used to estimate the users' channels at the APs. For simplicity, we assume that the channel coefficients are the same for all RUs within a RB. Therefore, channel estimation can be performed once per RB. During the UL training phase, UEs transmit τ_p pilot symbols on specific RUs within the RB. These RUs are carefully chosen to be orthogonal to each other in the frequency-domain in order to enable simultaneous transmission of pilots without causing interference with other UEs. This approach helps to mitigate pilot contamination and ensure accurate channel estimation. It is also important to highlight that by allocating different RBs to different UEs, OFDM technique can serve more UEs without interference in the frequency domain.

Let $\phi_{k,m}\in\mathbb{C}^{\tau_p\times 1}$ denotes the τ_p-length pilot sequence assigned to the k-th user, where $\phi_{k,m}\phi_{k,m}^H=\tau_p$ and $\phi_{k,m}\phi_{k',m}^H=0,\ \forall k\neq k'$. Considering NL-PA at UEs, the received signal, $\mathbf{Y}_{l,m}^p\in\mathbb{C}^{N\times\tau_p}$, at the l-th AP is given by

$$\mathbf{Y}_{l,m}^p=\mathbf{H}_{l,m}\left(\mathbf{P}\Phi_m^H+\mathbf{D}_m^p\right)+\mathbf{N}_{l,m}, \tag{5}$$

where $\mathbf{P}=\text{diag}(\kappa_0\sqrt{p_1},\ldots,\kappa_0\sqrt{p_K})$ is the transmit power matrix, $\mathbf{N}_{l,m}\in\mathbb{C}^{N\times\tau_p}$ is the AWGN matrix with variance σ^2 and $\mathbf{D}_m^p=[\mathbf{d}_{1,m}^p,\ldots,\mathbf{d}_{K,m}^p]$ is the matrix that contains the NLD vector of each user on the m-th subcarrier. This first step is to estimate the channels $\mathbf{H}_{l,m}$ is performing a de-spreading of the received pilot signal as follows [21]

$$\overline{\mathbf{H}}_{l,m}=\frac{1}{\sqrt{\tau_p}}\mathbf{Y}_{l,m}^p\Phi_m\in\mathbb{C}^{N\times\tau_p}, \tag{6}$$

where $\Phi_m=[\phi_{1,m},\ldots,\phi_{\tau_p,m}]\in\mathbb{C}^{\tau_p\times\tau_p}$. Then, based on the estimation theory in [22], by adopting the minimum mean square error (MMSE) estimation of the channel response, $\hat{\mathbf{h}}_{k,l,m}$ can be expressed as

$$\hat{\mathbf{h}}_{k,l,m}=c_{k,l}\overline{\mathbf{H}}_{l,m}\mathbf{e}_k, \tag{7}$$

where

$$c_{k,l}=\frac{\sqrt{p_k\tau_p}\kappa_0\beta_{k,l}}{p_k\tau_p|\kappa_0|^2\beta_{k,l}+\sum_{t=1}^K\sigma_d^2\beta_{t,l}+\sigma^2}, \tag{8}$$

represents the full-rank matrix of the channel estimates and $\mathbf{e}_k$ denotes the k-th column of $\mathbf{I}_{\tau_p}$. The channel estimate, $\hat{\mathbf{h}}_{k,l,m}$, and the channel estimation error, $\tilde{\mathbf{h}}_{k,l,m}=\mathbf{h}_{k,l,m}-\hat{\mathbf{h}}_{k,l,m}$, are independent with distributions $\mathcal{N}_{\mathbb{C}}\left(\mathbf{0},\gamma_{k,l}\mathbf{I}_N\right)$ and $\mathcal{N}_{\mathbb{C}}\left(\mathbf{0},(\beta_{k,l}-\gamma_{k,l})\mathbf{I}_N\right)$, respectively, where

$$\begin{aligned}\gamma_{k,l}&=\sqrt{p_k\tau_p}\kappa_0^*\beta_{k,l}c_{k,l},\\&=\frac{p_k\tau_p|\kappa_0|^2\beta_{k,l}^2}{p_k\tau_p|\kappa_0|^2\beta_{k,l}+\sum_{t=1}^K\sigma_d^2\beta_{t,l}+\sigma^2}.\end{aligned} \tag{9}$$

We note form (9) that the estimated channel of each UE is affected by NLD caused by both itself and other users in the network. Moreover, the level of NLD caused by other UEs varies based on their channel conditions.

III. PERFORMANCE ANALYSIS

In this section, we evaluate the achievable SE of distributed UL transmission in CF-mMIMO-OFDM system that consists of two stages: local processing and LSFD [6]. First, each AP independently estimates the channels and designs its combiner vector to perform data detection locally. The estimated data at all APs are then sent to the CPU that performs linear processing for joint detection. We consider LSFD technique at the CPU which relies solely on channel statistics since the channel estimates are not shared over the fronthaul links. Moreover, we study the system performance based on the following receive combining schemes: MR, FZF and PFZF.

A. UPLINK DATA TRANSMISSION

During the UL data transmission, the received data signal at the l-th AP on the m-th subcarrier is given by

$$\mathbf{y}_{l,m}^{ul} = \sum_{k=1}^{K} \mathbf{h}_{k,l,m} \left(\sqrt{p_k} \kappa_0 s_{k,m} + d_{k,m} \right) + \mathbf{n}_{l,m}, \tag{10}$$

where $s_{k,m}$ is the transmitted data by the k-th user on the m-th subcarrier and $\mathbf{n}_{l,m} \in \mathbb{C}^{N \times 1}$ is the AWGN vector with variance σ^2. The local data estimate of the k-th UE at the l-th AP based on the combining vector $\mathbf{v}_{k,l,m}$ is given by

$$\hat{s}_{k,l,m} = \mathbf{v}_{k,l,m}^H \mathbf{y}_{l,m}^{ul}. \tag{11}$$

Local data estimates are then sent to the CPU to be linearly combined using the LSFD coefficients. The received signal at the CPU is given by

$$\hat{s}_{k,m} = \sum_{l=1}^{L} a_{k,l,m}^* \hat{s}_{k,l,m}, \tag{12}$$

where $a_{k,l,m}$ is the complex LSFD coefficient for AP_l and UE_k on the m-th subcarrier. As to maximize the SINR of UE_k, $\forall k \in \{1, \ldots, K\}$, the LSFD coefficients are computed as follows

$$\mathbf{a}_{k,m} = \left(\sum_{t=1}^{K} p_k \mathbb{E}\{\mathbf{g}_{k,t,m} \mathbf{g}_{k,t,m}^H\} + \sigma^2 \mathbf{F}_k \right)^{-1} \mathbb{E}\{\mathbf{g}_{k,k,m}\}, \tag{13}$$

where $\mathbf{g}_{k,t,m} = [\mathbf{v}_{k,1,m}^H \mathbf{h}_{t,1,m}, \ldots, \mathbf{v}_{k,L,m}^H \mathbf{h}_{t,L,m}]^T$ and $\mathbf{F}_k = \text{diag}\left(\{\|\mathbf{v}_{k,1,m}\|^2\}, \ldots, \{\|\mathbf{v}_{k,L,m}\|^2\} \right)$.

B. LOCAL RECEIVE COMBINERS

In this subsection, we present the combining schemes considered in this investigation : MR, FZF and PFZF.

1) MR combining is the simplest receiving scheme that maximizes the power of the desired signal while neglecting the inter-user interference. The MR combing vector is given by

$$\mathbf{v}_{k,l,m}^{\text{MR}} = c_{k,l} \overline{\mathbf{H}}_{l,m} \mathbf{e}_k = \hat{\mathbf{h}}_{k,l,m}. \tag{14}$$

Based on (14), we can note that this combining scheme has a low computational complexity since the combining vector is the corresponding local channel estimate and no additional computation is needed.

2) FZF is an efficient UL combining scheme since it enables the inter-user interference cancellation while maintaining strong desired signal powers. The FZF combiner vector is given by

$$\mathbf{v}_{k,l,m}^{\text{FZF}} = c_{k,l} \theta_{k,l} \overline{\mathbf{H}}_{l,m} \left(\overline{\mathbf{H}}_{l,m}^H \overline{\mathbf{H}}_{l,m} \right)^{-1} \mathbf{e}_k, \tag{15}$$

where $\theta_{k,l} = \mathbb{E}\{|\overline{\mathbf{H}}_{l,m} \mathbf{e}_k|^2\} = \frac{\gamma_{k,l}}{c_{k,l}^2}$. In this combining scheme, one combiner vector is proposed per pilot symbol. The FZF could be used only if $N \geq \tau_p$. Moreover, while FZF combining can eliminate interference, it can also lead to noise amplification.

3) PFZF is a modified version of the FZF combining scheme that alleviates the constraint on the number of antennas at the APs versus the number of pilots. In PFZF combining, each AP divides the UEs into two disjoint subsets : $\mathcal{S}_l$ of strong users and $\mathcal{W}_l$ of weak users given by

$$\mathcal{S}_l = [j_{l1}, \ldots, j_{l\nu_l}] = \underset{\downarrow k \in \{1,\ldots,K\}}{\text{argsort}} \{\beta_{k,l}\}, \tag{16}$$

$$\mathcal{W}_l = \{\beta_{k,l}\} \backslash \mathcal{S}_l. \tag{17}$$

We use the same strategy as in [23] to define the value of $\nu_l = |\mathcal{S}_l|$ denoting the number of UEs having the best channel conditions to AP_l and contributing to $\delta\%$ of the total path loss. Hence, the strong users is defined as follows

$$\sum_{t=1}^{\nu_l} \frac{\overline{\beta}_{t,l}}{\sum_{j=1}^{K} \beta_{k,l}} \geq \delta\%, \tag{18}$$

where $\{\overline{\beta}_{t,l}\}$ is the set of sorted path-losses (in descending order) and the threshold δ indicates the percentage of the total path loss from all users. The strong users are assigned to the FZF combiner, while the weak users are assigned to the MR combiner. Hence, the PFZF combining vector is given by

$$\mathbf{v}_{k,l,m}^{\text{PFZF}} = \begin{cases} \mathbf{v}_{k,l,m}^{\text{FZF}} & \text{if} \quad k \in \mathcal{S}_l \\ \mathbf{v}_{k,l,m}^{\text{MR}} & \text{if} \quad k \in \mathcal{W}_l \end{cases} \tag{19}$$

The choice of the combiner scheme depends on the specific system requirements, such as the number of APs, the channel conditions, and the desired performance metrics.

C. UPLINK SPECTRAL EFFICIENCY

The CPU in a CF-mMIMO system is responsible for coordinating the transmissions from all the APs, and optimizing the overall system performance. The signal in (12) can be rewritten as shown in (20), shown at the bottom of the next page.

The detected signal at the CPU comprises five main parts : $\text{DS}_{k,m}$ and $\text{BU}_{k,m}$ represent the desired signal and the beamforming uncertainty for the k-th user, respectively. Moreover, the user-interference caused by the t-th user

$(t \neq k)$ is denoted by $\mathrm{UI}_{k,t,m}$. The term $\mathrm{NLD}_{k,t,m}$ presents the received NLD form the t-th user in the network. Finally, $\mathrm{GN}_{k,m}$ represents the total received noise. Based on (20), the SINR of UE_k is given by (21), as shown at the bottom of the next page.

Note that, the SINR expressions in (21) holds for any combining scheme used at the AP. The closed-form expression of the SINR of UE_k is given by

$$\mathrm{SINR}_{k,m} = \frac{p_k|\kappa_0|^2 \left|\sum_{l\in\mathcal{A}_{k,m}} a^*_{k,l,m}\gamma_{k,l} + \sum_{l\in\mathcal{B}_{k,m}} N a^*_{k,l,m}\gamma_{k,l}\right|^2}{\mathrm{A}_{k,m} + \mathrm{B}_{k,m} + \mathrm{C}_{k,m} + \mathrm{D}_{k,m}}, \tag{22}$$

where

$$\mathrm{A}_{k,m} = \sum_{t=1}^{K} p_t|\kappa_0|^2 \left(\sum_{l\in\mathcal{A}_{k,m}} |a^*_{k,l,m}|^2 \frac{\gamma_{k,l}(\beta_{t,l} - \gamma_{t,l})}{(N-\nu_l)} + \sum_{l\in\mathcal{B}_{k,m}} |a^*_{k,l,m}|^2 N\gamma_{k,l}\beta_{t,l} \right),$$

$$\mathrm{B}_{k,m} = p_t|\kappa_0|^2 \left(\sum_{l\in\mathcal{A}_{k,m}} a^*_{k,l,m}\gamma_{t,l} + \sum_{l\in\mathcal{B}_{k,m}} a^*_{k,l,m} N\gamma_{t,l} \right),$$

$$\mathrm{C}_{k,m} = \sigma_d^2 \left(\sum_{l\in\mathcal{A}_k} a^*_{k,l,m}\gamma_{k,l} + \sum_{l\in\mathcal{B}_k} a^*_{k,l,m} N\gamma_{k,l} \right) + \sum_{t=1}^{K} \sigma_d^2 \left(\sum_{l\in\mathcal{A}_k} |a^*_{k,l,m}|^2 \frac{\gamma_{k,l}(\beta_{t,l} - \gamma_{t,l})}{(N-\nu_l)} + \sum_{l\in\mathcal{B}_k} |a^*_{k,l,m}|^2 N\gamma_{k,l}\beta_{t,l} \right),$$

$$\mathrm{D}_{k,m} = \sigma^2 \left(\sum_{l\in\mathcal{A}_k} |a^*_{k,l,m}|^2 \frac{\gamma_{k,l}}{(N-\nu_l)} + \sum_{l\in\mathcal{B}_k} |a^*_{k,l,m}|^2 N\gamma_{k,l} \right),$$

$$\mathcal{A}_k = \{l = 1, \ldots, L, k \in \mathcal{S}_l\},$$

$$\mathcal{B}_k = \{l = 1, \ldots, L, k \in \mathcal{W}_l\}.$$

Proof: The proof is given in the Appendix.

From the users' perspective, $\mathcal{A}_k$ denotes the set of APs considering UE_k as a strong user and $\mathcal{B}_k$ refers to the set of APs considering UE_k as a weak user. Therefore, $|\mathcal{A}_k| + |\mathcal{B}_k| = L$. We note the $\mathcal{A}_k = \emptyset$ when MR combining scheme is used. However, when the FZF receive combining is used, we have $\mathcal{B}_k = \emptyset$ and $\nu_l = \tau_p$.

Thus, the achievable per-RB SE of the k-th user for the UL CF-mMIMO-OFDM system is given by

$$\mathrm{SE}_k = \xi\,\Delta_f N_{sc} \left(1 - \frac{\tau_p}{\tau_c}\right) \log_2\left(1 + \mathrm{SINR}_{k,m}\right), \tag{23}$$

where Δ_f is the subcarrier spacing (SCS) and ξ denotes the portion of coherence time interval dedicated to the UL data transmission phase.

IV. DISTRIBUTED NLD CANCELLATION SCHEMES

NL-PA at UEs impacts both channel estimation and UL transmission. Hence, we propose two NLD cancellation algorithms, adequate for the scalable CF-mMIMO architecture, that is able to improve the accuracy of the channel estimation and reduce the NLD impact on the UL transmission phase, leading to improved performance of UL CF-mMIMO systems under HWIs. Fig. 2 shows the different components involved in the studied system. As depicted, the system comprises several UEs equipped by NL-PAs served by multiple APs. Moreover, the reception process at the APs is also shown. Each AP estimates the initial UL channels based on the pilot sent by UEs. Then, before designing the combiner vector, each AP applies a NLD cancellation technique to the initial channel estimates. This technique aims to improve the accuracy of channel estimates by reducing the effect of the NLD. After designing the combiner vector, the AP performs data detection on the received data symbols. However, since the data symbols are also influenced by NLD, a second NLD cancellation algorithm is subsequently employed to process the detected data before it is transmitted to the CPU. Details of the proposed NLD cancellation techniques are detailed in next sections. It is also shown that the NLD cancellation is be performed locally at the different APs in a distributed manner. Interestingly, performing NLD cancellation at the APs can simplify the overall system architecture, as both the UEs and the CPU do not need to perform any additional processing. This can reduce the overall complexity of the system and avoid to transfer data between the APs and the CPU.

The proposed successive NLD cancellation techniques tries to mitigate the effects of using NL-PAs by approximating

$$\hat{s}_{k,m} = \underbrace{\mathbb{E}\left\{\sum_{l=1}^{L} \sqrt{p_k}\kappa_0 a^*_{k,l,m}\mathbf{v}^H_{k,l,m}\mathbf{h}_{k,l,m}\right\} s_{k,m}}_{\mathrm{DS}_{k,m}} + \underbrace{\sqrt{p_k}\kappa_0 \left(\sum_{l=1}^{L} a^*_{k,l,m}\mathbf{v}^H_{k,l,m}\mathbf{h}_{k,l,m} - \mathbb{E}\left\{\sum_{l=1}^{L} a^*_{k,l,m}\mathbf{v}^H_{k,l,m}\mathbf{h}_{k,l,m}\right\}\right) s_{k,m}}_{\mathrm{BU}_{k,m}}$$
$$+ \sum_{t=1,t\neq k}^{K} \underbrace{\sum_{l=1}^{L} \sqrt{p_t}\kappa_0 a^*_{k,l,m}\mathbf{v}^H_{k,l,m}\mathbf{h}_{t,l,m}}_{\mathrm{UI}_{k,t,m}} s_{t,m} + \sum_{t=1}^{K} \underbrace{\sum_{l=1}^{L} a^*_{k,l,m}\mathbf{v}^H_{k,l,m}\mathbf{h}_{t,l,m} d_{t,m}}_{\mathrm{NLD}_{k,t,m}} + \underbrace{\sum_{l=1}^{L} a^*_{k,l,m}\mathbf{v}^H_{k,l,m}\mathbf{n}_{l,m}}_{\mathrm{GN}_{k,m}} \tag{20}$$

and subtracting the distortion signal from the received signals (pilot symbols and data symbols). The initial estimated distortion is generally inaccurate, and the iterative process is used to refine the remaining distortion. The proposed iterative cancellation algorithms are mainly based on the following main steps: initial symbol estimation, distortion approximation and distortion removal.

A. DISTRIBUTED CHANNEL ESTIMATION WITH NLD CANCELLATION

As the presence of NLD can cause errors in the estimation of the channel state information, we propose a distributed NLD cancellation technique where each AP approximates the distortion of each user separately. This is because the non-linear distortion introduced by each user's PA is different, and needs to be approximated and removed separately to mitigate its impact on their received signals. It is worth mentioning that channel estimation is performed in the frequency domain. The proposed method is applied at the AP side after the initial MMSE channel estimation and is depicted in **Algorithm 1**.

The main steps of the algorithm are summarized as follows

- Distortion approximation (step 3): We assume that the AP knows the PA model and the assigned pilot sequence to each user. Therefore, distortion approximation can be performed before initiating the iterative process of the algorithm. Then, the distortion introduced by the k-th user can be computed, based on the recovered time-domain pilot sequences, and transformed to the frequency domain as follows

$$\hat{\mathbf{D}}^p = \mathrm{FFT}(f(\mathrm{IFFT}(\Phi)) - \kappa_0 \mathrm{IFFT}(\Phi)) \tag{24}$$

  This approximated distortion will be used in the iterative process to remove the received distortion from each user which is influenced by the characteristics of the wireless channel.
- Initial channel estimates (step 5): The cancellation process is based on the initial channel estimates obtained from the received pilot signal in (6) and is given by

$$\begin{aligned}\hat{\mathbf{H}}^{(0)}_{l,m} &= \frac{\kappa_0^*}{|\kappa_0|^2\sqrt{\tau_p}}\overline{\mathbf{H}}_{l,m} = \frac{\kappa_0^*}{|\kappa_0|^2\tau_p}\mathbf{Y}^p_{l,m}\Phi_m, \\ &= \mathbf{H}_{l,m}\mathbf{P}' + \frac{\kappa_0^*}{|\kappa_0|^2\tau_p}\mathbf{H}_{l,m}\mathbf{D}^p_m\Phi_m \\ &\quad + \frac{\kappa_0^*}{|\kappa_0|^2\tau_p}\mathbf{N}_{l,m}\Phi_m,\end{aligned} \tag{25}$$

  where $\mathbf{P}' = \mathrm{diag}(\sqrt{p_1}, \ldots, \sqrt{p_K})$. More precisely, for a given user, UE_k, the first estimated channel at AP_l is given by

$$\begin{aligned}\hat{\mathbf{h}}^{(0)}_{k,l,m} &= \sqrt{p_k}\mathbf{h}_{k,l,m} + \underbrace{\frac{\kappa_0^*}{|\kappa_0|^2\tau_p}\mathbf{h}_{k,l,m}\mathbf{d}^p_{k,m}\phi_{k,m}}_{\text{received individual distortion}} \\ &\quad + \frac{\kappa_0^*}{|\kappa_0|^2\tau_p}\sum_{t=1,t\neq k}^{K}\mathbf{h}_{t,l,m}\mathbf{d}^p_{t,m}\phi_{k,m} \\ &\quad + \frac{\kappa_0^*\mathbf{N}_{l,m}\phi_{k,m}}{|\kappa_0|^2\tau_p}.\end{aligned} \tag{26}$$

  The received distortion is uncorrelated but is UE's channel-dependent, which means that the distortion removal process needs to take the channel information into account.
- Distortion removal (step 7): The AP then uses an iterative cancellation process to remove the received distortion from the received signal. In the i-th iteration, it approximates the total NLD term by multiplying the approximated distortion by the estimated channel in the (i-1)th iteration.
- Stop criterion (step 8): The iterative cancellation process is repeated until either the maximum number of iterations is reached, $i^* = N_{itr}$ or until no further cancellation is done between two successive iterations, $i^* < N_{itr}$.

Algorithm 1 NLD Cancellation in Channel Estimation at the l-Th AP

1: **Input:** $\mathbf{Y}^p_l$, Φ, N_{itr}
2: **Output:** Channel estimates with NLD cancellation $\hat{\mathbf{H}}^{(i^*)}_{l,m}$
3: $\hat{\mathbf{D}}^p = \mathrm{FFT}(f(\mathrm{IFFT}(\Phi)) - \kappa_0\mathrm{IFFT}(\Phi))$
4: **for** $m \leftarrow 1$ to N_{rb} **do**
5: $\quad\hat{\mathbf{H}}^{(0)}_{l,m} = \frac{\kappa_0^*}{|\kappa_0|^2\tau_p}\mathbf{Y}^p_{l,m}\Phi_m$
6: $\quad$**for** $i \leftarrow 1$ to $(N_{itr}-1)$ **do**
7: $\quad\quad\hat{\mathbf{H}}^{(i)}_{l,m} = \hat{\mathbf{H}}^{(0)}_{l,m} - \frac{\kappa_0^*}{|\kappa_0|^2\sqrt{\tau_p}}\hat{\mathbf{H}}^{(i-1)}_{l,m}\hat{\mathbf{D}}^p_m\Phi_m$
8: $\quad\quad$**if** $\|\hat{\mathbf{H}}^{(i)}_{l,m} - \hat{\mathbf{H}}^{(i-1)}_{l,m}\|^2 <= 10^{-10}$ **then**
9: $\quad\quad\quad$**break**
10: $\quad\quad$**end if**
11: $\quad$**end for**
12: **end for**

Remark 1: It is worth noticing that channels are estimated per-RB and are assumed to be identical for RUs within one RB. Then, for a given RB_i, each AP can effectively construct τ_p channel coefficients per-subcarrier m as

$$\hat{\mathbf{h}}_{k,l,m} = \hat{\mathbf{h}}_{k,l,m'} \text{ for } m' \in \mathrm{RB}_i. \tag{27}$$

Similarly, the proposed algorithm is applied for each RB.

$$\mathrm{SINR}_{k,m} = \frac{|\mathrm{DS}_{k,m}|^2}{\mathbb{E}\{|\mathrm{BU}_{k,m}|^2\} + \sum_{t=1,t\neq k}^{K}\mathbb{E}\{|\mathrm{UI}_{k,t,m}|^2\} + \sum_{t=1}^{K}\mathbb{E}\{|\mathrm{NLD}_{k,t,m}|^2\} + \mathbb{E}\{|\mathrm{GN}_{k,m}|^2\}}, \tag{21}$$

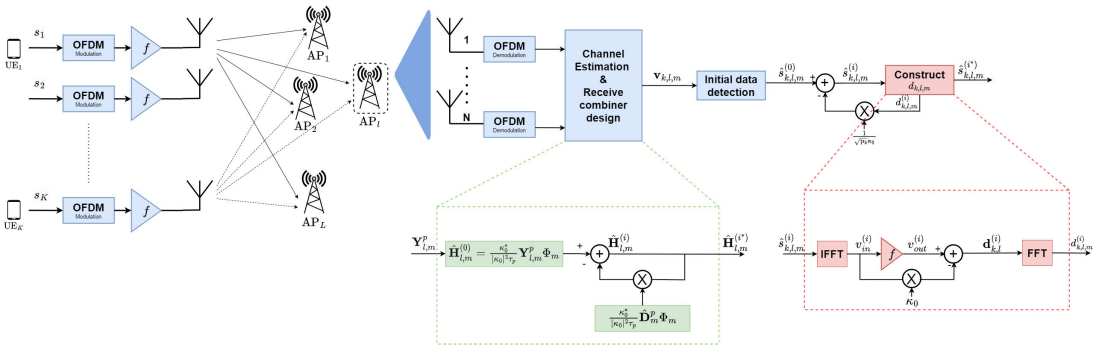


FIGURE 2. CF-mMIMO-OFDM system with NLD cancellation algorithms. The algorithms are implemented in a distributed manner, i.e, They are executed locally at each AP.

B. DISTRIBUTED JOINT UL DATA DETECTION WITH NLD CANCELLATION

The distributed NLD cancellation in data detection is illustrated in **Algorithm. 2**. The block diagram of the proposed algorithm is shown in Fig. 2. In the UL data transmission, each AP performs receive combining to obtain a combined signal that contains the contribution from multiple UEs. The resulting signal at each AP is then processed to approximate and remove the NLD introduced by the NL-PA of the UEs. The process of NLD cancellation technique involves two main parts: **individual distortion approximation** and **total NLD removal**. The first step consists in approximating the NLD introduced by each user. Contrary to NLD approximation in channel estimation, in the absence of the prior knowledge of the data, the approximation step is performed in the iterative process of the algorithm. Since the detected signal of each user is affected by the NLD received from all users, once the distortion introduced by each user has been approximated, the second step aims to cancel out the NLD effects caused by all users' transmission. This process is carried out at each AP for each user for all sub-carriers.

In the following, we detail the NLD cancellation process for a given user. Before starting the iterative process, it is important to note that a specific normalization step that depends on the used combining scheme is needed (step 3). The aim here is to separate the signal and distortion components from other components (i.e, multi-user interference, received NLD from other users and the thermal noise) in the initial detected signal of the user of interest. The resulting signal after this step is given by (28), as shown at the bottom of the next page, where

$$\begin{aligned}
&\eta_{k,l,m} \\
&= \sqrt{p_k}\kappa_0 \mathbf{v}_{k,l,m}^H \hat{\mathbf{h}}_{k,l,m}, \\
&= \begin{cases} \sqrt{p_k}\kappa_0 \hat{\mathbf{h}}_{k,l,m}^H \hat{\mathbf{h}}_{k,l,m} & \text{MR} \\ \sqrt{p_k}\kappa_0 \gamma_{k,l} & \text{FZF} \\ \sqrt{p_k}\kappa_0 \gamma_{k,l} \mathbb{1}_{k\in\mathcal{S}_l} + \sqrt{p_k}\kappa_0 \hat{\mathbf{h}}_{k,l,m}^H \hat{\mathbf{h}}_{k,l,m} \mathbb{1}_{k\in\mathcal{W}_l} & \text{PFZF} \end{cases}
\end{aligned} \tag{29}$$

The resulting signal is the input signal of the **individual distortion approximation** process which is carried out in an iterative manner. At each iteration, first, the distortion introduced the k-th user, $d_{k,l,m}$, is approximated by imitating the behavior of PAs and then the term $\frac{\hat{d}_{k,l,m}}{\sqrt{p_i}\kappa_0}$ is subtracted from the initial detected signal, $\hat{s}_{k,l,m}^{(0)}$. The iterative process of the algorithm is basically based on three steps:

- (step 6): The input signal is OFDM modulated into time-domain by applying the IFFT operation.
- (step 7): The resulting signal from (step 6) is amplified using the approximated PA model and then subtracted from its original version. Thus, we are constructing the NLD part introduced by the k-th user in the time-domain then reproduce its frequency-domain representation via the FFT operation.
- (step 9): Finally, the approximated NLD of the k-th user, $\hat{d}_{k,l,m}^{(i)}$, is subtracted from the initial normalized detected

Algorithm 2 NLD Cancellation in UL Data Detection at the l-Th AP

1: **Input:** $\hat{\mathbf{s}}_{k,l}, N_{itr}$
2: **Output:** Estimated data with NLD cancellation $\hat{\mathbf{s}}_{k,l}^{(i^*)}$
3: **Initialization:** $\hat{s}_{k,l,m}^{(0)} = \frac{\hat{s}_{k,l,m}}{\eta_{k,l,m}}, \forall k\ \forall m$
4: **for** $k \leftarrow 1$ to K **do**
5: **for** $i \leftarrow 0$ to $(N_{itr} - 1)$ **do**
6: $v_{in}^{(i)} = \text{IFFT}\left(\hat{\mathbf{s}}_{k,l}^{(i)}\right)$
7: $\hat{d}_{k,l}^{(i)} = \text{FFT}\left(f(v_{in}^{(i)}) - \kappa_0 v_{in}^{(i)}\right)$
8: **for** $m \leftarrow 1$ to N_{sub} **do**
9: $\hat{\mathbf{s}}_{k,l,m}^{(i+1)} = \hat{\mathbf{s}}_{k,l,m}^{(0)} - \frac{\hat{d}_{k,l,m}^{(i)}}{\sqrt{p_k}\kappa_0}$
10: **end for**
11: **if** $\|\hat{\mathbf{s}}_{k,l}^{(i+1)} - \hat{\mathbf{s}}_{k,l}^{(i)}\|^2 <= 10^{-10}$ **then**
12: **break**
13: **end if**
14: **end for**
15: **end for**
16: $i^* = i$
17: **for** $k \leftarrow 1$ to K **do**
18: **for** $m \leftarrow 1$ to N_{sub} **do**
19: $\hat{s}_{k,l,m}^{(i^*)} = \hat{s}_{k,l,m} - \sum_{t=1}^{K} \mathbf{v}_{k,l,m}^H \hat{\mathbf{h}}_{t,l,m} \hat{d}_{t,l,m}^{(i^*)}$
20: **end for**
21: **end for**

TABLE 2. Complexity analysis of channel estimation.

Channel estimation	
Scheme	Complexity
MMSE	$N_{rb}N\tau_p^2$
Algorithm 1	$N_{sub}\log(N_{sub})+N_{rb}N_{itr}(NK\tau_p+N^2\tau_p)$

signal in (28) for all subcarriers. As such, its NLD is decreased in an iterative manner.

In this way, if the AP was able to approximate the distortion, the outcome of i-th iteration is less-distorted compared to that of the (i-1)-th iteration and the algorithm tends to converge to the non-distorted data after a number of iterations. The iterative process ends either when the maximum number of iterations is reached, $i^* = N_{itr}$, or when almost the same corrected signals are achieved within two consecutive iterations (step 11), $i^* < N_{itr}$. Once the NLD introduced by each user has been approximated, the second step, **total NLD removal**, consists in subtracting this distortion from the initial detected signal of each user for all subcarriers based on channel estimates (step 19).

C. COMPLEXITY ANALYSIS OF THE PROPOSED TECHNIQUES

In this section, the computational complexity analysis is provided for the proposed schemes compared with the conventional UL channel estimation and UL data detection schemes. In terms of complexity analysis, we consider the number of complex operations, mostly complex multiplications, needed for the proposed algorithms. The added complexities of the proposed algorithms are given as follows

1) Algorithm 1: The added computational complexity of NLD cancellation to MMSE channel estimation process arises from two main operations in the algorithm: the IFFT/FFT (step 3) in the initial NLD approximation and the iterative NLD subtraction (step 7). The complexity of the IFFT/FFT is $\mathcal{O}(N_{sub}\log(N_{sub}))$, while the NLD subtraction operation has a complexity of $\mathcal{O}(N_{itr}(NK\tau_p + N^2\tau_p))$. This process is performed per-RB.

TABLE 2 provides the complexity comparison between **Algorithm 1** and the conventional MMSE channel estimation scheme.

2) Algorithm 2: Regarding the individual NLD approximation step in the proposed algorithm, it is evident that the additional computational, compared to data detection in the traditional CF-mMIMO combining scheme, mainly stems from two key operations: the IFFT/FFT (steps 6 and 7) and the NLD subtraction (step 9). The complexity of the IFFT/FFT is $\mathcal{O}(N_{sub}\log(N_{sub}))$, while the NLD subtraction operation has a complexity of $\mathcal{O}(N_{sub})$. Thus the total additional computational complexity per iteration is given by $\mathcal{O}(K(N_{sub}+2N_{sub}\log(N_{sub})))$. The second part of the algorithm, total NLD removal, the complexity is $\mathcal{O}(KN_{sub}N^2)$ for all users at all subcarriers. The parts of the algorithm are completely independent, then the computational complexity of the overall data NLD cancellation process is the sum of their individual complexities.

TABLE 3. Complexity analysis of UL data detection.

Combining design & Initial data detection	
Scheme	Complexity
MR	-
FZF	$N_{sub}\left(\frac{3N\tau_p^2}{2}+\frac{N\tau_p}{2}+\frac{\tau_p^3-\tau_p}{3}\right)$
PFZF	$N_{sub}\left(\frac{3N\nu_l^2}{2}+\frac{N\nu_l}{2}+\frac{\nu_l^3-\nu_l}{3}\right)$
Initial data detection	$KN_{sub}N^2$
Algorithm 2	
Scheme	Complexity
Algorithm 2 (Individual NLD approximation)	$KN_{itr}(N_{sub}+2N_{sub}\log(N_{sub}))$
Algorithm 2 (Total NLD removal)	$KN_{sub}N^2$

TABLE 3 provides the complexity comparison between the conventional UL data detection (i.e., combining design and initial UL data detection) and different parts of **Algorithm 2**.

V. SIMULATION RESULTS

This section presents numerical results that show the effectiveness of the proposed scheme. The system being considered is a CF-mMIMO-OFDM, consisting of L APs and K single-antenna users, operating in a 1km×1km square-shaped area. The locations of both APs and UEs are randomly and uniformly distributed. The large-scale channel coefficient between UE_k and AP_l is modeled as follows

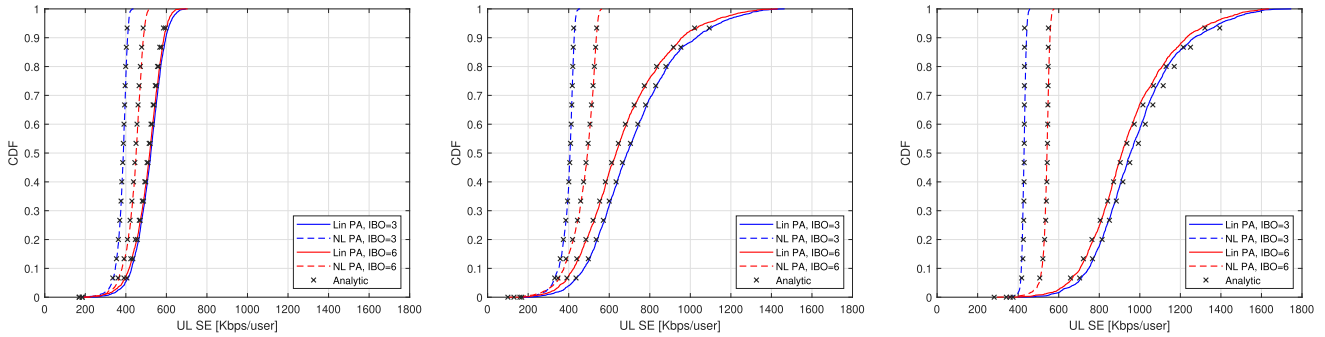
$$\beta_{k,l} = \text{PL}_{k,l}10^{\frac{\sigma_{sh}z_{k,l}}{10}}, \tag{30}$$

where σ_{sh} is the standard deviation, $z_{k,l} \sim \mathcal{N}(0,1)$ and $PL_{k,l}$ is the path loss from UE_k to AP_l. Considering the micro cell path loss models of the 3GPP model which is given by [6]

$$\text{PL}_{k,l}[dB] = -36.7\log_{10}\left(\frac{\text{Dist}_{k,l}}{1m}\right) - 22.7 - 26\log_{10}\left(\frac{F_c}{1\text{GHz}}\right), \tag{31}$$

where $\text{Dist}_{k,l}$ is the distance between UE_k and AP_l and $F_c = 3.5$GHz is the carrier frequency. The key simulation parameters are summarized in Table 4. The parameters are chosen to align with those used in the literature [18]. The parameters of PAs are set as follows G = 16, Vsat = 1.9, p = 1.1, A = -345, B = 0.17 and q = 4. Hereafter, "Lin PA" refers to the ideal case where there are no HWIs and "NL PA" denotes the case under HWIs and without NLD cancellation. For the PFZF combining scheme, we assume $\delta = 95$.

$$\hat{s}^{(0)}_{k,l,m} = \frac{\hat{s}_{k,l,m}}{\eta_{k,l,m}} = s_{k,m} + \frac{d_{k,m}}{\sqrt{p_k}\kappa_0} + \sum_{t=1,t\neq k}^{K}\sqrt{p_t}\kappa_0\frac{\mathbf{v}^H_{k,l,m}\mathbf{h}_{t,l,m}}{\eta_{k,l,m}}\times s_{t,m} + \sum_{t=1,t\neq k}^{K}\frac{\mathbf{v}^H_{k,l,m}\mathbf{h}_{t,l,m}}{\eta_{k,l,m}}\times d_{t,m} + \frac{\mathbf{v}^H_{k,l,m}\mathbf{n}_{l,m}}{\eta_{k,l,m}}, \tag{28}$$



(a) MR (b) FZF (c) PFZF

FIGURE 3. CDF of uplink SE per-RB and per-user when L=100, N=16, $K = \tau_p = 15$ and $\sigma^2 = -93$dBm for IBO=[3, 6]dB.

TABLE 4. Simulation parameters.

Parameter	Value	Parameter	Value
σ_{sh}	4 dB	ξ	0.5
N_{sub}	256	N_{sc}	12
Δ_f	15 kHz	N_c	14
Coherence time	1 ms	Coherence bandwidth	168

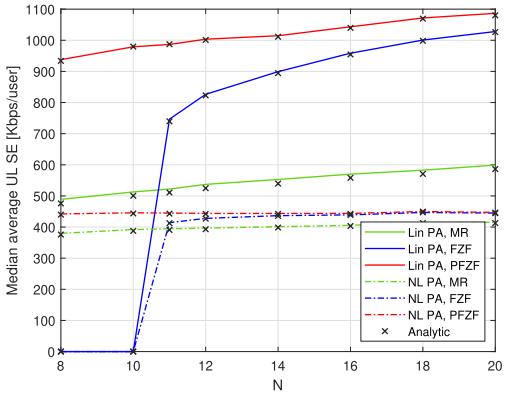


FIGURE 4. Median of average uplink SE per-RB when $L = 100$, $K = \tau_p = 10$, IBO = 3 dB and $\sigma^2 = -93$ dBm.

A. IMPACT OF HWI ON THE UPLINK SE OF CF-MMIMO-OFDM

The impact of the NLD on the CDF of the uplink SE of CF-mMIMO-OFDM system is shown in Fig. 3, for $L = 100$, $N = 16$ and $K = 15$.

First, it can be seen that the results obtained from closed-form expressions (black cross) are close to the ones given by Monte-Carlo simulations for linear (solid curves) and non-linear cases (dashed curves). Moreover, regardless the used combining scheme, we note that the impact of NLD reduces when the IBO increases. However, it may also be noted that the system achieves better SE for "Lin PA" case as the IBO factor decreases. In fact, increasing the IBO reduces the NLD effects at the cost of the power efficiency. It should be noted that the impairments effect becomes more severe for better users with better channel gains (high percentiles). In fact, the total NLD is proportional to users' channel quality. On the other hand, the achieved SE of UEs with poor channel conditions is dominated by noise and inter-user interference and hence the impact of PA distortions is not clearly seen. As illustrated, at CDF of 95% for IBO = 3dB, NLD causes a loss of 33% 63% and 69% in SE of the considered system based on MR, FZF and PFZF, respectively. While the MR scheme is less affected due to the presence of noise and multi-user interference, the FZF and PFZF schemes are more susceptible to NLD, leading to a sever degradation of the system SE.

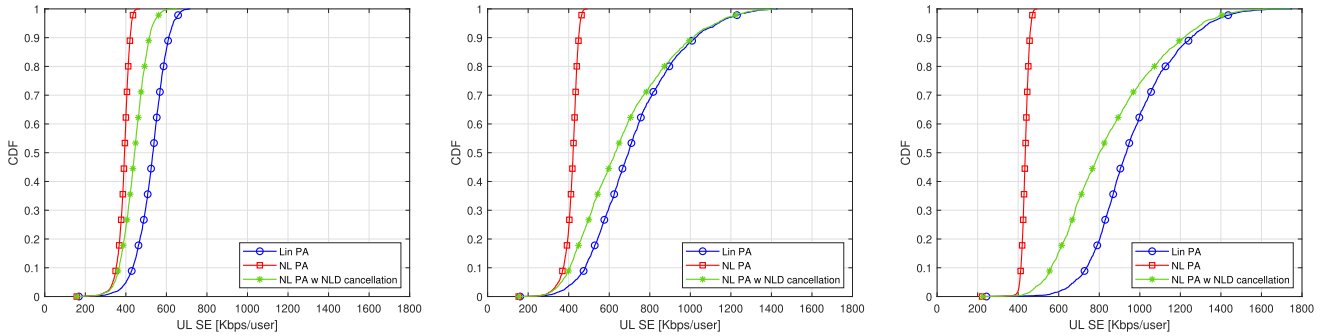
Fig. 4 shows the comparison of the median average SE of the system based on different combining schemes with and without NLD cancellation, while varying the number of antennas per AP. We note here that the median can provide a better representation of the SE experienced by most users since the average SE might be affected by the outline values. First, for the "Lin PA" case, we observe that increasing N leads to higher average spectral efficiency for all combining schemes. By increasing the number of antennas, the system benefits from increased spatial degrees of freedom, improved beamforming, and enhanced interference suppression capabilities. Second, for the "Lin PA" case, we note that the PFZF combining scheme tends to provide the highest average SE, followed by FZF, while MR offers the lowest average SE. Based on the MR combining scheme, the received signals are already affected by noise and multi-user interference. Therefore, the impact of NLD on the MR combining scheme may be relatively less prominent compared to PFZF and FZF. However, the overall performance of MR can still be affected by the presence of NLD.

On the other hand, we can observe that, for the "NL PA" case, the MR combining scheme still performs worse than FZF and PFZF. However, the achievable average SE is almost the same for FZF and PFZF combining schemes. In fact, the distortion can degrade the performance of interference suppression and compromise the benefits of these schemes.

B. PERFORMANCE EVALUATION OF THE PROPOSED NLD CANCELLATION SCHEMES

1) NLD CANCELLATION/SPECTRAL EFFICIENCY

Fig. 5 shows the CDFs of the SEs achieved by the CF-mMIMO-OFDM adopting, respectively, the MR, FZF and PFZF combining schemes. For the sake of comparison, the figure depicts the achievable SE after applying the



(a) MR (b) FZF (c) PFZF

FIGURE 5. CDF of uplink SE per-RB and per-user when $L = 100$, $N = 16$, $K = \tau_p = 15$, $\sigma^2 = -93$dBm and IBO = 3dB.

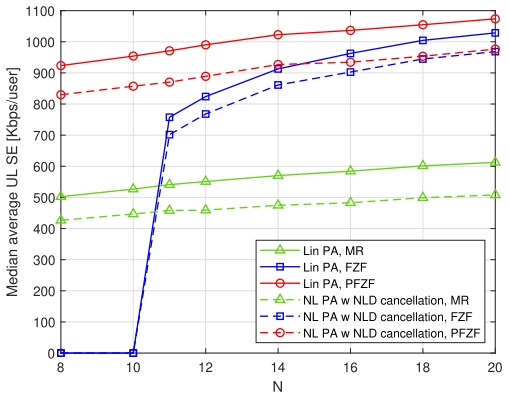


FIGURE 6. Median of average uplink SE per-RB versus the number of antennas per AP (N) when $L = 100$, $K = \tau_p = 10$, IBO = 3dB and $\sigma^2 = -93$dBm.

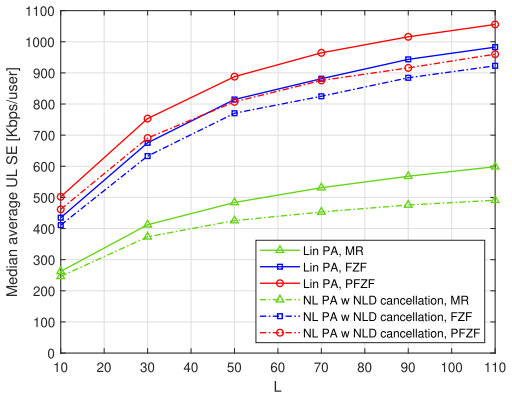


FIGURE 7. Median of average uplink SE per-RB versus the number of APs (L) when $N = 16$, $K = \tau_p = 10$, IBO = 3dB and $\sigma^2 = -93$ dBm.

proposed NLD cancellation techniques and "Lin PA" and "NL PA" cases. We can clearly note that, for all combining schemes, the NLD cancellation technique is generally more effective for high percentiles of users with good channel conditions, rather than users with poor channel conditions.

We can also see that the effectiveness of the proposed scheme is limited when the MR combining scheme is used. Particularly, at CDF of 95%, the achievable SE is improved from 428 to 536 kbps/user (gain of 25%). For low percentiles, at CDF of 20%, the provided gain decreases to 6%. When using the FZF combining scheme, the achievable gain after NLD cancellation is only 16.7%, at CDF of 2%. At CDF of 95%, the achievable SE through FZF combining is 2.5 times better than the achieved SE in the "NL PA" case. Moreover, the gain in the achievable SE, when the PFZF combining scheme is used, is about 49% at CDF of 2%. At CDF of 95%, the achievable SE through PFZF combining is 2.8 times better than the achieved SE in the "NL PA" case. While the MR combining scheme is limited by noise and multi-user interference, the FZF combining scheme faces challenges related to noise amplification and channel estimation errors that can affect the effectiveness of NLD cancellation in data detection.

Fig. 6 shows the median average SE with $K = \tau_p = 10$ versus N for the three considered combiners. The figure also compares the median average SE of "Lin PA" and "NL PA w NLD cancellation" cases. We can note that, for all the combining schemes, the average SE increases when increasing the number of antennas at the APs. Moreover, the PFZF still provides the highest SE while MR gives the lowest SE.

To further evaluate the effectiveness of the proposed schemes, we plot in Fig. 7 the average UL per-RB SE as function of the number of APs (L). Based on the obtained results, our proposed solution demonstrates an improvement in UL SE. Particularly, we observe a gain enhancement as the number of APs increases.

2) CONVERGENCE ANALYSIS

In this section, we provide a convergence analysis of the proposed iterative process in **Algorithm 1** and **Algorithm 2**. For the sake of illustration, we consider the system configuration illustrated in Fig. 8 with 100 APs and 15 UEs randomly distributed within an area of size 1km×1km.

We compare the convergence behavior and accuracy of the proposed NLD cancellation algorithms for different AP locations and based on different combining schemes. Since the NLD cancellation is performed per user, we assume that UE_1 is the user of interest. For the two proposed distortion cancellation methods, we present the mean square error (MSE) between the original data of UE_1 and its estimated data after NLD cancellation at the following APs: AP_{34},

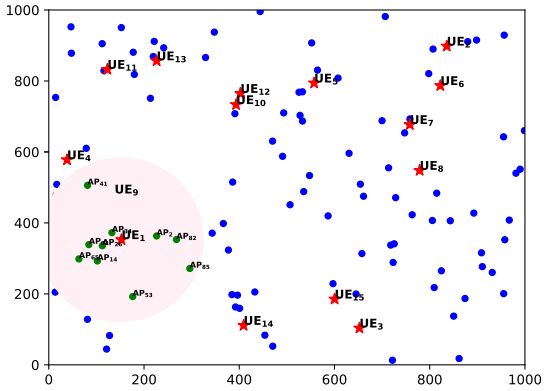
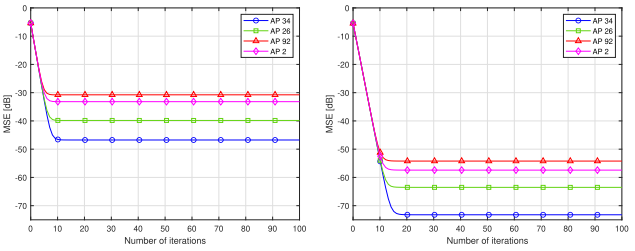


FIGURE 8. Random network system topology of a CF-mMIMO system where 15 UEs and 100 APs with $N = 16$ are randomly distributed inside the 1km x 1km square.

AP_{26}, AP_{92}, AP_2 versus the number of iterations. The algorithm convergence is analyzed for two noise powers: $\sigma^2 = -93$ dBm and $\sigma^2 = -120$ dB.

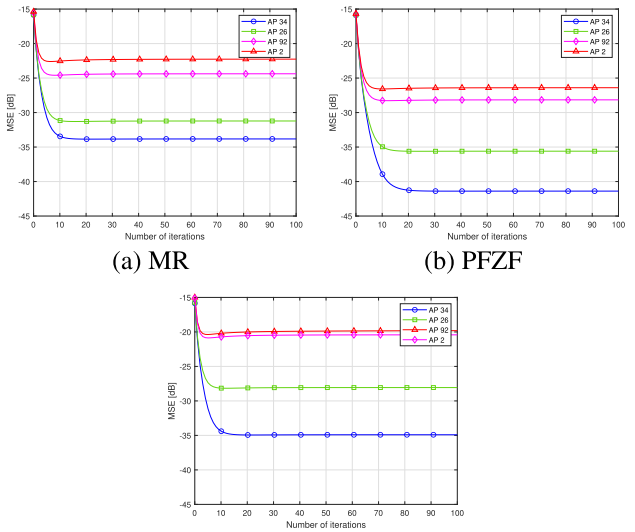
Starting with the distributed NLD cancellation for channel estimation, we show in Fig. 9 the number of iterations needed for the convergence of the proposed **Algorithm 1** for the two considered noise figures. It is evident that the MSE of all APs decreases as the noise power decreases. Specifically, we observe that at a noise power level of $\sigma^2 = -120$dBm, the algorithm requires 20 iterations for convergence at AP_{34}, resulting in an MSE value of -73dB. In contrast, at the same AP but with a noise power level of $\sigma^2 = -93$dB, only 12 iterations are needed, yielding to an MSE of -46dB. Let's consider AP_{92}, which is located farther away from UE_1. At a noise power level of $\sigma^2 = -120$dBm, the algorithm converges within 23 iterations, achieving an MSE of -54dB. Conversely, at a noise power level of $\sigma^2 = -93$dB, convergence is reached in 10 iterations, with an MSE value of -30dB.

Clearly, a decrease in noise power leads to more accurate distortion approximation and improved cancellation performance, resulting in a lower MSE in the distortion approximation. However, it should be noted that as the noise level decreases, the number of iterations required for convergence may increase. This is because the algorithm can identify and address small amount of distortions that could be hidden by the high noise levels. Hence, the algorithm might need to perform more iterations to fine-tune the corrections for these remaining amounts of distortion, leading to slower convergence. Moreover, when considering a fixed noise power, the achievable MSE in NLD approximation can vary among different APs. Specifically, APs situated closer to UE_1 generally possess the potential to attain a lower MSE. Nevertheless, achieving a low MSE value may require a higher number of iterations in the approximation process. This is primarily because APs in closer proximity to the user typically experience stronger received signal power compared to those located farther away. The higher signal power results in an improved SINR, thereby enhancing the accuracy of the approximation process.



(a) $\sigma^2 = -93$ dBm (b) $\sigma^2 = -120$ dBm

FIGURE 9. MSE versus the number of iterations N_{itr} for channel estimation when $N = 16$, $K = \tau_p = 15$, $L = 100$ and IBO= 3dB.

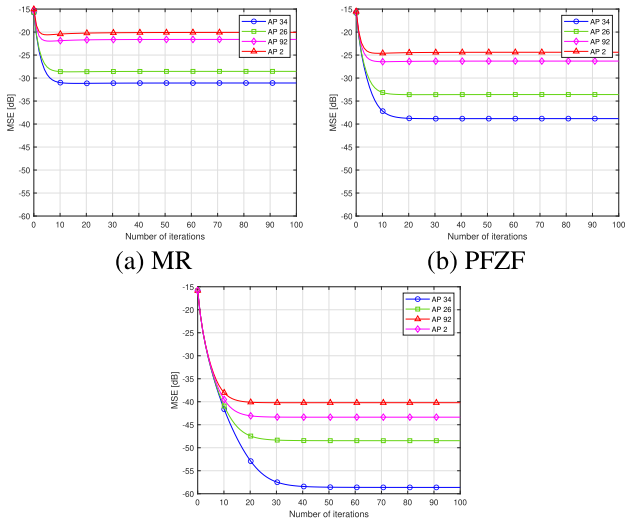


(a) MR (b) PFZF

(c) FZF

FIGURE 10. MSE versus the number of iterations N_{itr} for Algorithm 1 when $N = 16$, $K = \tau_p = 15$, $L = 100$, IBO= 3dB and $\sigma^2 = -93$dBm.

In Fig.10 and 11, we plot the MSE performance of NLD cancellation in data detection based on MR, PFZF and FZF combining for the same configuration when $\sigma^2 = -93$dBm and $\sigma^2 = -120$dBm, respectively. First, we can observe from the two figures that, regardless the noise level, the effectiveness of NLD cancellation depends on the used combining scheme at the AP. As depicted in Figure 10, we observe that NLD approximation is less effective when the noise power level is $\sigma^2 = -93$ dBm. This is particularly evident for APs located farther away from the user UE_1, such as AP_{92} and AP_2. In fact, in scenarios with high noise levels, the presence of noise significantly deteriorates the quality of the received signal, posing a greater challenge for NLD approximation, especially for APs with weaker channel conditions towards UE_1. Furthermore, it is worth noticing that the FZF combining scheme exhibits a relatively high error floor value (-28 dB) at AP_{26} compared to MR and PFZF combining schemes, which achieve -31 dB and -35 dB, respectively. This is due to the fact that the noise amplification associated to FZF combining scheme limits its effectiveness in accurately estimating and canceling NLD, particularly when encountering unfavorable channel conditions. On the other hand, MR combining scheme may be susceptible to multi-user interference. However,



(a) MR (b) PFZF

(c) FZF

FIGURE 11. MSE versus the number of iterations N_{itr} for Algorithm 1 when $N = 16$, $K = \tau_p = 15$, $L = 100$, IBO= 3 and $\sigma^2 = -120$dBm.

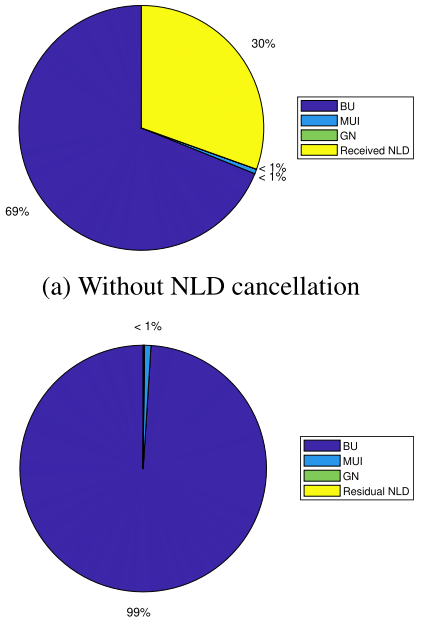
being a compromise between MR and FZF schemes, PFZF combining scheme provides a better NLD approximation capability.

Moreover, it is worth noticing that both MR and FZF combining schemes converge in fewer iterations (10 iterations), whereas PFZF requires 20 iterations. This arises from the fact that additional cancellation attempts may not significantly improve the distortion approximation due to the amplified noise level in the case of FZF and the interference level in the case of MR.

In fact, As shown in Figure 11, in low noise conditions, FZF can effectively suppress interference and amplify the desired signal with minimal noise amplification. This allows for accurate approximation and cancellation of NLD. On the other hand, MR combining may be more susceptible to interference dominance, which can hinder the accurate approximation of NLD in low noise conditions. On the other hand, despite reducing noise amplification compared to FZF combining, PFZF combining scheme do not achieve the same level of interference suppression.

Remark 2: The convergence speed of the iterative NLD approximation depends on various factors, including the used combining scheme and the quality of initial detected signal of each user. The cancellation algorithm needs to be carefully designed to approximate and cancel total NLD in the presence of multi-user interference and considering the distributed nature of the APs.

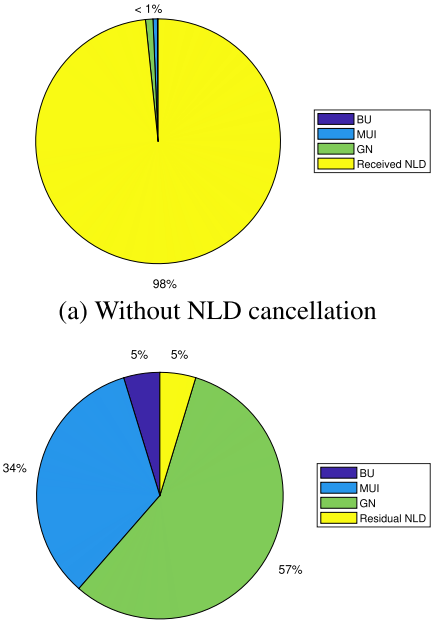
The power of unwanted signals affecting the quality of the desired signal of UE_1 at AP_{34} are shown in Fig. 12, 13 and 14 when using MR, FZF and PFZF combining schemes, respectively, with and without NLD cancellation. Unwanted signals consist of multiple components including beamforming uncertainty (BU), multi-user interference (MUI), the total received NLD (Received NLD), the residual NLD after applying the proposed cancellation schemes (Residual NLD) and the noise term (GN).



(a) Without NLD cancellation

(b) After NLD cancellation

FIGURE 12. The power of unwanted signals in the desired signal of UE_1 at AP_{34} before and after NLD cancellation using the MR combining.

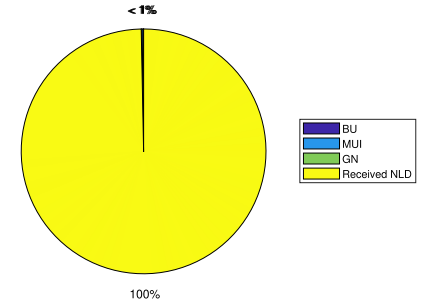


(a) Without NLD cancellation

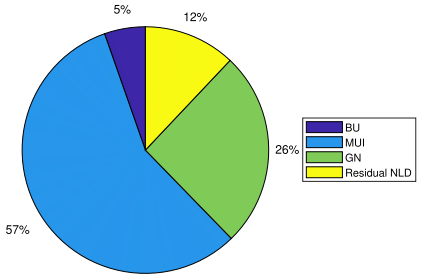
(b) After NLD cancellation

FIGURE 13. The power of unwanted signals in the desired signal of UE_1 at AP_{34} before and after NLD cancellation using the FZF combining.

We can see from Fig. 12 that, without NLD cancellation, the NLD can still affect the MR combining performance which is highly susceptible to BU. After NLD cancellation, the level of the received distortion is reduced but BU still have a more significant impact on the performance of MR combining. As shown in Fig. 13, when NLD is present, it dominates the non-desired signals such as MUI and GN when using FZF combining. After NLD cancellation,



(a) Without NLD cancellation



(b) After NLD cancellation

FIGURE 14. The power of unwanted signals in the desired signal of UE_1 at AP_{34} before and after NLD cancellation using the PFZF combining.

the level of the received distortion is reduced, and GN becomes the dominant factor affecting the performance of FZF combining. Moreover, as depicted in Fig. 14, the PFZF combining scheme is highly impacted by NLD that dominates the undesired signals without NLD cancellation. However, after NLD cancellation the level of received distortions becomes negligible compared to MUI.

Remark 3: With NLD cancellation, regardless the used combining scheme and at the different APs, the proposed scheme reduces the level of the total received NLD. When NLD cancellation is performed successfully, it effectively removes or reduces the amount of the NLD component from the received signal. As a result, the distortion level is significantly reduced and becomes negligible compared to other factors affecting the signal quality, such as multi-user interference and noise power. We can conclude that the effectiveness of NLD cancellation in UL CF-mMIMO is indeed influenced by the SINR of the received signal. Particularly, a higher SINR generally allows for better approximation of the NLD introduced by each user. In fact, when the SINR is high, it means that the desired user's signal is strong compared to the interference and noise levels. In this case, the system has a clearer and more distinguishable signal from the desired user, making it easier to approximate the NLD introduced by that user and more effective total NLD cancellation.

3) COMPLEXITY EVALUATION

Based on the complexity analysis provided in TABLE 3, the complexity of the FZF combining scheme serves as a reference to evaluate the computational requirements of UL data detection in CF-mMIMO-OFDM systems. For $N_{sub} = 256$, $\tau_p = K = 15$, $N = 16$, $\sigma^2 = -93$dBm and $N_{itr} = 10$, the added complexity related to the NLD cancellation process in **Algorithm 2** is approximately 50% of the complexity of the UL data detection process using FZF combining scheme. Therefore, the proposed NLD cancellation scheme is an efficient and low-complexity solution that mitigates the impact of HWIs caused by PAs at UEs while maintaining a reasonable computational requirements.

VI. CONCLUSION

In this study, we investigated the performance of a CF-mMIMO systems in the presence of HWIs at UEs. We derived the achievable SE for an UL CF-mMIMO-OFDM system affected by non-linear power amplifier (NL PA) distortion based on three combining schemes: MR, FZF and PFZF. Our simulation results aligned closely with the analytical derivation. We proposed two distributed cancellation techniques to effectively address NLD effects on both channel estimation and data detection. The goal was to mitigate the impact of NLD and improve the overall system performance. Importantly, we found that the NLD cancellation schemes depends on channel conditions between the users and APs. Moreover, data detection depends also on the specific combining scheme used (MR, FZF, or PFZF). Both the FZF and PFZF combining schemes are observed to be more effective in mitigating the distortions caused by NLD compared to the MR combining scheme. On the other hand, PFZF combining scheme still achieves better SE compared to the FZF combining scheme. Numerical results show the effectiveness of the proposed NLD cancellation algorithm. Particularly, the proposed NLD cancellation algorithms leads to a SE gain of $2.5X$ and $2.8X$ at a CDF of 95% when using FZF and PFZF combining schemes, respectively. As a perspective, we aim to introduce a distributed machine learning architecture with local inference to enhance the trade-off between performance and time-complexity in CF-mMIMO-OFDM systems under HWIs.

APPENDIX. SINR COMPUTATION

To have the analytical expression for SINR in (21), we need to derive the expressions of $|\mathrm{DS}_{k,m}|^2$, $\mathbb{E}\{|\mathrm{BU}_{k,m}|^2\}$, $\mathbb{E}\{|\mathrm{UI}_{k,t,m}|^2\}$, $\mathbb{E}\{|\mathrm{NLD}_{k,t,m}|^2\}$ and $\mathbb{E}\{|\mathrm{GN}_{k,m}|^2\}$. We note that the SINR expression is based on the received signal in (20) which depends on the used combining vector. In the following, we give the proof of the SINR expression when using the PFZF combining vector. On the other hand, we note that the SINR expressions when using MR and FZF receive combining vectors could be given when ($\mathcal{A}_k = \emptyset$) and ($\mathcal{B}_k = \emptyset$ and $\nu_l = \tau_p$), respectively.

According to (27) and (28) in [23] and as UEs do not share pilots, $\forall k$ when $l \in \mathcal{A}_{k,m}$, we have

$$(\mathbf{v}_{k,l,m}^{PFZF})^H \hat{\mathbf{h}}_{t,l,m} = (\mathbf{v}_{k,l,m}^{FZF})^H \hat{\mathbf{h}}_{t,l,m} = \begin{cases} \gamma_{k,l}, & t = k \\ 0, & t \neq k \end{cases} \quad (32)$$

Otherwise, when $l \in \mathcal{B}_{k,m}$

$$\mathbb{E}\left\{(\mathbf{v}_{k,l,m}^{PFZF})^H \hat{\mathbf{h}}_{k,l,m}\right\} = \mathbb{E}\left\{(\mathbf{v}_{k,l,m}^{MR})^H \hat{\mathbf{h}}_{k,l,m}\right\}, \quad (33)$$
$$= \begin{cases} N\gamma_{k,l}, & t = k \\ 0, & t \neq k \end{cases}$$

According to (32) and (33), $|\mathrm{DS}_{k,m}|^2$ is given by

$$\begin{aligned} &|\mathrm{DS}_{k,m}|^2 \\ &= \left|\mathbb{E}\left\{\sum_{l=1}^{L} \sqrt{p_k}\kappa_0 a_{k,l,m}^* (\mathbf{v}_{k,l,m}^{PFZF})^H \mathbf{h}_{k,l,m}\right\}\right|^2, \\ &= \left|\sqrt{p_k}\kappa_0 a_{k,l,m}^* \mathbb{E}\left\{\sum_{l\in\mathcal{A}_{k,m}} (\mathbf{v}_{k,l,m}^{FZF})^H \mathbf{h}_{k,l,m}\right.\right. \\ &\quad \left.\left. + \sum_{l\in\mathcal{B}_{k,m}} (\mathbf{v}_{k,l,m}^{MR})^H \mathbf{h}_{k,l,m}\right\}\right|^2, \\ &= p_k|\kappa_0|^2 \left|\sum_{l\in\mathcal{A}_{k,m}} a_{k,l,m}^* \gamma_{k,l} + \sum_{l\in\mathcal{B}_{k,m}} N a_{k,l,m}^* \gamma_{k,l}\right|^2. \end{aligned} \quad (34)$$

The term $\mathbb{E}\{|\mathrm{BU}_{k,m}|^2\}$ is given by

$$\begin{aligned} &\mathbb{E}\{|\mathrm{BU}_{k,m}|^2\} \\ &= p_k|\kappa_0|^2 \mathbb{E}\left\{\left|\sum_{l=1}^{L} a_{k,l,m}^* (\mathbf{v}_{k,l,m}^{PFZF})^H \mathbf{h}_{k,l,m}\right|^2\right\} - |\mathrm{DS}_{k,m}|^2, \end{aligned} \quad (35)$$

where

$$\begin{aligned} &\mathbb{E}\left\{\left|\sum_{l=1}^{L} a_{k,l,m}^* (\mathbf{v}_{k,l,m}^{PFZF})^H \mathbf{h}_{k,l,m}\right|^2\right\} \\ &= \mathbb{E}\left\{\left|\sum_{l\in\mathcal{A}_{k,m}} a_{k,l,m}^* (\mathbf{v}_{k,l,m}^{FZF})^H \mathbf{h}_{k,l,m}\right.\right. \\ &\quad \left.\left. + \sum_{l\in\mathcal{B}_{k,m}} a_{k,l,m}^* (\mathbf{v}_{k,l,m}^{MR})^H \mathbf{h}_{k,l,m}\right|^2\right\}, \\ &= \mathbb{E}\left\{\left|\sum_{l\in\mathcal{A}_{k,m}} a_{k,l,m}^* (\mathbf{v}_{k,l,m}^{FZF})^H \mathbf{h}_{k,l,m}\right|^2\right\} \\ &\quad + 2\mathbb{E}\left\{\left(\sum_{l\in\mathcal{A}_{k,m}} a_{k,l,m}^* (\mathbf{v}_{k,l,m}^{FZF})^H \mathbf{h}_{k,l,m}\right)\right. \\ &\quad \left.\times \left(\sum_{l\in\mathcal{B}_{k,m}} a_{k,l,m}^* (\mathbf{v}_{k,l,m}^{MR})^H \mathbf{h}_{k,l,m}\right)\right\} \\ &\quad + \mathbb{E}\left\{\left|\sum_{l\in\mathcal{B}_{k,m}} a_{k,l,m}^* (\mathbf{v}_{k,l,m}^{MR})^H \mathbf{h}_{k,l,m}\right|^2\right\}, \end{aligned}$$

$$\begin{aligned} &\overset{(a)}{=} \left(\sum_{l\in\mathcal{A}_{k,m}} |a_{k,l,m}^*|^2 \frac{\gamma_{k,l}(\beta_{k,l} - \gamma_{k,l})}{(N - \nu_l)}\right. \\ &\quad \left. + \sum_{l\in\mathcal{B}_{k,m}} |a_{k,l,m}^*|^2 N\gamma_{k,l}\beta_{k,l}\right). \end{aligned} \quad (36)$$

where (a) is computed as in [23].

Then, following the analysis in [23], the inter-user interference term, $\mathbb{E}\{|\mathrm{UI}_{k,t,m}|^2\}$ $(t \neq k)$, is given by

$$\begin{aligned} \mathbb{E}\{|\mathrm{UI}_{k,t,m}|^2\} &= \mathbb{E}\left\{\left|\sum_{l=1}^{L} \sqrt{p_t}\kappa_0 a_{k,l,m}^* (\mathbf{v}_{k,l,m}^{PFZF})^H \mathbf{h}_{t,l,m}\right|^2\right\}, \\ &= p_t|\kappa_0|^2 \left(\sum_{l\in\mathcal{A}_{k,m}} |a_{k,l,m}^*|^2 \frac{\gamma_{k,l}(\beta_{t,l} - \gamma_{t,l})}{(N - \nu_l)}\right. \\ &\quad \left. + \sum_{l\in\mathcal{B}_{k,m}} |a_{k,l,m}^*|^2 N\gamma_{k,l}\beta_{t,l}\right). \end{aligned} \quad (37)$$

Since $(\mathbf{v}_{k,l,m}^{PFZF})^H \mathbf{h}_{t,l,m}$ and $d_{t,m}$ are independent, the NLD term is given by

$$\begin{aligned} &\mathbb{E}\{|\mathrm{NLD}_{k,t,m}|^2\} \\ &= \mathbb{E}\left\{\left|\sum_{l=1}^{L} a_{k,l,m}^* (\mathbf{v}_{k,l,m}^{PFZF})^H \mathbf{h}_{t,l,m} d_{t,m}\right|^2\right\}, \\ &= \mathbb{E}\left\{\left|\sum_{l=1}^{L} a_{k,l,m}^* (\mathbf{v}_{k,l,m}^{PFZF})^H \mathbf{h}_{t,l,m}\right|^2\right\} \mathbb{E}\left\{\left|d_{t,m}\right|^2\right\}. \end{aligned} \quad (38)$$

Thus, following (36) and (37), $\mathbb{E}\{|\mathrm{NLD}_{k,t,m}|^2\}$ is computed as follows,

If $t = k$ we have

$$\begin{aligned} &\mathbb{E}\{|\mathrm{NLD}_{k,t,m}|^2\} \\ &= \sigma_d^2 \left(\sum_{l\in\mathcal{A}_k} a_{k,l,m}^* \gamma_{k,l} + \sum_{l\in\mathcal{B}_k} a_{k,l,m}^* N\gamma_{k,l}\right) \\ &\quad + \sigma_d^2 \left(\sum_{l\in\mathcal{A}_k} |a_{k,l,m}^*|^2 \frac{\gamma_{k,l}(\beta_{t,l} - \gamma_{t,l})}{(N - \nu_l)}\right. \\ &\quad \left. + \sum_{l\in\mathcal{B}_k} |a_{k,l,m}^*|^2 N\gamma_{k,l}\beta_{t,l}\right), \end{aligned} \quad (39)$$

Otherwise, if $t \neq k$

$$\begin{aligned} \mathbb{E}\{|\mathrm{NLD}_{k,t,m}|^2\} &= \sigma_d^2 \left(\sum_{l\in\mathcal{A}_k} |a_{k,l,m}^*|^2 \frac{\gamma_{k,l}(\beta_{t,l} - \gamma_{t,l})}{(N - \nu_l)}\right. \\ &\quad \left. + \sum_{l\in\mathcal{B}_k} |a_{k,l,m}^*|^2 N\gamma_{k,l}\beta_{t,l}\right) \end{aligned} \quad (40)$$

Finally, the noise term is given as follows

$$\begin{aligned}
&\mathbb{E}\{|\mathrm{GN}_{k,m}|^2\} \\
&= \mathbb{E}\left\{\left|\sum_{l\in\mathcal{A}_k} a^*_{k,l,m}(\mathbf{v}^{FZF}_{k,l,m})^H \mathbf{h}_{k,l,m}\mathbf{n}_{l,m}\right|^2\right\} \\
&\quad + \mathbb{E}\left\{\left|\sum_{l\in\mathcal{B}_k} a^*_{k,l,m}(\mathbf{v}^{MR}_{k,l,m})^H \mathbf{h}_{k,l,m}\mathbf{n}_{l,m}\right|^2\right\}, \\
&= \sigma^2 \sum_{l\in\mathcal{A}_k} |a^*_{k,l,m}|^2 \frac{\gamma_{k,l}}{(N-\nu_l)} + \sigma^2 \sum_{l\in\mathcal{B}_k} |a^*_{k,l,m}|^2 N\gamma_{k,l}.
\end{aligned} \tag{41}$$

REFERENCES

[1] I. F. Akyildiz, A. Kak, and S. Nie, "6G and beyond: The future of wireless communications systems," *IEEE Access*, vol. 8, pp. 133995–134030, 2020.

[2] S. Chen, J. Zhang, J. Zhang, E. Björnson, and B. Ai, "A survey on user-centric cell-free massive MIMO systems," *Digit. Commun. Netw.*, vol. 8, no. 5, pp. 695–719, Oct. 2022.

[3] M. Alsabah, M. A. Naser, B. M. Mahmmod, S. H. Abdulhussain, M. R. Eissa, A. Al-Baidhani, N. K. Noordin, S. M. Sait, K. A. Al-Utaibi, and F. Hashim, "6G wireless communications networks: A comprehensive survey," *IEEE Access*, vol. 9, pp. 148191–148243, 2021.

[4] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1834–1850, Mar. 2017.

[5] Ö. T. Demir et al., "Foundations of user-centric cell-free massive mimo," *Found. Trends Signal Process.*, vol. 14, nos. 3–4, pp. 162–472, 2021.

[6] E. Björnson and L. Sanguinetti, "Making cell-free massive MIMO competitive with MMSE processing and centralized implementation," *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 77–90, Jan. 2020.

[7] E. Björnson and L. Sanguinetti, "Scalable cell-free massive MIMO systems," *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 4247–4261, Jul. 2020.

[8] Q. Wu, G. Y. Li, W. Chen, D. W. K. Ng, and R. Schober, "An overview of sustainable green 5G networks," *IEEE Wireless Commun.*, vol. 24, no. 4, pp. 72–80, Aug. 2017.

[9] J. Zhang, Y. Wei, E. Björnson, Y. Han, and X. Li, "Spectral and energy efficiency of cell-free massive MIMO systems with hardware impairments," in *Proc. 9th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Oct. 2017, pp. 1–6.

[10] J. Zhang, Y. Wei, E. Björnson, Y. Han, and S. Jin, "Performance analysis and power control of cell-free massive MIMO systems with hardware impairments," *IEEE Access*, vol. 6, pp. 55302–55314, 2018.

[11] J. Zheng, J. Zhang, L. Zhang, X. Zhang, and B. Ai, "Efficient receiver design for uplink cell-free massive MIMO with hardware impairments," *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4537–4541, Apr. 2020.

[12] Z. Mokhtari and R. Dinis, "Sum-rate of cell free massive MIMO systems with power amplifier non-linearity," *IEEE Access*, vol. 9, pp. 141927–141937, 2021.

[13] X. Zhang, D. Guo, K. An, and B. Zhang, "Secure communications over cell-free massive MIMO networks with hardware impairments," *IEEE Syst. J.*, vol. 14, no. 2, pp. 1909–1920, Jun. 2020.

[14] Y. Zhang, Q. Zhang, H. Hu, L. Yang, and H. Zhu, "Cell-free massive MIMO systems with non-ideal hardware: Phase drifts and distortion noise," *IEEE Trans. Veh. Technol.*, vol. 70, no. 11, pp. 11604–11618, Nov. 2021.

[15] X. Hu, C. Zhong, X. Chen, W. Xu, H. Lin, and Z. Zhang, "Cell-free massive MIMO systems with low resolution ADCs," *IEEE Trans. Commun.*, vol. 67, no. 10, pp. 6844–6857, Oct. 2019.

[16] H. Masoumi and M. J. Emadi, "Performance analysis of cell-free massive MIMO system with limited fronthaul capacity and hardware impairments," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 1038–1053, Feb. 2020.

[17] Y. Zhang, M. Zhou, X. Qiao, H. Cao, and L. Yang, "On the performance of cell-free massive MIMO with low-resolution ADCs," *IEEE Access*, vol. 7, pp. 117968–117977, 2019.

[18] R. Zayani, J.-B. Doré, B. Miscopein, and D. Demmer, "Local PAPR-aware precoding for energy-efficient cell-free massive MIMO-OFDM systems," *IEEE Trans. Green Commun. Netw.*, vol. 7, no. 3, pp. 1267–1284, Sep. 2023.

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

[20] R. Price, "A useful theorem for nonlinear devices having Gaussian inputs," *IEEE Trans. Inf. Theory*, vol. IT-4, no. 2, pp. 69–72, Jun. 1958.

[21] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge, U.K.: Cambridge Univ. Press, 2016.

[22] S. M. Kay, *Fundamentals of Statistical Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, 1993.

[23] J. Zhang, J. Zhang, E. Björnson, and B. Ai, "Local partial zero-forcing combining for cell-free massive MIMO systems," *IEEE Trans. Commun.*, vol. 69, no. 12, pp. 8459–8473, Dec. 2021.

ASMA MABROUK received the engineering degree in computer science and network engineering from the National School of Computer Sciences (ENSI), University of Manouba, Tunisia, in 2013, and the Ph.D. degree in computer science from ENSI, in 2020. From 2015 to 2022, she was an Associate Professor with the TEK-UP Engineering School and the ESPRIT Engineering School, Tunisia. Currently, she is a Postdoctoral Research Associate with Commissariat ã l'Energie Atomique et aux Energies Alternatives, CEA-Leti, Grenoble, France. Her research interests include cell-free massive MIMO, hardware impairments, wireless physical layer security, cooperative communications, energy harvesting and simultaneous wireless information, and power transfer.

RAFIK ZAYANI received the engineering, M.Sc., and Ph.D. degrees from Ecole Nationale d'Ingénieurs de Tunis (ENIT), in 2003, 2004, and 2009, respectively, and the Habilitation ã Diriger des Recherches (H.D.R.) degree from CNAM-Paris, in 2020. Since 2005, he has been with the Innov'COM Laboratory, Sup'Com School, Tunisia. Since 2009, he has been an Assistant Professor with the ISI/University of Tunis El Manar, Tunisia. Since 2010, he has also been an Associate Researcher with the CEDRIC Laboratory, Conservatoire National des Arts et Métiers, France. In 2021, he joined Commissariat ã l'Energie Atomique et aux Energies Alternatives, CEA-Leti, Grenoble, France. He is currently an established researcher with long experience in multi-carrier and multi-antenna communications, energy efficiency enhancement by transmitter linearization techniques (baseband DPD) and PAPR reduction, high power amplifier characterization, neural networks, identification modeling and equalization, MIMO, massive MIMO, and cell-free massive MIMO technologies. He was involved in developing enhanced multicarrier waveforms, such as FBMC-OQAM, UFMC, GFDM, BF-OFDM, and WOLA-OFDM. He has contributed to several European (EMPHATIC) and French (ANR-WONG5, POSEIDON, and PEPR-5G) projects that aim at designing flexible air-interfaces for future wireless communications (5G and Beyond). He was awarded a H2020 Marie Sklodowska-Curie Actions Individual Fellowship (MSCA-IF) Grant for his ADAM5 project proposal (2018–2020) that studied hardware-aware and energy-efficient solutions for massive MIMO based 5G systems. He was also awarded a CEA-Tech Science Impulse (SI) Grant for his ALEX6 project proposal that investigated new energy-efficient transceivers design for distributed massive MIMO-based B5G and 6G systems.