



HAL
open science

Sequential hardware-aware precoding for CF-mMIMO-OFDM with serial fronthaul

Antoine Durant, Rafik Zayani, David Demmer

► **To cite this version:**

Antoine Durant, Rafik Zayani, David Demmer. Sequential hardware-aware precoding for CF-mMIMO-OFDM with serial fronthaul. IEEE Wireless Communications and Networking Conference 2024, Apr 2024, Dubaï, United Arab Emirates. 10.1109/WCNC57260.2024.10570629 . cea-04662897

HAL Id: cea-04662897

<https://cea.hal.science/cea-04662897>

Submitted on 26 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Sequential Hardware-aware Precoding for CF-mMIMO-OFDM with serial fronthaul

Antoine Durant, Rafik Zayani, David Demmer
CEA-Leti, Univ. Grenoble Alpes, F-38000 Grenoble, France
{antoine.durant, rafik.zayani, david.demmer}@cea.fr

Abstract—Scalable and cost-effective solutions are of the foremost importance for the deployment of distributed/cell-free massive MIMO (CF-mMIMO) networks. Moreover, in the quest for cheaper hardware, power amplifiers (PAs) are expected to be operated on highly non-linear operation regimes, where strong distortions are impacting the system’s performance. In this paper, we first analyze the impact of the hardware impairment (HWI) caused by the PAs on an orthogonal frequency division multiplexing (OFDM) based CF-mMIMO system. Then, we introduce a novel sequential precoding scheme aiming at cancelling the PA-induced HWI by taking advantage of a serial fronthaul topology. Building on the so-called Bussgang decomposition, we derive a HW-aware precoding, in the form of a closed-form linear system equation. It takes advantage of a serial fronthaul connecting multiple APs, which offers the possibility of communicating HWI information between successive APs in a branch. Numerical results show that the proposed method is able to provide substantial improvement of the achievable spectral efficiency (SE) when the the PA is operated near its saturation region, where its power efficiency is high. Most importantly, it shows promising outcomes to reach high energy efficiency in CF-mMIMO system while maintaining network scalability.

Index Terms—B5G/6G, multi-user MIMO, distributed MIMO, massive MIMO, local precoding

I. INTRODUCTION

The current trends in wireless communications systems, such as cell-densification and co-located mMIMO, are inherently limited by inter-cell interference and a lack of macro spatial-diversity respectively [1]. To alleviate these effects, a network paradigm called Cell-Free massive MIMO (CF-mMIMO) has been investigated and proposed in the recent years [2]. In cell-free networks, the cell borders disappear and each user is simultaneously served by a group of spatially distributed APs. It naturally generates huge amounts of data and signaling to be transferred in back and front-hauls links as user data, channel state information (CSI) and multi-user precoders must be shared among involved APs. The scalability issues for CF-mMIMO have been addressed [3]. To limit the need of AP cooperation, local precoding schemes, where the multi-user precoders are computed locally at each AP with no need for CSI exchange, have been proposed [4] with a limited performance degradation with respect to full AP cooperation [3]. Some network topologies also help to limit inter-AP exchanges like serial-fronthaul and radio stripes [5]. Indeed with a serial connection, a compute-and-forward precoder where cooperation is limited to adjacent APs can be considered. Team-MMSE (minimum mean square estimate) precoding has been proposed to this end [6].

Due to the high number of radio-frequency (RF) elements in massive-MIMO systems, low-cost and energy efficient components are required in order to ensure economical viability. These cheaper alternatives are sources of extra signal distortion known as hardware impairment (HWI), which have been thoroughly studied in the non-serial fronthaul based CF-mMIMO literature [7]. An important part of the HWI is brought in the form of additive distortion noise by the power amplifier (PA) [8]. Two conditions are to be considered for efficient low-cost and low-energy PAs. First, the peak-to-average power ratio (PAPR) needs to be kept low in order to maximize the energy efficiency of the system. Secondly, the distortion, caused by the non-linear compression and the clipping, need to be minimized.

A. Related works

On the one hand, PA-induced HWI has already been extensively studied for traditional co-localized massive MIMO and solutions based on PAPR reduction [9], [10] and digital predistortion (DPD) [11] have been proposed. The extension to CF-MIMO networks is not straightforward because the proposed techniques rely on the massive number of antenna elements available at co-localized mMIMO APs. This criterion is not met in cell-free [2] due to the antennas being distributed over multiple APs and limited AP cooperation. Furthermore, in [12], authors studied local PAPR-aware Precoding solution for CF-mMIMO-OFDM. Although the enhanced performance offered by this method, it is still limited to quite high input values of input back-off (IBO), i.e. moderate energy-efficiency performance. On the other hand, studies focusing on serial fronthaul-based CF-mMIMO have been carried only for ideal hardware [5], [6]. Therefore, to the best of our knowledge, compensation of HWI such as PA-induced distortions in OFDM based CF-mMIMO systems with serial fronthaul is still not addressed in current literature.

B. Contributions

In the context of emerging CF-mMIMO-OFDM serial fronthaul network topologies, we propose a novel hardware-aware precoding solution that takes advantages of unidirectional HWI information sharing. The contributions of this paper are as follows :

- Unlike [7], instead of considering single-carrier based CF-mMIMO, we investigate the impact of HWI caused

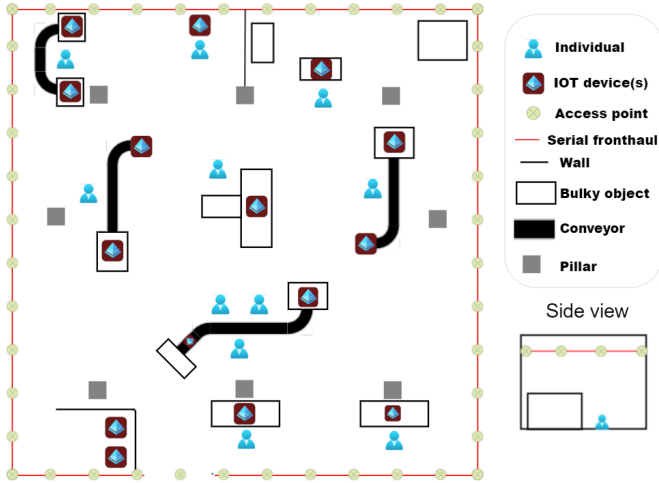


Fig. 1: Schematic of an indoor factory with a CF-mMIMO serial fronthaul deployed on the outer walls: a high number of scatterers and high UE density makes it a rough environment.

by nonlinear PAs on the performance of CF-mMIMO-OFDM, which suffer from high PAPR.

- Building of the Bussgang decomposition, we develop a novel sequential HW-aware precoding scheme adequate to cell-free architecture while taking advantage of a serial fronthaul that connects multiple APs and offers the possibility of communicating HWI information between APs in a successive way. Most importantly, the proposed scheme is fully scalable, since the same quantity of HWI information is required between each two successive APs as opposed to the team MMSE precoding scheme proposed in [6], where the exchanged CSI is getting bulkier along each AP of a branch.
- We derive the HW-aware precoding design, in the form of a closed-form linear system equation, leading to efficient performance assessment in terms of per-user SE under a per-AP power constraint, contrary to the work in [6] where a sum-power constraint is considered which is not practical. Numerical results show an interest of adopting the proposed scheme against a non-compensated hardware impaired system.

II. OFDM-BASED DOWNLINK CF-MMIMO UNDER HWI

A. Network & Scenario

We considered a CF-mMIMO system with L APs connected to a CPU through a single serial connection. Each AP is equipped with M antennas. As this work focuses on the downlink (DL) transmission, we considered K UEs with a single antenna. According to the user-centric CF-mMIMO [2] paradigm, we have $L \gg K$. We consider an indoor/factory scenario, which is adapted to the deployment of a serial fronthaul topology. The APs are placed on the outer edge of a square area and the UEs are randomly distributed within the surface, as depicted on Fig. 1.

B. Channel model

Our system is operating in time-division duplexing (TDD). We assume perfect channel reciprocity between the uplink

and downlink. In addition, thanks to OFDM modulation, we consider a flat frequency response over the frequency slot allocated to any sub-carrier n . The total number of sub-carrier is denoted by N and the system bandwidth by B . We consider a block-fading channel model where a channel is considered static for a duration called the coherence interval. Due to the highly scattered environment and absence of clear line of sight, we consider a frequency-selective Rayleigh fading channel for each coherence interval. It can be modelled as a linear time-varying filter $\mathbf{h}_{l,k}^t$ as follows (1).

$$\begin{aligned} \mathbf{H}_{l,k}^t &= [\mathbf{h}_{l,k}^0, \mathbf{h}_{l,k}^1, \dots, \mathbf{h}_{l,k}^I] \in \mathbb{C}^{M \times I} \\ &= [\sqrt{\beta_{l,k}} \mathbf{g}_{l,k}^0, \sqrt{\beta_{l,k}} \mathbf{g}_{l,k}^1, \dots, \sqrt{\beta_{l,k}} \mathbf{g}_{l,k}^I] \end{aligned} \quad (1)$$

with I and $\mathbf{g}_{l,k}^i \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$ the number and the magnitude of time pulses respectively. The large-scale fading coefficient $\beta_{l,k}$ between AP l and UE k varies slowly along several coherence interval and is assumed to be known from a prior evaluation step. The channel estimates in the frequency domain $\hat{\mathbf{h}}_{l,k,n} \sim \mathcal{CN}(\mathbf{0}, \gamma_{l,k} \mathbf{I}_M)$ with $\gamma_{l,k} \triangleq \mathbb{E}\{|\hat{h}_{l,k}|^2\}$ the mean-square of the estimate for any antenna m , are obtained through an UL training phase using the minimum mean-squared error estimator given in [4]. This estimate is an approximation of the real channel coefficient such that $\mathbf{h}_{l,k,n} = \hat{\mathbf{h}}_{l,k,n} + \tilde{\mathbf{h}}_{l,k,n}$ with $\tilde{\mathbf{h}}_{l,k,n} \sim \mathcal{CN}(\mathbf{0}, \beta_{l,k} - \gamma_{l,k}) \mathbf{I}_M$ the estimation error caused by thermal noise and the HWI.

C. Downlink transmission

The time resource, corresponding to a TDD frame, is split between an uplink (UL) training phase and a DL/UL data transmission. During this interval, a number of OFDM symbols N_c is transmitted over N_{sc} consecutive sub-carriers, and we denote $\tau_c = N_{sc} N_c$ the length in sample of the transmission. However, τ_p samples are reserved to be used as pilots for the training phase, leaving $\tau_d = \tau_c - \tau_p$ samples for the data transmission. The latter are split among the UL and the DL depending on a factor ξ where $0 < \xi < 1$.

We denote the data symbols transmitted through the n -th subcarrier to the K UEs by $\mathbf{s}_n \in \mathbb{C}^{K \times 1}$. They are independent, have unit power such that $\mathbb{E}\{\|\mathbf{s}_n\|^2\} = 1$ and are uncorrelated. For our OFDM system, we consider a guard band of size N_{GB} at both edges of the spectrum where no power is allocated for the data transmit. The sub-carriers are thus divided in two subsets: (i) Ξ , used for data transmission and (ii) its complementary Ξ^c , used for guard-band. Consequently, we have $\mathbf{s}_n = \mathbf{0}_{K \times 1}$ for $n \in \Xi^c$. The data signal transmitted by AP l to all UEs is given by (2).

$$\mathbf{x}_{l,n} = \mathbf{W}_{l,n} \mathbf{P}_l \mathbf{s}_n, \quad (2)$$

with $\mathbf{W}_{l,n} \in \mathbb{C}^{M \times K}$ the precoding matrix for sub-carrier n at AP l and $\mathbf{P}_l \in \mathbb{C}^{K \times K}$ a diagonal matrix whose element $\sqrt{\eta_{l,k}}$, $k = 1, \dots, K$ are the normalized DL transmit power with regard to the noise power and satisfies a per-AP constraint ρ_{max}^{DL} following the power control policy in [3].

For the precoder, we use the common CF-mMIMO local normalized full-pilot zero-forcing (FZF) scheme [4] which is

efficient at removing the inter-user interference while maintaining a fully distributed and scalable nature as no information sharing between AP is required. It is given by

$$\mathbf{w}_{l,k,n} = \hat{\mathbf{H}}_{l,n} \left(\hat{\mathbf{H}}_{l,n}^H \hat{\mathbf{H}}_{l,n} \right)^{-1} \mathbf{e}_{i_k} \sqrt{(M - \tau_p) \theta_{l,k}} \quad (3)$$

with \mathbf{e}_{i_k} the i -th column of \mathbf{I}_{τ_p} and $\theta_{l,k} = \gamma_{l,k} / c_{l,k}^2$ as analytically derived in [4]. It is worth noting that the normalization factor $\sqrt{(M - \tau_p) \theta_{l,k}}$ limits the set of spatially independent target (UE) to $M - 1$.

D. Hardware impairment

Each antenna element m of an AP receives the OFDM-modulated signal for the DL transmission through an RF chain. Due to the focus of this work, we reduce the model of the RF chain to the effect of the PA. We describe the effect of an arbitrary PA for different operation regimes through the input back-off (IBO) parameter. It is the difference (in dB) between the carrier input power at the operating point and the point of saturation of the device. Consequently, a low IBO leads to higher energy efficiency but increases the level of non-linear distortion. We denote by $\mathbf{a}_{l,m}^t$ the OFDM signal corresponding to the inverse fast-fourier transform (IFFT) of the DL precoded symbols, as in (2), and $\mathbf{z}_{l,m}^t$ the output of the DAC. The non-linearities encountered in a PA can be described through two types of conversion: amplitude to amplitude (AM/AM) and amplitude to phase (AM/PM). Hence, the transfer function representing the nonlinear operation of the PA $\mathbf{f}(\cdot)$, assumed to be identical over all antennas and memory-less, can be written as in (4).

$$\mathbf{z}_{l,m}^t = \mathbf{f}(\mathbf{a}_{l,m}^t) = \Omega(\alpha\rho) \exp(j\psi + \Psi(\alpha\rho)) \quad (4)$$

where Ω and Ψ denotes the denote the AM/AM and AM/PM conversions respectively, ρ and ψ are the magnitude and the phase of the PA's input sample and α is the multiplicative coefficient that is applied to the input of the PA to achieve the desired PA operating point based on the specified IBO. It is defined as $\alpha = \frac{V_{sat}}{G\sqrt{p_{in}}}$ with V_{sat} and G being the saturation voltage and the gain of the PA respectively, while p_{in} is the average power of the input signal. To accurately model the behavior of NL-PAs, we use the memoryless modified Rapp model given by [13] and described as in (5).

$$\Omega(\rho) = \frac{G\rho}{\left(1 + \left|\frac{G\rho}{V_{sat}}\right|^{2\rho}\right)^{\frac{1}{2\rho}}}, \quad \Psi(\rho) = \frac{A\rho^q}{\left(1 + \left(\frac{\rho}{B}\right)^q\right)} \quad (5)$$

with ρ the smoothness factor and A , B and q being fitting parameters. From the Bussgang theorem [14], we know that the output of a non-linear function (4) can be expressed as a unique decomposition given by (6).

$$\mathbf{z}_{l,m}^t = K_0 \mathbf{a}_{l,m}^t + \mathbf{d}_{l,m}^t, \quad (6)$$

with K_0 a scaled version of the input signal and $\mathbf{d}_{l,m}^t$ an uncorrelated zero mean noise with a variance σ_d^2 . While it isn't Gaussian at the receiver (UE) side, after OFDM demodulation, the distortion can be considered as a Gaussian distribution

among all the N sub-carriers. While it is not in the scope of this paper, it is worth noting that the HWI parameters K_0 and σ_d^2 can be obtained analytically for the PA to develop closed-form expressions and evaluate the in-band performance as in [12].

E. Performance Indicator

The performance is evaluated at the user side as the downlink spectral efficiency. The frequency-domain received signal at UE k and subcarrier n can be given by (7).

$$\begin{aligned} y_{k,n} &= \sum_{l=1}^L \mathbf{h}_{l,k,n}^H \mathbf{z}_{l,n} + b_{k,n} \\ &= \sum_{l=1}^L \mathbf{h}_{l,k,n}^H \mathbf{K}_0 \mathbf{x}_{l,n} + \sum_{l=1}^L \mathbf{h}_{l,k,n}^H \mathbf{d}_{l,n} + b_{k,n}, \end{aligned} \quad (7)$$

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 $b_{k,n} \sim \mathcal{CN}(0, 1)$ is an i.i.d. Gaussian noise.

By plugging (2) in (7), $y_{k,n}$ can be expanded as

$$\begin{aligned} y_{k,n} &= \sum_{l=1}^L \mathbf{h}_{l,k,n}^H \mathbf{K}_0 \mathbf{W}_{l,n} \mathbf{P}_l \mathbf{s}_n + \sum_{l=1}^L \mathbf{h}_{l,k,n}^H \mathbf{d}_{l,n} + b_{k,n} \quad (8) \\ &= \underbrace{\sum_{l=1}^L \sqrt{\eta_{l,k}} \mathbf{h}_{l,k,n}^H \mathbf{K}_0 \mathbf{w}_{l,n,k} s_{k,n}}_{\text{Desired signal}} + \underbrace{b_{k,n}}_{\text{Noise}} + \\ &\quad \underbrace{\sum_{l=1}^L \sum_{t \neq k}^K \sqrt{\eta_{l,t}} \mathbf{h}_{l,k,n}^H \mathbf{K}_0 \mathbf{w}_{l,n,t} s_{t,n}}_{\text{Multi-user interference}} + \underbrace{\sum_{l=1}^L \mathbf{h}_{l,k,n}^H \mathbf{d}_{l,n}}_{\text{HWI}} \quad (9) \end{aligned}$$

The per-user SE in the downlink can be computed [4] as

$$\text{SE}_k = \xi \left(1 - \frac{\tau_p}{\tau_c} \right) \log_2 (1 + \text{SINR}_{k,n}), \quad (10)$$

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

$$\begin{aligned} \text{SINR}_{k,n} &= \frac{|\text{CP}_{k,n}|^2}{\text{E}\{|\text{PU}_{k,n}|^2\} + \sum_{t \neq k}^K \text{E}\{|\text{UI}_{k,t,n}|^2\} + \text{E}\{|\text{HWI}_{k,n}|^2\} + 1} \quad (11) \end{aligned}$$

where $\text{CP}_{k,n}$, $\text{PU}_{k,n}$, $\text{UI}_{k,n}$ and $\text{HWI}_{k,n}$ are the coherent precoding, the precoding gain uncertainty, the multi-user interference and the hardware impairment respectively, and are given by (12, 13, 14, 15).

$$\text{CP}_{k,n} = \sum_{l=1}^L \sqrt{\eta_{l,k}} \text{E}\{\mathbf{h}_{l,k,n}^H \mathbf{K}_0 \mathbf{w}_{l,n,i_k}\} \quad (12)$$

$$\text{PU}_{k,n} = \sum_{l=1}^L \left(\sqrt{\eta_{l,k}} \mathbf{h}_{l,k,n}^H \mathbf{K}_0 \mathbf{w}_{l,n,i_k} - \sqrt{\eta_{l,k}} \text{E}\{\mathbf{h}_{l,k,n}^H \mathbf{K}_0 \mathbf{w}_{l,n,i_k}\} \right) \quad (13)$$

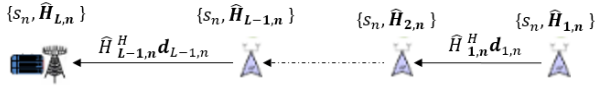


Fig. 2: Serial fronthaul: L APs interconnected in series with sequential hardware aware precoding

$$\text{UI}_{k,t,n} = \sum_{l=1}^L \sqrt{\eta_{l,k}} \mathbf{h}_{l,k,n}^H \mathbf{K}_0 \mathbf{w}_{l,n,i_t} \quad (14)$$

$$\text{HWI}_{k,n} = \sum_{l=1}^L \mathbf{h}_{l,k,n}^H \mathbf{d}_{l,n} \quad (15)$$

with i_k and i_t the indices of the pilot sequence used by UE k and t . It corresponds to the column vector in the precoder matrix designed for their training sequences.

III. SEQUENTIAL HARDWARE-AWARE PRECODING

In this section, we describe the proposed sequential hardware-aware precoding scheme. We consider a serially connected network of APs as depicted in Fig.1. Each AP obtains its local channel estimate thanks to the uplink training phase. In the meantime, the hardware-related distortion $\{\mathbf{d}_{l,n}, \forall n\}$ that will be generated when transmitting the precoded signals via the RF chains approximated at each AP l using the impairment model (4). The goal of this method is to maintain the constructive addition of the contribution of each AP while canceling out the HWI received by each user on each subcarrier. It is done by sharing the local CSI and the estimated distortion noise to the subsequent AP as depicted on Fig. 2.

For the first AP of the branch (AP 1), and only for this one, the transmitted signal is similar as in (2). The local FZF precoder (3) is used as the precoding filter $\mathbf{W}_{l,n}$. At the K UEs, the hardware-related distortions from the contribution of AP 1 can be computed as $\{\mathbf{H}_{1,n}^H \mathbf{d}_{1,n}, \forall n\}$. This latter ($\mathbf{d}_{1,n}$) represents the HWI information at AP 1 that will be transmitted to AP 2. For the rest of the network branch, when AP l receives HWI from its preceding AP ($l-1$), along the serial fronthaul, it exploits the received information together with its own local channel estimates $\{\hat{\mathbf{H}}_{l,n}^H, \forall n\}$ to generate its local precoding filters. The optimal frequency-domain precoded signals $\mathbf{x}_{l,n}$, to be transmitted from the AP l on subcarrier n to the K UEs, are given by the unique solution in (16).

$$\mathbf{x}_{l,n} = \mathbf{W}_{l,n} \mathbf{P}_l \mathbf{s}_n - \mathbf{W}_{l,n} \mathbf{C}'_l \hat{\mathbf{H}}_{l-1,n}^H \mathbf{d}_{l-1,n} \quad (16)$$

with $\mathbf{C}'_l \in \mathbb{C}^{K \times K}$ a normalization matrix satisfying $\{E\{\|\mathbf{H}_{l,n}^H \mathbf{W}_{l,n}\|^2\} = \mathbf{I}_K, \forall k\}$. It can be optimized in an optimal diagonal matrix whose elements are equal to $\frac{1}{\sqrt{(M-K)\beta_{l,k}}}$.

The process is repeated sequentially along the serial fronthaul, until the CPU is reached. The hardware distortions, noted by $\{\mathbf{H}_{l,n}^H \mathbf{d}_{l,n}, \forall n\}$, represent the effective distortions received by the K served users after the frequency-domain precoded signals $\{\mathbf{x}_{l,n}, \forall n\}$ are created, modulated, passed through the RF chains of AP l and transmitted via the UEs channels. One

advantage of the method is the limited signaling it requires: only MN complex samples are transmitted to the subsequent AP. Interestingly, The amount of signaling data, ie. HWI information remain constant along all the serial fronthaul.

The frequency-domain received signal at UE k and subcarrier n is be given by (17). Here, only residual HWI remains, hence, the effective SINR of UE k at subcarrier n remains as in (11), but only residual HWI are left after compensation, which become as defined by (18).

$$\text{HWI}_{k,n} = \sum_{l=1}^{L-1} \left[\mathbf{h}_{l,k,n}^H - \mathbf{h}_{l+1,k,n}^H \mathbf{W}_{l+1,n} \hat{\mathbf{H}}_{l,n}^H \right] \mathbf{d}_{l,n} + \mathbf{h}_{L,k,n}^H \mathbf{d}_{L,n} \quad (18)$$

The last term $\mathbf{h}_{L,k,n}^H \mathbf{d}_{L,n}$ being the uncompensated distortions brought by the last AP in the branch. It can be minimized by using a base station with sufficient resources and PAs operated with a high IBO as the last AP.

IV. SIMULATION AND EXPERIMENTAL RESULTS

A. Simulation parameters

We consider a scenario with a serial fronthaul network wrapped around a 60 by 60 meter square environment as in Fig. 1. The large-scale fading coefficients $\beta_{l,k}$ are modeled as in [4] by (19).

$$\beta_{l,k} = \text{PL}_{l,k} \cdot 10^{\frac{\sigma_{sh} z_{l,k}}{10}} \quad (19)$$

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 path-loss model, which is given by equation (20) when assuming a sub 6 GHz carrier frequency [4].

$$\text{PL}_{l,k} [\text{dB}] = -22.7 - 36.7 \log_{10} \left(\frac{d_{l,k}}{1 \text{ m}} \right) - 26 \log_{10} \left(\frac{f_c}{1 \text{ GHz}} \right) \quad (20)$$

where f_c is the carrier frequency and $d_{l,k}$ denotes the distance between the l -th AP and the k -th UE including AP and UE's heights. The per-AP DL transmit power is capped by the IBO such that $\rho_{max}^{DL} = M(\text{V}_{\text{sat}}/G)10^{\text{IBO}/10}$. The simulation settings are reported in Table I. The UL and noise power are used for the UL training described in [4], and a number R of independent channel realisations and transmission sequences are performed per snapshot before performance evaluation.

TABLE I: Simulation settings.

Description	Value	Description	Value
D (simulation area)	60 × 60m	UE distribution	random
Noise power	-93 dBm	AP/UE antenna height	10/1.5 m
UL transmit power	20 dBm	V _{sat}	1.9 V
τ _c	168	G	16
ξ	0.5	σ _{sh}	4 dB
Carrier frequency	3.5 GHz	Bandwidth	20 MHz
N	256	R	25

B. Results

The cumulative density functions (CDF) of the downlink SE for the non-impaired system (red), the hardware impaired (black) and the hardware-impaired with our proposed HWI-aware precoding scheme systems (blue) are shown in Fig. 3

$$\begin{aligned}
y_{k,n} = & \sum_{l=1}^L \mathbf{h}_{l,k,n}^H \mathbf{K}_0 \mathbf{x}_{l,n} + \sum_{l=1}^L \mathbf{h}_{l,k,n}^H \mathbf{d}_{l,n} + b_{k,n} \underbrace{\sum_{l=1}^L \sqrt{\eta_{l,k}} \mathbf{h}_{l,k,n}^H \mathbf{K}_0 \mathbf{w}_{l,n,k} s_{k,n}}_{\text{Desired signal}} + \underbrace{\sum_{l=1}^L \sum_{t \neq k}^K \sqrt{\eta_{l,t}} \mathbf{h}_{l,k,n}^H \mathbf{K}_0 \mathbf{w}_{l,n,t} s_{t,n}}_{\text{Multi-user interference}} \\
& + \underbrace{\sum_{l=1}^{L-1} \left[\mathbf{h}_{l,k,n}^H - \mathbf{h}_{l+1,k,n}^H \mathbf{W}_{l+1,n} \hat{\mathbf{H}}_{l,n} \right] \mathbf{d}_{l,n} + \mathbf{h}_{L,k,n}^H \mathbf{d}_{L,n}}_{\text{Residual HWI}} + \underbrace{b_{k,n}}_{\text{Noise}}
\end{aligned} \tag{17}$$

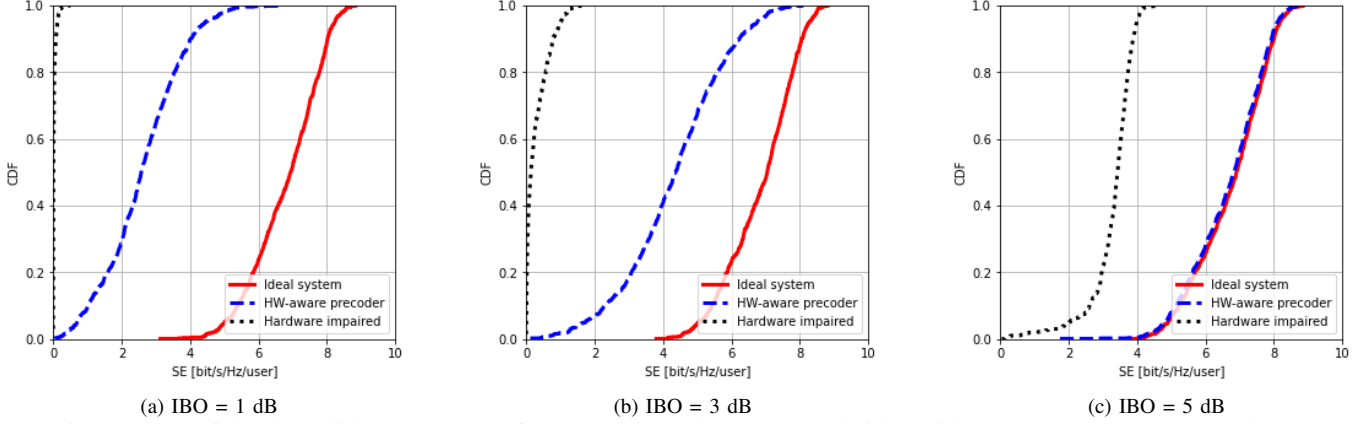


Fig. 3: CDF of the downlink per-user SE for three IBO values (1, 3 and 5dB) with $L = 48$, $K=7$, $\tau_p = 7$ and $M=8$

for three levels of IBO. First of all, it seems worth reminding that we consider an OFDM system with high PAPR. This means that we need to operate the power amplifier with a high IBO to avoid the PA-induced HWI. However, increasing the IBO severely limits the energy efficiency. One can observe the performance degradation induced by the presence of the HWI by comparing the red and black curves in Fig. 3. When the IBO is very low, such as 1 dB, the power of the HWI is too high (compared to the power of the useful signal) and severely degrades the signal quality and the capacity of the large majority of users tends to zero. Increasing the IBO up to 5 dB, helps limiting the distortion induced by the HWI, however the channel capacity is still limited with respect to the ideal case (*i.e.* linear case): half of the ideal capacity for the top 5% of the best channels and less than 2.0 bits/s/Hz for the top 5% of the worst channels. One can observe that our proposed precoding technique, depicted with the blue curve in Fig. 3, greatly helps reducing the effects of the HWIs. Indeed, coupled with an IBO of 5 dB, it is sufficient to reach the performance of a linear system. It represents a huge gain in energy efficiency of about 5 dB. In order to clearly see the effective of our proposed method in mitigating HWI and without loss of generality, results in Fig. 3 are performed with no pilot contamination (*i.e.*, $K = \tau_p$). Then, the observed performance degradation is due to the presence of HWI (noisy CSI is considered for both linear and HWI cases).

Fig. 4 illustrates the impact of the AP density for a fixed IBO. For this figure, we consider the median SE over 750 snapshots. One can observe that the median SE for

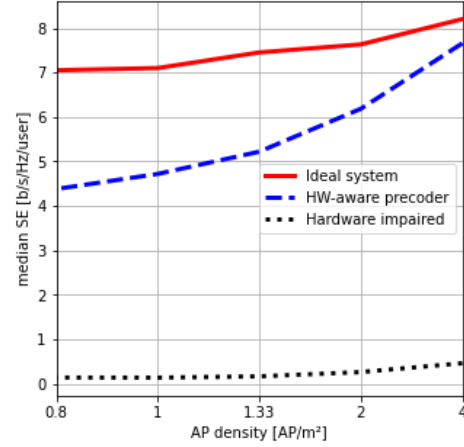


Fig. 4: Median SE against AP density with an IBO of 3dB

the linear case increases with the number of APs. For the non-compensated HWI case, increasing the number of APs barely increase the performance and the median SE is almost null. This result is expected because the proportion of good channels increases with the number of APs but the HWI term now prevails in the signal in (11). However, with the HW-aware precoder, the increased AP density leads to more frequent HWI cancellation, while guarantying a high SE floor even at a reasonable network density (0.8 AP/m²). In the best cases, our proposed solution tends to the ideal, non-hardware impaired equivalent when the AP density is extreme: less than 1 bits/s/Hz/UE of difference with at least 4 APs/m².

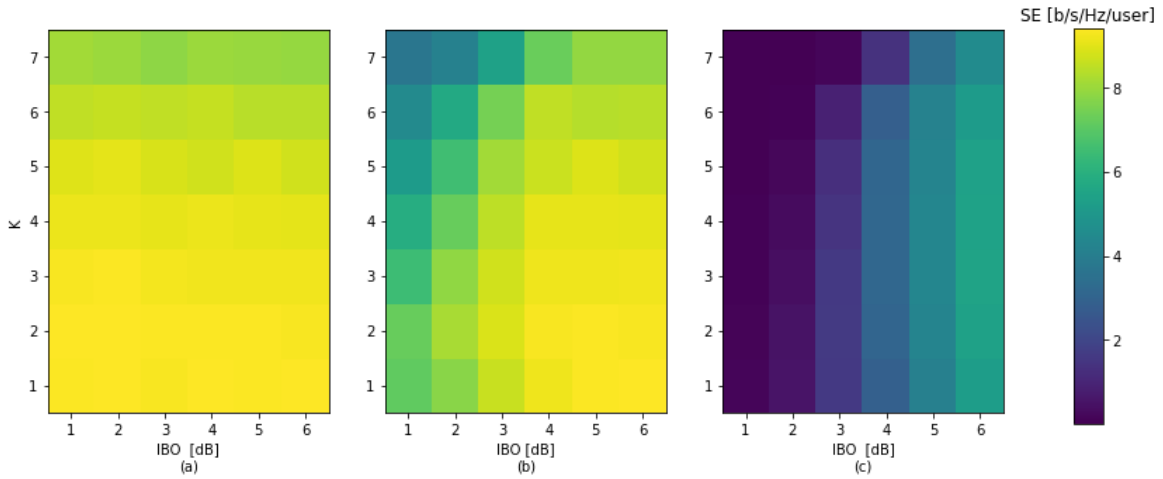


Fig. 5: 2D Color map of the achieved median SE against number of users and IBO for the linear, hardware aware precoding-based correction and the hardware impaired case. (a) Ideal system (b) HWI system with HW-aware precoder (c) HWI system

Last but not least, it is important to assess the performance of the proposed precoding scheme against the degrees of freedom (DoF) available for the beamforming. Fig. 5 presents the results of the performance evaluation of the per-UE mean SE for a varying number of UE K without pilot contamination $K = \tau_p$. The SE is proportional to the IBO as expected from the results in Fig. 3. However, due to the limited per-UE DoF available when $K \rightarrow M$, we observe that the HW-aware precoder (b) exhibits reduced performance as compared to the ideal (a), with a 65% decrease in performance for a IBO of 1 dB and $K = 7$ while the decrease is only of 33% for $K = 1$. It can be expected that the FZF pre-filter is responsible for this degradation due to a reduced array gain when the DoF are low, as seen from the similar behavior of the non-compensated equivalent (c).

CONCLUSION

In this paper, the impact of PA-induced HWI on a CF-mMIMO OFDM system has been evaluated. For the most critically impaired case, the SE exhibits a more than 95% decrease. Exploiting a serial fronthaul topology, a sequential hardware-aware precoding has been proposed where the HWI information are extracted and transmitted from AP to AP in a successive, unidirectional manner. Numerical results showed that the proposed HWI cancellation algorithm leads to a 50% increase of the median SE when the PA is operated at an IBO of 5dB, providing performance comparable to an ideal (linear) PA. In the meantime, at least 35% of the SE of the ideal system is obtained in the worst hardware conditions (IBO of 1dB).

ACKNOWLEDGMENT

This work was supported in part by the ANR under the France 2030 program, grant "NF-PERSEUS: ANR-22-PEFT-0004" and by the framework of the ALEX6 Project receiving fund by the French CARNOT-ANR under Carnot-SCIENCE-ALEX6.

REFERENCES

- [1] A. Lozano, R. W. Heath, and J. G. Andrews, "Fundamental limits of cooperation," *IEEE Transactions on Information Theory*, vol. 59, no. 9, pp. 5213–5226, 2013.
- [2] S. Chen, J. Zhang, J. Zhang, E. Björnson, and b. ai, "A survey on user-centric cell-free massive MIMO systems," *Digital Communications and Networks*, vol. 8, 12 2021.
- [3] E. Björnson and L. Sanguinetti, "Scalable Cell-Free Massive MIMO Systems," *IEEE Transactions on Communications*, vol. 68, no. 7, pp. 4247–4261, 2020.
- [4] G. Interdonato, M. Karlsson, E. Björnson, and E. G. Larsson, "Local Partial Zero-Forcing Precoding for Cell-Free Massive MIMO," *IEEE Transactions on Wireless Communications*, vol. 19, no. 7, pp. 4758–4774, 2020.
- [5] Z. H. Shaik, E. Björnson, and E. G. Larsson, "Cell-Free Massive MIMO with Radio Stripes and Sequential Uplink Processing," in *Proc. IEEE International Conference on Communications Workshops (ICC Workshops)*, 2020, pp. 1–6.
- [6] L. Miretti, E. Björnson, and D. Gesbert, "Team MMSE Precoding With Applications to Cell-Free Massive MIMO," *IEEE Transactions on Wireless Communications*, vol. 21, no. 8, pp. 6242–6255, 2022.
- [7] 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. 55 302–55 314, 2018.
- [8] Z. Mokhtari and R. Dinis, "Sum-rate of cell free massive mimo systems with power amplifier non-linearity," *IEEE Access*, vol. 9, pp. 141 927–141 937, 2021.
- [9] D. Wulich and L. Goldfeld, "Reduction of peak factor in orthogonal multicarrier modulation by amplitude limiting and coding," *IEEE Transactions on Communications*, vol. 47, no. 1, pp. 18–21, 1999.
- [10] R. Zayani, H. Shaiek, and D. Roviras, "PAPR-Aware Massive MIMO-OFDM Downlink," *IEEE Access*, vol. 7, pp. 25 474–25 484, 2019.
- [11] H. W. Kang, Y. S. Cho, and D. H. Youn, "On compensating nonlinear distortions of an OFDM system using an efficient adaptive predistorter," *IEEE Transactions on Communications*, vol. 47, no. 4, pp. 522–526, 1999.
- [12] R. Zayani, J.-B. Doré, B. Miscopein, and D. Demmer, "Local PAPR-Aware Precoding for Energy-Efficient Cell-Free Massive MIMO-OFDM Systems," *IEEE Transactions on Green Communications and Networking*, vol. 7, no. 3, pp. 1267–1284, 2023.
- [13] Nokia, "Realistic Power Amplifier Model for the New Radio Evaluation, document R4-163314," *3GPP TSG-RAN WG4 Meeting 79*, 2016.
- [14] O. T. Demir and E. Björnson, "The Busgang Decomposition of Non-linear Systems: Basic Theory and MIMO Extensions [Lecture Notes]," *IEEE Signal Processing Magazine*, vol. 38, no. 1, pp. 131–136, 2021.