

Iterative AP Selection, MMSE Precoding and Power Allocation in Cell-Free Massive MIMO Systems

Victoria M. T. Palhares^{1*}, Rodrigo C. de Lamare^{1,2}, André R. Flores¹, Lukas T. N. Landau¹

¹ Center for Telecommunications Studies, Pontifical Catholic University of Rio de Janeiro, Rio de Janeiro, Brazil

² Department of Electronic Engineering, University of York, York, U.K.

* E-mail: victoriapalhares@cetuc.puc-rio.br

Abstract: In this work, we propose iterative access point (AP) selection (APS), linear minimum mean-square error (MMSE) precoding and power allocation techniques for Cell-Free Massive multiple-input multiple-output (MIMO) systems. We consider the downlink channel with single-antenna users and multiple-antenna APs. We derive sum-rate expressions for the proposed iterative APS techniques followed by MMSE precoding and optimal, adaptive, and uniform power allocation schemes. Simulations show that the proposed approach outperforms existing conjugate beamforming (CB) and zero-forcing (ZF) schemes and that performance remains excellent with APS, in the presence of perfect and imperfect channel state information (CSI).

1 Introduction

In the fifth generation (5G) of wireless communications, massive multiple-input multiple-output (MIMO) systems have been used to improve the reliability, efficiency and throughput of wireless networks. In these systems, multiple antennas simultaneously serve multiple users in the same time-frequency resource [1]. In 5G, distributed antenna systems (DAS) are being considered in order to provide higher coverage probability, flexible resource management, higher power efficiency and larger capacities by the exploitation of smaller distances between base stations and users, and of spatial diversity [2]. The combination of Massive MIMO and DAS systems has been exploited in the past, where it has been proven that distributed setups can provide higher rates than collocated antenna systems (CAS) [3]. In [4], it is shown that large-scale DAS (L-DAS) systems are energy-efficient and have simpler precoding and power control, when associated with antenna selection and user clustering methods.

As an evolution of Massive MIMO and DAS, in Cell-Free Massive MIMO, many randomly distributed access points (APs) are connected to a central processing unit (CPU) and serve simultaneously a much smaller number of users. At the CPU, precoding techniques and power allocation algorithms can be performed. The main goal is to use advanced backhaul in order to provide uniformly good service for all users [5]. Compared to a cellular system, cell-free concepts have been shown to increase energy efficiency and per-user throughput, in rural and urban scenarios [6]. Moreover, they can have simple signal processing which facilitates the exploitation of phenomena such as favorable propagation with channel hardening [7]. Nevertheless, channel hardening can be achieved in specific cell-free scenarios [8]. If multiple-antenna APs are considered, channel hardening can be experienced at the cost of losses in macro-diversity. In a single-antenna APs scenario, increasing AP density does not lead to more channel hardening. An exception is made if a small path loss component is assumed, increasing the chances of this phenomenon with the increment in AP density.

Many precoding and receive processing schemes have been previously studied for cell-free networks. In the uplink, a matched filter (MF) has been exploited in [9], whereas in [10, 11], a minimum mean-square error (MMSE) and large-scale fading decoding (LSFD) receivers have demonstrated to provide higher outage rate than the former. Moreover, in [12], high spectral efficiency was achieved through an uplink framework combined with local partial MMSE and maximum ratio (MR) combining. On the other hand, in the

downlink, a conjugate beamforming (CB) precoder has been studied for computational simple signal processing, at the cost of lower performance [9, 13, 14]. With more backhaul requirements, a zero-forcing (ZF) precoding design has been extensively studied in order to improve system performance [13] and maximize the energy efficiency [15]. In [16], two distributed precoding schemes based on the ZF criterion have been developed, which are called local partial ZF and local protective partial ZF. In [17], an MMSE-type precoder was obtained through an uplink-downlink duality based on the MMSE combiner from [10]. The technique was combined with a clustering method based on pilot assignment and a heuristic solution for power allocation. Moreover, in a scenario which considers cell-free networks and non-orthogonal multiple access (NOMA), precoders based on MR transmission, full pilot ZF precoding and modified regularized ZF precoding [18] have been evaluated.

Since uniformly good service for all users is a key point of cell-free systems, many precoding and receive processing techniques have been combined with power control algorithms. Different criteria have been used in the literature to satisfy different aspects of this network concept. In [19], power allocation is performed on the uplink to maximize the minimum user rate under per-user power constraints. In [9], the MF and CB receiver and precoder, respectively, are combined with an optimization that maximizes the smallest of all user rates under per-AP constraints. In [13], a similar procedure is performed, where CB and ZF precoders are combined with the maximization of the minimum signal-to-interference-plus-noise ratio (SINR) from all users also under per-AP power constraints. With a different criterion, ZF precoders were also combined with an energy efficiency maximization, under a per-AP power constraint and a per-user spectral efficiency constraint [14] considering backhaul power consumption and imperfect channel state information (CSI) [15]. In [20], power control algorithms focused on maximizing the sum-rate or the minimum rate are developed and combined with a ZF precoder for cell-free systems.

To decrease the backhaul power consumption in cell-free systems, some works have proposed AP selection (APS), where each user is served by a subset of APs. In [21], two APS schemes have been proposed, one based on the received power and the other on the largest large-scale-fading coefficients. Similarly, a sequential approach to connect each user only to certain APs based on the channel gain and the channel quality has been proposed in [22]. Most works addressing Cell-Free Massive MIMO systems consider single-antenna APs. However, in [8], multiple-antenna APs have been analysed in order to improve channel hardening and to increase the likelihood that

favorable propagation occurs. In the same way, [23] explores the downlink spectral efficiency in a scenario with multiple-antenna APs and users. Additionally, in [24] it is shown that, in terms of costs, it is better to add more antennas to an AP than to install more APs.

In this work, we propose iterative APS, linear MMSE precoding and power allocation techniques. Specifically, APS, precoding and power allocation schemes are presented, where the precoder is calculated based on initial parameters, used in power allocation and recalculated based on power allocation coefficients. MMSE channel estimates are considered, similarly to previous works [13]. Furthermore, optimal, adaptive and uniform power allocation techniques are devised along optimal and suboptimal APS, and compared with existing CB and ZF precoding and power allocation techniques [9, 13] in terms of sum-rate, minimum SINR and bit error rate (BER), taking into account perfect and imperfect CSI. Moreover, analytical expressions are derived to compute the achievable rates of the proposed approaches. Differently from [25] which considers an asymptotic rate analysis in the number of antennas and users, we consider them to be finite.

In summary, the main contributions of this work are:

- Optimal APS scheme based on exhaustive search (ES) and a sub-optimal solution based on the large-scale fading coefficients applied to the obtained channel parameters. In contrast with previous works, the optimal scenario has not been explored to date, and the proposed APS schemes have not been combined with the studied precoding and power allocation strategies.
- The complete derivation of an iterative linear MMSE precoder for Cell-Free Massive MIMO systems, which takes into account a power allocation matrix, unlike existing approaches [17, 26].
- Optimal and uniform power allocation techniques applied to the proposed scheme in order to maximize the minimum SINR, differently from [17] where the MMSE precoder is combined with a heuristic power allocation approach.
- Adaptive power allocation based on the Stochastic Gradient (SG) with the mean-square error (MSE) criterion, differently from existing power control solutions.
- An analytical expression of the achievable rate, specifically for the proposed MMSE precoder.
- An analysis of the computational complexity of the proposed and existing schemes.
- A simulation study of the proposed and existing techniques in terms of sum-rate and minimum SINR. Unlike existing results from the literature, we illustrate the performance of the proposed system in terms of BER as well.

The rest of this paper is organized as follows. In Section II the Cell-Free Massive MIMO system model and CSI scenarios are detailed. In Section III, an iterative APS, combined with MMSE precoder and power allocation is presented. Sum-rate performance and the computational cost are evaluated in Section IV. In Section V, numerical results and discussions are presented, whereas in Section VI conclusions are drawn.

Notation: Uppercase and bold symbols denote matrices and vectors respectively. The superscripts $()^*$, $()^T$, $()^H$ stand for complex conjugate, transpose and Hermitian operations, respectively. The expectation, trace of a matrix, real part of the argument, Euclidean norm and Frobenius norm are denoted by $\mathbb{E}[\cdot]$, $\text{tr}(\cdot)$, $\text{Re}(\cdot)$, $\|\cdot\|_2$ and $\|\cdot\|_F$, respectively. The Hadamard product is expressed by \odot . The operator $\text{diag}\{\mathbf{A}\}$ retains the main diagonal elements of \mathbf{A} in a column vector. The $D \times D$ identity matrix is \mathbf{I}_D . The notation $x \sim \mathcal{N}(0, \sigma^2)$ refers to a Gaussian random variable (RV) x with zero mean and variance σ^2 and $x \sim \mathcal{CN}(0, \sigma^2)$ denotes a circularly symmetric complex Gaussian RV x with zero mean and variance σ^2 .

Table 1 summarizes the most important parameters of this work.

Table 1 Parameters

| | |
|----------------------------------|--|
| L | number of APs |
| N | number of antennas per AP |
| M | total number of antennas |
| S | number of selected APs |
| K | number of users |
| \mathbf{G} | channel matrix where $[\mathbf{G}]_{m,k} = g_{m,k}$ |
| $\hat{\mathbf{G}}$ | channel estimation matrix where $[\hat{\mathbf{G}}]_{m,k} = \hat{g}_{m,k}$ |
| $\tilde{\mathbf{G}}$ | channel estimation error matrix where $[\tilde{\mathbf{G}}]_{m,k} = \tilde{g}_{m,k}$ |
| $\beta_{m,k}$ | large-scale fading coefficient between the m th antenna and the k th user |
| $h_{m,k}$ | small-scale fading coefficient between the m th antenna and the k th user |
| $\alpha_{m,k}$ | variance of $g_{m,k}$ |
| $n_{m,k}$ | CSI accuracy for the m th antenna and the k th user |
| ρ_f | maximum transmit power of each antenna |
| \mathbf{s} | symbol vector |
| \mathbf{w} | additive noise vector |
| \mathbf{Q} | APS matrix, where $[\mathbf{Q}]_{m,k} = q_{m,k}$ and $q_{m,k} \in \{0, 1\}$ |
| SINR_k | SINR of the k th user |
| A_1 | desired signal |
| $A_{2,i}$ | interference caused by user i for $i \neq k, i = 1, \dots, K$ |
| A_3 | CSI error |
| $()$ | variable after APS |
| \mathbf{P}_{MMSE} | MMSE precoder |
| \mathbf{N} | intermediate power allocation matrix |
| $\hat{\mathbf{N}}_{\text{MMSE}}$ | final power allocation matrix |
| η_k | power coefficient of the k th user |
| f_{MMSE} | normalization factor for the MMSE precoder |
| \mathbf{C}_s | symbol covariance matrix |
| \mathbf{C}_w | noise covariance matrix |
| $T_{\text{APA}}, T_{\text{OPA}}$ | number of iterations for APA and OPA algorithm, respectively |

2 System Model

The downlink of a Cell-Free Massive MIMO system is considered with L randomly distributed APs equipped with N antenna elements each and K single-antenna users, where $LN = M$ is the total number of antenna elements and $M \gg K$. In this system, all APs are connected to a CPU and serve simultaneously all users, as shown in Fig. 1. Each AP obtains CSI and sends them to the CPU, which performs APS, precoding and power allocation whose parameters are then fed back to the APs.

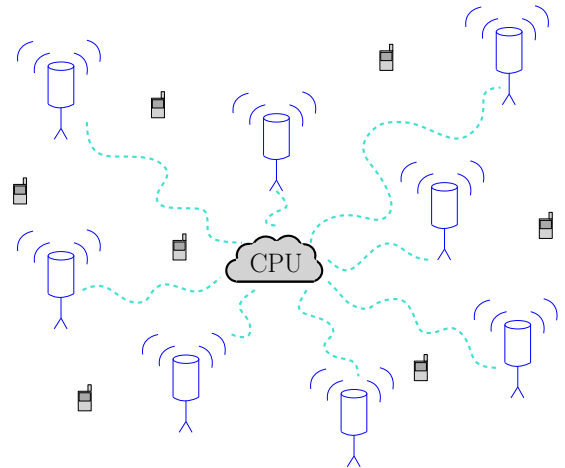


Fig. 1: Cell-Free Massive MIMO System.

The channel coefficients between the m th antenna element and the k th user are defined as [9]

$$g_{m,k} = \sqrt{\beta_{m,k}} h_{m,k}, \quad (1)$$

where $\beta_{m,k}$ is the large-scale fading (LS) coefficient (path loss and shadowing effects) and $h_{m,k} \sim \mathcal{CN}(0, 1)$ is the small-scale fading coefficient, defined as independent and identically distributed (i.i.d) RVs that remain constant during a coherence interval and are independent over different coherence intervals. However, the LS coefficients change less frequently, being constant for several coherence intervals. In this case, we assume that they change at least 40 times slower than $h_{m,k}$ [9]. Assuming a carrier frequency of 1.9 GHz and low user's mobility, the coherence interval is large and many orthogonal pilots can be used in channel estimation. Therefore, considering that the same pilot is assigned for users far away from each other, pilot contamination is assumed negligible [13].

Since each AP has N antenna elements, we have for the l th AP

$$\beta_{(l-1) \cdot N + 1, k} = \beta_{(l-1) \cdot N + 2, k} = \dots = \beta_{l \cdot N, k}, \quad \text{for } l = 1, \dots, L, \quad (2)$$

where we assume that the links between the antenna elements of an AP and the users have the same distance and are affected by the same path loss and shadowing effects [8, 24].

In order to reduce interference between signals intended for different users, considering that each user is served by all APs, channel coefficients need to be taken into account when forming transmitted signals. The system employs the time division duplex (TDD) protocol, where the channel is estimated through uplink training. First, all users send simultaneously and synchronously pilot sequences, $\mathbf{\Pi}_1, \dots, \mathbf{\Pi}_k, \dots, \mathbf{\Pi}_K \in \mathbb{C}^\tau$, to each antenna, where $\|\mathbf{\Pi}_k\|_2^2 = 1$. Then, APs perform MMSE channel estimation and obtain $\hat{g}_{m,k}$, the estimate of the channel coefficient $g_{m,k}$ between the m th antenna and the k th user. With $\hat{g}_{m,k}$, the data are transmitted to all users.

In the training step, the received signal sequence by the m th antenna is given by

$$\mathbf{y}_m = \sqrt{\rho_r \tau} \sum_{k=1}^K g_{m,k} \mathbf{\Pi}_k + \mathbf{w}_m, \quad (3)$$

where ρ_r is the uplink power, τ is the length of pilot sequences and $\mathbf{w}_m \sim \mathcal{CN}(0, \mathbf{I}_\tau)$ is the additive noise.

The MMSE estimate of $g_{m,k}$ is given by

$$\hat{g}_{m,k} = \frac{\sqrt{\rho_r \tau} \beta_{m,k}}{1 + \rho_r \tau \beta_{m,k}} \mathbf{\Pi}_k^H \mathbf{y}_m. \quad (4)$$

Since

$$\tilde{g}_{m,k} = g_{m,k} - \hat{g}_{m,k} \quad (5)$$

where $\tilde{g}_{m,k}$ is the CSI error between the m th antenna element and the k th user, we define

$$\hat{g}_{m,k} \sim \mathcal{CN}(0, \alpha_{m,k}), \quad \tilde{g}_{m,k} \sim \mathcal{CN}(0, \beta_{m,k} - \alpha_{m,k}), \quad (6)$$

and

$$\alpha_{m,k} = \frac{\rho_r \tau \beta_{m,k}^2}{1 + \rho_r \tau \beta_{m,k}}. \quad (7)$$

We notice that $\alpha_{m,k}$ as well as $\hat{g}_{m,k}$, are functions of $\beta_{m,k}$. Therefore, to evaluate perfect and different levels of imperfect CSI we consider $\alpha_{m,k}$ as an adjustable percentage of $\beta_{m,k}$ ($0 \leq \alpha_{m,k} \leq \beta_{m,k}$). Thus, we have

$$\begin{aligned} \alpha_{m,k} &= n_{m,k} \beta_{m,k} \\ \tilde{g}_{m,k} &= g_{m,k} - \hat{g}_{m,k}, \text{ and} \\ \mathbb{E} [|\tilde{g}_{m,k}|^2] &= (1 - n_{m,k}) \beta_{m,k}. \end{aligned} \quad (8)$$

We assume channel reciprocity, which means that the channel coefficients for the uplink and downlink are the same. After channel estimation is completed, the data is transmitted to all users. The signal received by the k th user is described by

$$\mathbf{y}_k = \sqrt{\rho_f} \mathbf{g}_k^T \mathbf{P} \mathbf{s} + w_k, \quad (9)$$

where ρ_f is the maximum transmit power of each antenna, $\mathbf{g}_k = [g_{1,k}, \dots, g_{M,k}]^T$ are the channel coefficients for user k , $\mathbf{P} \in$

$\mathbb{C}^{M \times K}$ is a generic linear precoding matrix, $\mathbf{s} = [s_1, \dots, s_K]^T$ is the zero mean symbol vector, with $\sigma_s^2 = \mathbb{E}(|s_k|^2)$, s_k is the data symbol intended for user k (uncorrelated between users), $w_k \sim \mathcal{CN}(0, \sigma_w^2)$ is the additive noise for user k and σ_w^2 is the noise variance. Therefore, for all users combined, we have the model

$$\mathbf{y} = \sqrt{\rho_f} \mathbf{G}^T \mathbf{P} \mathbf{s} + \mathbf{w}, \quad (10)$$

where $\mathbf{G} \in \mathbb{C}^{M \times K}$ is the channel matrix with elements $[\mathbf{G}]_{m,k} = g_{m,k}$ and $\mathbf{w} = [w_1, \dots, w_K]^T$ is the noise vector.

3 Proposed Iterative APS, MMSE Precoding and Power Allocation

In this section, we present the proposed iterative APS, linear MMSE precoding and power allocation techniques. In particular, the APS, precoding and power allocation tasks are integrated into a cell-free framework. Furthermore, APS algorithms based on an ES and LS coefficients, linear MMSE precoding and power allocation techniques based on max-min fairness and adaptive techniques are developed to improve the performance of existing approaches.

3.1 APS Techniques

We present two APS techniques for reducing power consumption, and consequently, increasing the energy efficiency in cell-free settings. First, we explain how to perform APS in the proposed and existing precoding techniques, taking into consideration the power allocation problems that are examined in this work. An ES APS (ES-APS) scheme and a suboptimal LS based APS (LS-APS) with low computational cost are devised [21].

3.1.1 Precoding with APS: To perform APS with precoding, different approaches need to be used in the optimization problems. First, an auxiliary matrix $\mathbf{Q} \in \mathbb{N}^{M \times K}$, where $[\mathbf{Q}]_{m,k} = q_{m,k} \in \{0, 1\}$ is considered in order to define which links between APs and users are taken into account (as 1) and the ones that are discarded (as 0). Since we perform APS rather than select antenna elements, all antennas from the l th AP should be considered or discarded. Therefore, the same procedure carried out for the LS coefficients ($\beta_{m,k}$) is done for $q_{m,k}$ where, $q_{(l-1) \cdot N + 1, k} = q_{(l-1) \cdot N + 2, k} = \dots = q_{l \cdot N, k}$, for $l = 1, \dots, L$.

Before calculating the MMSE precoder, each coefficient $q_{m,k}$ will be multiplied by $\alpha_{m,k}$, $\beta_{m,k}$, $\hat{g}_{m,k}$ and $\tilde{g}_{m,k}$ so that $\alpha'_{m,k} = q_{m,k} \cdot \alpha_{m,k}$, $\beta'_{m,k} = q_{m,k} \cdot \beta_{m,k}$, $\hat{g}'_{m,k} = q_{m,k} \cdot \hat{g}_{m,k}$ and $\tilde{g}'_{m,k} = q_{m,k} \cdot \tilde{g}_{m,k}$. Then, precoding and power allocation algorithms are performed with these parameters. Similarly, we apply the same technique to the ZF precoder from [13]. On the other hand, to perform APS in the CB precoder from [9, 13], the max-min fairness power allocation algorithm turns into a mixed continuous/discrete optimization problem [27].

3.1.2 Exhaustive Search Selection (ES-APS): In this scheme, all possible sets considering a total of L APs, S selected APs and K users, will be tested so that the combination that maximizes the minimum SINR is chosen. Each possible vector will be a column of matrix \mathbf{V} , formed by S 1s and $(L - S)$ 0s. By choosing a set of K columns of matrix \mathbf{V} , performing all combinations between them, $\binom{L}{S}^K$, and by replicating every row N times to represent the selection of all N antenna elements of a certain AP, we form all possibilities for \mathbf{Q} .

The proposed optimization problem that performs APS using ES is given by

$$\max_{\mathbf{Q}} \min_k \text{SINR}_k \quad (11a)$$

$$\text{s.t.} \sum_{j=1}^M q_{j,k} = SN, \quad k = 1, \dots, K, \quad (11b)$$

where the SINR for user k reads as

$$\text{SINR}_k = \frac{\mathbb{E}[|A_1|^2]}{\sigma_w^2 + \sum_{i=1, i \neq k}^K \mathbb{E}[|A_{2,i}|^2] + \mathbb{E}[|A_3|^2]}, \quad (12)$$

which is different to the expression in [13] that relies on ZF precoding explicitly. In equation (12)

$$A_1 = \sqrt{\rho_f} \hat{\mathbf{g}}_k^T \mathbf{P} \mathbf{s}_k, \quad (13)$$

describes the desired signal, σ_w^2 is the noise variance,

$$A_{2,i} = \sqrt{\rho_f} \hat{\mathbf{g}}_k^T \mathbf{P} \mathbf{s}_i, \quad (14)$$

is the interference caused by user i for $i \neq k, i = 1, \dots, K$,

$$A_3 = \sqrt{\rho_f} \tilde{\mathbf{g}}_k^T \mathbf{P} \mathbf{s} \quad (15)$$

is the CSI error, all based on (9), and $q_{j,k}$ is the j th element of the k th column of \mathbf{Q} . Note that for the ES-APS approach, the whole procedure is done, including the precoding and power allocation steps, for each possibility. After calculation of the minimum SINRs for all arrangements, we choose the setup that yields the maximum minimum SINR.

3.1.3 Large-Scale-Fading-Based Selection with Fixed Number of APs (LS-APS): As an alternative to ES-APS, we devise a LS based selection method, where user k will only be associated with $S \leq L$ APs corresponding to the SN largest LS coefficients for user k , inspired by the algorithm in [21]. Unlike the approach in [21] that chooses the number of selected APs based on their contribution to the sum of the LS coefficients, LS-APS sets a fixed S and then performs APS. This allows a fair comparison with ES-APS.

Algorithm 1 LS-APS

- 1: Estimate $\beta_k = [\beta_{1,k}, \dots, \beta_{M,k}]^T$ for user k and sort the elements in descending order.
 - 2: Choose the largest SN elements of β_k .
 - 3: In the auxiliary vector \mathbf{q}_k , assign 1s to the corresponding SN largest elements of β_k and 0s to the remaining ones.
 - 4: Let $\mathbf{q}_k, k = 1, \dots, K$ be the columns of the matrix \mathbf{Q} .
 - 5: Let $\alpha_{m,k} = q_{m,k} \cdot \alpha_{m,k}, \beta_{m,k} = q_{m,k} \cdot \beta_{m,k}, \hat{g}_{m,k} = q_{m,k} \cdot \hat{g}_{m,k}$ and $\tilde{g}_{m,k} = q_{m,k} \cdot \tilde{g}_{m,k}$, where $q_{m,k}$ is the m th element of vector \mathbf{q}_k .
 - 6: Perform the precoding and power allocation with the new parameters.
-

To perform APS with LS-APS, as detailed in Algorithm 1, we first need to estimate $\beta_k = [\beta_{1,k}, \dots, \beta_{M,k}]^T$ for user k and sort the elements in descending order. Then, we assign in an auxiliary vector \mathbf{q}_k 1s corresponding to the SN largest elements of β_k and 0s to the remaining entries. By grouping \mathbf{q}_k as the columns of \mathbf{Q} , we obtain the matrix \mathbf{Q} . This completes the APS stage.

3.2 Precoder Design

In this subsection, we detail the derivation of the proposed linear MMSE precoder. Unlike existing approaches [13, 26], in the proposed precoding technique, we consider power allocation in the derivation and take into account the CSI matrix after APS, $\hat{\mathbf{G}}'$, instead of the actual channel matrix \mathbf{G} , since the APs have imperfect CSI. Furthermore, we consider an iterative MMSE precoder with power allocation in order to maximize the minimum SINR. If the conventional MMSE precoder in [26] was applied to this type of system, the performance would be degraded due to the lack of appropriate power allocation. In the precoder design we also take

into account a normalization factor f^{-1} at the receivers, which can be interpreted as an automatic gain control [26].

By modifying the expression in (9), taking into account both the precoding and the power allocation matrices, the signal received by the k th user is given by

$$y_k = \sqrt{\rho_f} \hat{\mathbf{g}}_k^T \mathbf{P} \mathbf{N} \mathbf{s} + w_k, \quad (16)$$

where $\mathbf{N} \in \mathbb{R}_+^{K \times K}$ is the power allocation diagonal matrix with $\sqrt{\eta_1}, \dots, \sqrt{\eta_K}$ on its diagonal and η_k is the power coefficient of user k . Therefore, for all users combined, we have

$$\mathbf{y} = \sqrt{\rho_f} \mathbf{G}^T \mathbf{P} \mathbf{N} \mathbf{s} + \mathbf{w}, \quad (17)$$

To obtain the proposed MMSE precoder, the following optimization is solved [26]

$$\{\mathbf{P}_{\text{MMSE}}, \mathbf{N}, f_{\text{MMSE}}\} = \underset{\{\mathbf{P}, \mathbf{N}, f\}}{\text{argmin}} \mathbb{E} \left[\left\| \mathbf{s} - f^{-1} \mathbf{y} \right\|_2^2 \right] \quad (18a)$$

$$\text{s.t.} \mathbb{E} \left[\left\| \mathbf{x} \right\|_2^2 \right] = E_{tr}, \quad (18b)$$

where the transmitted signal is given by

$$\mathbf{x} = \sqrt{\rho_f} \mathbf{P} \mathbf{N} \mathbf{s}. \quad (19)$$

The average transmit power is described by

$$\mathbb{E} \left[\left\| \mathbf{x} \right\|_2^2 \right] = \rho_f \text{tr} \left(\mathbf{P} \mathbf{N} \mathbf{C}_s \mathbf{N}^H \mathbf{P}^H \right) = E_{tr}, \quad (20)$$

where $\mathbf{C}_s = \mathbb{E} [\mathbf{s} \mathbf{s}^H]$ is the symbol covariance matrix.

By constructing the Lagrangian function with the Lagrange multiplier λ , setting its derivatives with respect to the precoder and the normalization factor to zero, and considering a power allocation matrix \mathbf{N} , we can compute the precoder \mathbf{P} and the normalization f , as shown below:

$$\begin{aligned} \mathcal{L}(\mathbf{P}, \mathbf{N}, f, \lambda) = & \mathbb{E} \left[\left\| \mathbf{s} - f^{-1} \mathbf{y} \right\|_2^2 \right] + \lambda \left(\rho_f \text{tr} \left(\mathbf{P} \mathbf{N} \mathbf{C}_s \mathbf{N}^H \mathbf{P}^H \right) \right. \\ & - E_{tr} \Big) = \text{tr}(\mathbf{C}_s) - f^{-1} \sqrt{\rho_f} \text{tr} \left(\hat{\mathbf{G}}'^T \mathbf{P} \mathbf{N} \mathbf{C}_s \right) \\ & - f^{-1} \sqrt{\rho_f} \text{tr} \left(\hat{\mathbf{G}}'^* \mathbf{C}_s \mathbf{N}^H \mathbf{P}^H \right) + f^{-2} \text{tr}(\mathbf{C}_w) \\ & + f^{-2} \rho_f \text{tr} \left(\hat{\mathbf{G}}'^* \hat{\mathbf{G}}'^T \mathbf{P} \mathbf{N} \mathbf{C}_s \mathbf{N}^H \mathbf{P}^H \right) \\ & + \lambda \left(\rho_f \text{tr} \left(\mathbf{P} \mathbf{N} \mathbf{C}_s \mathbf{N}^H \mathbf{P}^H \right) - E_{tr} \right), \end{aligned} \quad (21)$$

where $\mathbf{C}_w = \mathbb{E} [\mathbf{w} \mathbf{w}^H]$ is the noise covariance matrix.

Using Wirtinger's calculus and the result of the partial derivative $\partial \text{tr}(\mathbf{B} \mathbf{X}^H) / \partial \mathbf{X}^* = \mathbf{B}$, we have

$$\begin{aligned} \frac{\partial \mathcal{L}(\mathbf{P}, \mathbf{N}, f, \lambda)}{\partial \mathbf{P}^*} = & f^{-2} \rho_f \hat{\mathbf{G}}'^* \hat{\mathbf{G}}'^T \mathbf{P} \mathbf{N} \mathbf{C}_s \mathbf{N}^H \\ & + \lambda \rho_f \left(\mathbf{P} \mathbf{N} \mathbf{C}_s \mathbf{N}^H \right) - f^{-1} \sqrt{\rho_f} \hat{\mathbf{G}}'^* \mathbf{C}_s \mathbf{N}^H = 0 \end{aligned} \quad (22)$$

and

$$\begin{aligned} \frac{\partial \mathcal{L}(\mathbf{P}, \mathbf{N}, f, \lambda)}{\partial f} = & 2f^{-3} \text{tr} \left(- \left(\rho_f \hat{\mathbf{G}}'^T \mathbf{P} \mathbf{N} \mathbf{C}_s \mathbf{N}^H \mathbf{P}^H \hat{\mathbf{G}}'^* \right. \right. \\ & \left. \left. + \mathbf{C}_w \right) + f \sqrt{\rho_f} \text{Re} \left(\hat{\mathbf{G}}'^T \mathbf{P} \mathbf{N} \mathbf{C}_s \right) \right) = 0. \end{aligned} \quad (23)$$

Solving for (22), we obtain

$$\mathbf{P} = \frac{f}{\sqrt{\rho_f}} \underbrace{\left(\hat{\mathbf{G}}'^* \hat{\mathbf{G}}'^T + \lambda f^2 \mathbf{I}_M \right)^{-1} \hat{\mathbf{G}}'^*}_{\tilde{\mathbf{P}}} \mathbf{N}^{-1} = \frac{f}{\sqrt{\rho_f}} \tilde{\mathbf{P}} \mathbf{N}^{-1}. \quad (24)$$

By using the expression in (23), we arrive at

$$\text{tr} \left(f^2 \hat{\mathbf{G}}'^T \tilde{\mathbf{P}} \mathbf{C}_s \tilde{\mathbf{P}}^H \hat{\mathbf{G}}'^* + \mathbf{C}_w - f^2 \text{Re} \left(\hat{\mathbf{G}}'^T \tilde{\mathbf{P}} \mathbf{C}_s \right) \right) = 0. \quad (25)$$

Considering $\epsilon = \lambda f^2$ and using the relation presented in [26], described by

$$\text{tr} \left(\text{Re} \left(\hat{\mathbf{G}}'^T \tilde{\mathbf{P}} \mathbf{C}_s \right) \right) = \text{tr} \left(\left(\hat{\mathbf{G}}'^* \hat{\mathbf{G}}'^T + \epsilon \mathbf{I}_M \right) \tilde{\mathbf{P}} \mathbf{C}_s \tilde{\mathbf{P}}^H \right), \quad (26)$$

we obtain

$$\text{tr}(\mathbf{C}_w) - \epsilon f^2 \text{tr}(\tilde{\mathbf{P}} \mathbf{C}_s \tilde{\mathbf{P}}^H) = 0 \quad (27)$$

$$\text{tr}(\mathbf{C}_w) - \epsilon E_{tr} = 0 \quad (28)$$

$$\epsilon = \frac{\text{tr}(\mathbf{C}_w)}{E_{tr}} = \lambda f^2. \quad (29)$$

As shown above, λ is not directly required for finding the optimal solution. With this, the auxiliary precoding matrix reads as

$$\tilde{\mathbf{P}} = \left(\hat{\mathbf{G}}'^* \hat{\mathbf{G}}'^T + \frac{\text{tr}(\mathbf{C}_w)}{E_{tr}} \mathbf{I}_M \right)^{-1} \hat{\mathbf{G}}'^*. \quad (30)$$

By inserting (24) in (20)

$$\begin{aligned} E_{tr} &= \rho_f \text{tr}(\mathbf{P} \mathbf{N} \mathbf{C}_s \mathbf{N}^H \mathbf{P}^H) \\ &= \rho_f \text{tr} \left(\frac{f}{\sqrt{\rho_f}} \tilde{\mathbf{P}} \mathbf{N}^{-1} \mathbf{N} \mathbf{C}_s \mathbf{N}^H (\mathbf{N}^{-1})^H \tilde{\mathbf{P}}^H \frac{f}{\sqrt{\rho_f}} \right) \\ &= f^2 \text{tr}(\tilde{\mathbf{P}} \mathbf{C}_s \tilde{\mathbf{P}}^H), \end{aligned} \quad (31)$$

we find that

$$f_{\text{MMSE}} = \sqrt{\frac{E_{tr}}{\text{tr}(\tilde{\mathbf{P}} \mathbf{C}_s \tilde{\mathbf{P}}^H)}}. \quad (32)$$

Therefore, the proposed MMSE precoder that takes into account power allocation for cell-free systems is given by

$$\mathbf{P}_{\text{MMSE}} = \frac{f_{\text{MMSE}}}{\sqrt{\rho_f}} \left(\hat{\mathbf{G}}'^* \hat{\mathbf{G}}'^T + \frac{\text{tr}(\mathbf{C}_w)}{E_{tr}} \mathbf{I}_M \right)^{-1} \hat{\mathbf{G}}'^* \mathbf{N}^{-1}, \quad (33)$$

where $[\hat{\mathbf{G}}']_{m,k} = \hat{g}_{m,k}'$ is the CSI matrix after APS and we consider $\text{tr}(\mathbf{C}_w) = K\sigma_w^2$. We initialize the precoder considering a power allocation matrix $\mathbf{N} = \mathbf{I}_K$. After the MMSE precoder is obtained, we perform power allocation. With the new power allocation matrix \mathbf{N} , we substitute it in \mathbf{P}_{MMSE} . The last step of the iteration is to recalculate the matrix \mathbf{N} by employing \mathbf{P}_{MMSE} to perform power allocation. The final power allocation matrix, denoted as \mathbf{N}_{MMSE} , ensures that the power constraint is satisfied. Note that the \mathbf{N} present in the \mathbf{P}_{MMSE} expression is different from the final \mathbf{N}_{MMSE} . Therefore, \mathbf{N} and \mathbf{N}_{MMSE} will not cancel each other.

3.3 Power Allocation

We present Optimal Power Allocation (OPA), Adaptive Power Allocation (APA) and Uniform Power Allocation (UPA) techniques to obtain \mathbf{N} , a diagonal matrix with $\sqrt{\eta_1}, \dots, \sqrt{\eta_K}$ on its main diagonal, which is used to recompute the precoder \mathbf{P}_{MMSE} and find the matrix \mathbf{N}_{MMSE} .

3.3.1 Optimal Power Allocation (OPA): The max-min fairness power allocation problem with antenna constraint is given by

$$\max_{\boldsymbol{\eta}} \min_k \text{SINR}_k(\boldsymbol{\eta}) \quad (34a)$$

$$\text{s.t.} \sum_{i=1}^K \eta_i \delta_{m,i} \leq 1, m = 1, \dots, M, \quad (34b)$$

where

$$\delta_m = \text{diag} \left\{ \mathbb{E} \left[\mathbf{p}_m^T \mathbf{p}_m^* \right] \right\}, \quad (35)$$

$\delta_{m,i}$ is the i th element of vector δ_m and $\mathbf{p}_m = [p_{m,1}, \dots, p_{m,K}]$ is the m th row of the precoder \mathbf{P}_{MMSE} .

The optimization problem in an epigraph form employs the bisection method at each step and is described by

$$\text{find } \boldsymbol{\eta} \quad (36a)$$

$$\text{s.t.} \text{SINR}_k(\boldsymbol{\eta}) \geq t, k = 1, \dots, K, \quad (36b)$$

$$\sum_{i=1}^K \eta_i \delta_{m,i} \leq 1, m = 1, \dots, M, \quad (36c)$$

where $t = \frac{t_b + t_e}{2}$ is the midpoint of a chosen interval (t_b, t_e) that contains the optimal value t^* , as in [28]. The value T_{OPA} is the number of iterations for the bisection method.

3.3.2 Adaptive Power Allocation (APA): We propose an adaptive SG learning algorithm, which takes into account the gradient of the error in (18) to perform APA and has a per-antenna power constraint. Our main objective here is to propose alternatives to the OPA and UPA algorithms. The proposed APA technique aims to adjust the power coefficients η_k so that they minimize the effect of the interference at the received signal vector \mathbf{y} . First, an unconstrained optimization is performed in accordance with (37) to minimize the interference. Then, a per-antenna power constraint is applied in order to fulfill the cell-free network transmit power requirements.

To start with the optimization we will first calculate the cost function:

$$\begin{aligned} \mathcal{C}(\mathbf{N}) &= \mathbb{E} \left[\left\| \mathbf{s} - f^{-1} \mathbf{y} \right\|_2^2 \right] = \text{tr}(\mathbf{C}_s) + f^{-2} \text{tr}(\mathbf{C}_w) \\ &\quad - f^{-1} \sqrt{\rho_f} \text{tr} \left(\hat{\mathbf{G}}'^T \mathbf{P}_{\text{MMSE}} \mathbf{N} \mathbf{C}_s \right) \\ &\quad - f^{-1} \sqrt{\rho_f} \text{tr} \left(\mathbf{P}_{\text{MMSE}}^H \hat{\mathbf{G}}'^* \mathbf{C}_s \mathbf{N}^H \right) \\ &\quad + f^{-2} \rho_f \text{tr} \left(\mathbf{P}_{\text{MMSE}}^H \hat{\mathbf{G}}'^* \hat{\mathbf{G}}'^T \mathbf{P}_{\text{MMSE}} \mathbf{N} \mathbf{C}_s \mathbf{N}^H \right). \end{aligned} \quad (37)$$

The instantaneous gradient with respect to \mathbf{N}^* is

$$\begin{aligned} \hat{\nabla}_{\mathbf{N}^*} \mathcal{C}(\mathbf{N}) &= -f^{-1} \sqrt{\rho_f} \mathbf{P}_{\text{MMSE}}^H \hat{\mathbf{G}}'^* \mathbf{C}_s \\ &\quad + f^{-2} \rho_f \mathbf{P}_{\text{MMSE}}^H \hat{\mathbf{G}}'^* \hat{\mathbf{G}}'^T \mathbf{P}_{\text{MMSE}} \mathbf{N} \mathbf{C}_s. \end{aligned} \quad (38)$$

Therefore, the matrix \mathbf{N} is updated by

$$\begin{aligned} \mathbf{N}[i+1] &= \text{Re} \left(\mathbf{N}[i] - \mu \left(-f^{-1} \sqrt{\rho_f} \mathbf{P}_{\text{MMSE}}^H \hat{\mathbf{G}}'^* \mathbf{C}_s \right. \right. \\ &\quad \left. \left. + f^{-2} \rho_f \mathbf{P}_{\text{MMSE}}^H \hat{\mathbf{G}}'^* \hat{\mathbf{G}}'^T \mathbf{P}_{\text{MMSE}} \mathbf{N}[i] \mathbf{C}_s \right) \right). \end{aligned} \quad (39)$$

As previously stated, the updated \mathbf{N} matrix will be calculated based on the initial MMSE precoder, \mathbf{P}_{MMSE} , which considered $\mathbf{N} = \mathbf{I}_K$. With this new \mathbf{N} , we substitute it in \mathbf{P}_{MMSE} . Finally, with the new \mathbf{P}_{MMSE} matrix, we calculate the final \mathbf{N} , called \mathbf{N}_{MMSE} . It is important to emphasize that in every calculation of \mathbf{N} , the per-antenna power constraint is reinforced.

In Algorithm 2, the proposed adaptive SG learning strategy that performs APA is explained in further detail.

Algorithm 2 SG Adaptive Power Allocation Algorithm

- 1: **Parameters:** μ (step size) and T_{APA} (number of iterations).
 - 2: **Initialization:** $\eta_k[0] = 10^{-3}, k = 1, \dots, K$.
 - 3: **For** $i = 0:T_{\text{APA}}$
 - 4: Set $\mathbf{N}[i]$ as a diagonal matrix with $\sqrt{\eta[i]}$ on its diagonal.
 - 5: Define $\mathcal{C}(\mathbf{N}) = \left[\left\| \mathbf{s} - \mathbf{f}^{-1} \mathbf{y} \right\|_2^2 \right]$.
 - 6: Compute $\hat{\mathbf{N}}^* \mathcal{C}(\mathbf{N})$.
 - 7: Calculate $\mathbf{N}[i+1] = \mathbf{N}[i] - \mu \hat{\mathbf{N}}^* \mathcal{C}(\mathbf{N})$.
 - 8: Obtain $\mathbf{N}[i+1] = \text{Re}(\mathbf{N}[i+1] \odot \mathbf{I}_K)$.
 - 9: Compute $\eta_k[i+1] = |\mathbf{N}_{k,k}[i+1]|^2, k = 1, \dots, K$.
 - 10: Scale according to the per-antenna power constraint $\delta_m \cdot \eta[i+1] \leq 1, m = 1, \dots, M$ to adjust $\mathbf{N}[i+1]$
 - 11: **end**
 - 12: Obtain $\mathbf{N} = \mathbf{N}[i+1]$.
-

The proposed SG power allocation algorithm for APA has a complexity of $\mathcal{O}(T_{\text{APA}}MK^2)$ and converges within only a few iterations. In Fig. 2 we illustrate the performance of the proposed APA algorithm using a step size of $\mu = 0.25$ that converges within $T_{\text{APA}} = 5$ iterations.

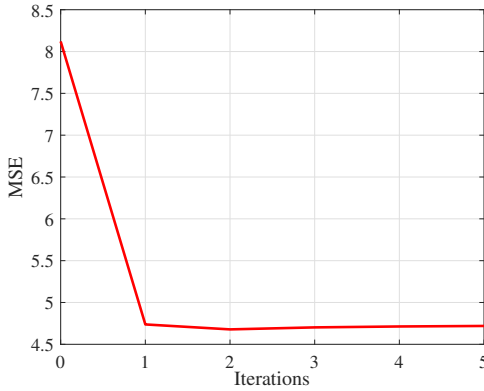


Fig. 2: Learning Algorithm - $L = 24, N = 4, S = 12, K = 8, n = 1, \text{SNR} = 25 \text{ dB}, \mu = 0.25$ and $E_{tr} = M\rho_f$.

3.3.3 Uniform Power Allocation (UPA): As an alternative to APA, an UPA scheme is proposed, based on the one in [13] where a per-antenna power constraint is taken into account. Consider a case where we wish to obtain equal η_k and that a certain antenna element m transmits with full power. Then, with η_k as its minimum possible value, we have

$$\eta_k = 1 / \left(\max_m \sum_{i=1}^K \delta_{m,i} \right), \quad k = 1, \dots, K, \quad (40)$$

where $\delta_{m,i}$ is the i th element of vector δ_m . Although (40) is a suboptimal solution, it has low complexity, is simple and has the potential to show the benefits of the MMSE precoder.

4 Analysis

In this section, we present a sum-rate analysis of the proposed techniques along with the computational complexity of the proposed and existing algorithms.

4.1 Sum-Rate

First, we will expand expressions (16) and (33) to obtain the received signal by user k :

$$\begin{aligned} y_k &= \sqrt{\rho_f} \mathbf{g}_k^T \mathbf{P}_{\text{MMSE}} \mathbf{N}_{\text{MMSE}} \mathbf{s} + w_k \\ &= \sqrt{\rho_f} \left(\hat{\mathbf{g}}_k' + \tilde{\mathbf{g}}_k' \right)^T \frac{f_{\text{MMSE}}}{\sqrt{\rho_f}} \left(\hat{\mathbf{G}}'^T \hat{\mathbf{G}}'^T + \frac{K\sigma_w^2}{E_{tr}} \mathbf{I}_M \right)^{-1} \hat{\mathbf{G}}' \mathbf{N}^{-1} \\ &\quad \mathbf{N}_{\text{MMSE}} \mathbf{s} + w_k \\ &= \underbrace{\sqrt{\rho_f} \hat{\mathbf{g}}_k'^T \frac{f_{\text{MMSE}}}{\sqrt{\rho_f}} \left(\hat{\mathbf{G}}'^T \hat{\mathbf{G}}'^T + \frac{K\sigma_w^2}{E_{tr}} \mathbf{I}_M \right)^{-1} \hat{\mathbf{G}}' \mathbf{N}^{-1} \mathbf{N}_{\text{MMSE}} \mathbf{s}}_{\text{desired signal + interference}} \\ &\quad + \underbrace{\sqrt{\rho_f} \tilde{\mathbf{g}}_k'^T \frac{f_{\text{MMSE}}}{\sqrt{\rho_f}} \left(\hat{\mathbf{G}}'^T \hat{\mathbf{G}}'^T + \frac{K\sigma_w^2}{E_{tr}} \mathbf{I}_M \right)^{-1} \hat{\mathbf{G}}' \mathbf{N}^{-1} \mathbf{N}_{\text{MMSE}} \mathbf{s}}_{\text{CSI error}} \\ &\quad + w_k, \end{aligned} \quad (41)$$

where $\hat{\mathbf{g}}_k' = [\hat{g}_{1,k}', \dots, \hat{g}_{M,k}']^T$ is the CSI vector for user k and $\tilde{\mathbf{g}}_k' = [\tilde{g}_{1,k}', \dots, \tilde{g}_{M,k}']^T$ is the CSI error vector for user k , both after APS.

Assuming uncorrelated Gaussian noise, the achievable rate of user k with the proposed iterative MMSE precoder is given by

$$R_{k,\text{MMSE}} = \log_2(1 + \text{SINR}_{k,\text{MMSE}}). \quad (42)$$

The sum-rate is then given by

$$R_{\text{MMSE}} = \sum_{k=1}^K \log_2(1 + \text{SINR}_{k,\text{MMSE}}), \quad (43)$$

where the SINR reads as

$$\text{SINR}_{k,\text{MMSE}} = \frac{\mathbb{E}[|A_1|^2]}{\sigma_w^2 + \sum_{i=1, i \neq k}^K \mathbb{E}[|A_{2,i}|^2] + \mathbb{E}[|A_3|^2]}. \quad (44)$$

Note that the considered SINR is different to the SINR expression used in [13], which relies on ZF precoding. In the expression above, the quantity

$$A_1 = \sqrt{\rho_f} \hat{\mathbf{g}}_k'^T \mathbf{p}_k \sqrt{\eta_k} s_k, \quad (45)$$

is the desired signal, the parameter

$$A_{2,i} = \sqrt{\rho_f} \hat{\mathbf{g}}_k'^T \mathbf{p}_i \sqrt{\eta_i} s_i, \quad (46)$$

is the interference caused by user i and

$$A_3 = \sqrt{\rho_f} \tilde{\mathbf{g}}_k'^T \mathbf{P}_{\text{MMSE}} \mathbf{N}_{\text{MMSE}} \mathbf{s} \quad (47)$$

refers to CSI error. The mean-square values of A_1 , $A_{2,i}$ and A_3 are computed as follows:

$$\begin{aligned} \mathbb{E}[|A_1|^2] &= \mathbb{E} \left[\left(\sqrt{\rho_f} \hat{\mathbf{g}}_k'^T \mathbf{p}_k \sqrt{\eta_k} s_k \right)^* \left(\sqrt{\rho_f} \hat{\mathbf{g}}_k'^T \mathbf{p}_k \sqrt{\eta_k} s_k \right) \right] \\ &= \rho_f \eta_k \psi_k \end{aligned} \quad (48)$$

$$\begin{aligned} \mathbb{E}[|A_{2,i}|^2] &= \mathbb{E} \left[\left(\sqrt{\rho_f} \hat{\mathbf{g}}_k'^T \mathbf{p}_i \sqrt{\eta_i} s_i \right)^* \left(\sqrt{\rho_f} \hat{\mathbf{g}}_k'^T \mathbf{p}_i \sqrt{\eta_i} s_i \right) \right] \\ &= \rho_f \eta_i \phi_{k,i} \end{aligned} \quad (49)$$

$$\begin{aligned}\mathbb{E}[|A_3|^2] &= \mathbb{E}\left[\left|\sqrt{\rho_f}\tilde{\mathbf{g}}_k^T \mathbf{P}_{\text{MMSE}} \mathbf{N}_{\text{MMSE}}\right|^2\right] \\ &= \rho_f \sum_{i=1}^K \eta_i \gamma_{k,i}\end{aligned}\quad (50)$$

In the expressions above, we have

$$\psi_k = \mathbf{p}_k^H \tilde{\mathbf{g}}_k^* \tilde{\mathbf{g}}_k^T \mathbf{p}_k, \text{ for } k = 1, \dots, K, \quad (51)$$

$$\phi_{k,i} = \mathbf{p}_i^H \tilde{\mathbf{g}}_k^* \tilde{\mathbf{g}}_k^T \mathbf{p}_i, \text{ for } i \neq k, i = 1, \dots, K, \quad (52)$$

$$\gamma_k = \text{diag}\left\{\mathbf{P}_{\text{MMSE}}^H \mathbb{E}\left[\tilde{\mathbf{g}}_k^* \tilde{\mathbf{g}}_k^T\right] \mathbf{P}_{\text{MMSE}}\right\}, \quad (53)$$

where $\mathbf{p}_k = [p_{1,k}, \dots, p_{M,k}]^T$ is the column k of matrix \mathbf{P}_{MMSE} , ψ_k is the k th element of vector ψ , $\phi_{k,i}$ is the i th element of vector ϕ_k , $\gamma_{k,i}$ is the i th element of vector γ_k , and $\mathbb{E}[\tilde{\mathbf{g}}_k^* \tilde{\mathbf{g}}_k^T]$ is a diagonal matrix with $((1 - n_{m,k})\beta_{mk}')$ on its m th diagonal element.

By substituting (48), (49) and (50) in the $\text{SINR}_{k,\text{MMSE}}$ expression we get

$$\text{SINR}_{k,\text{MMSE}} = \frac{\rho_f \eta_k \psi_k}{\sigma_w^2 + \rho_f \sum_{i=1, i \neq k}^K \eta_i \phi_{k,i} + \rho_f \sum_{i=1}^K \eta_i \gamma_{k,i}}. \quad (54)$$

Note that in $\text{SINR}_{k,\text{MMSE}}(\eta)$, the numerator and denominator are linear functions of η . Consequently, $\text{SINR}_{k,\text{MMSE}}(\eta)$ is a quasilinear function, enabling us to use the bisection method [28]. The proof of expression (54) is detailed in Appendix 6.

4.2 Computational Complexity

Here, we evaluate the computational complexity of the proposed and existing techniques [29].

In Table 2, we notice that the overall complexity of the proposed iterative LS-APS with MMSE precoding and OPA is comparable to the existing ZF precoder and is higher than that of the CB precoder from [9, 13]. Depending on the number of iterations of the bisection method, T_{OPA} , the complexity of the OPA algorithm may prevail, or the one of the MMSE and ZF precoder, if $M^3 > T_{\text{OPA}} K^{3.5}$. The same can be said regarding the APA algorithm, depending on $M^3 > T_{\text{APA}} M K^2$, or the opposite. When evaluating the impact of APS on precoding and SINR computation, there is a reduction in computational complexity from $\mathcal{O}(M^3)$ to $\mathcal{O}((SN)^3)$ and from $\mathcal{O}(M^2 K^2)$ to $\mathcal{O}((SN)^2 K^2)$. For the power allocation techniques proposed, APS only affects the proposed APA and UPA algorithms, with a reduction in complexity from $\mathcal{O}(T_{\text{APA}} M K^2)$ to $\mathcal{O}(T_{\text{APA}} S N K^2)$ and from $\mathcal{O}(M K^2)$ to $\mathcal{O}(S N K^2)$, respectively, where $M = L N > S N$. For the precoders with ES-APS, we notice that the complexity is $\mathcal{O}(L!)$, which is very costly when compared to LS-APS. Thus, for ES-APS we only consider a very small system with $L = 5$ single-antenna APs.

5 Simulations

In this section, we assess the proposed and existing techniques using simulations. The proposed APS is first applied, followed by MMSE precoder, which is initially calculated with $\mathbf{N} = \mathbf{I}_K$, and then power allocation is performed. With the corresponding power coefficients, the precoder is recomputed and the final power allocation is performed, completing two iterations in total. In all experiments, we performed 120 channel realizations and assumed $\sigma_s^2 = 1$. Moreover, we consider that $n_{m,k} = n$, which means that all channel coefficients are estimated with the same accuracy.

We consider L APs with N antenna elements, a total of $M = L N$ antennas, and K single-antenna users uniformly distributed at

Table 2 Computational Complexity

| APS Techniques | Proposed LS-APS Proposed ES-APS | $\mathcal{O}(M^2)$ $\mathcal{O}(L!)$ |
|------------------|---|---|
| Precoding | Proposed MMSE ZF CB [9, 13] | $\mathcal{O}(M^3)$ $\mathcal{O}(M^3)$ $\mathcal{O}(M K)$ |
| SINR Computation | Proposed MMSE ZF CB [9, 13] | $\mathcal{O}(M^2 K^2)$ $\mathcal{O}(M^2 K^2)$ $\mathcal{O}(M K^2)$ |
| Power Allocation | Proposed OPA Proposed SG APA Algorithm Proposed UPA | $\mathcal{O}(T_{\text{OPA}} K^{3.5})$ [30] $\mathcal{O}(T_{\text{APA}} M K^2)$ $\mathcal{O}(M K^2)$ |

random within an area of 1 km^2 . The LS coefficients from (1) are modeled by

$$\beta_{m,k} = \text{PL}_{m,k} \cdot 10^{\frac{\sigma_{sh} z_{m,k}}{10}}, \quad (55)$$

where $\text{PL}_{m,k}$ is the path loss and $10^{\frac{\sigma_{sh} z_{m,k}}{10}}$ refers to the shadow fading with standard deviation $\sigma_{sh} = 8 \text{ dB}$ and $z_{m,k} \sim \mathcal{N}(0, 1)$. The path loss is based on a three-slope model [31], in dB, defined as

$$\text{PL}_{m,k} = \begin{cases} -L - 35 \log_{10}(d_{m,k}), & \text{if } d_{m,k} > d_1 \\ -L - 15 \log_{10}(d_1) - 20 \log_{10}(d_{m,k}), & \text{if } d_0 < d_{m,k} \leq d_1 \\ -L - 15 \log_{10}(d_1) - 20 \log_{10}(d_0), & \text{if } d_{m,k} \leq d_0 \end{cases} \quad (56)$$

where

$$\begin{aligned}L &\triangleq 46.3 + 33.9 \log_{10}(f_{\text{freq}}) - 13.82 \log_{10}(h_{\text{AP}}) \\ &\quad - (1.1 \log_{10}(f_{\text{freq}}) - 0.7) h_u + (1.56 \log_{10}(f_{\text{freq}}) - 0.8),\end{aligned} \quad (57)$$

$d_{m,k}$ is the distance between the m th antenna element and the k th user, $d_1 = 50 \text{ m}$, $d_0 = 10 \text{ m}$, $f_{\text{freq}} = 1900 \text{ MHz}$ is the carrier frequency in MHz, $h_{\text{AP}} = 15 \text{ m}$ is the AP antenna height in meters and $h_u = 1.65 \text{ m}$ is the user antenna height in meters, as in [9]. When $d_{m,k} \leq d_1$ there is no shadowing.

We consider strong path loss, which is typical of Cell-Free Massive MIMO systems, and define ρ_f based on the signal-to-noise ratio (SNR) given by [32]

$$\rho_f = \frac{\text{SNR} \cdot \text{tr}(\mathbf{C}_w)}{\mathbb{E}[\|\hat{\mathbf{G}}\|_F^2]} = \frac{\text{SNR} \cdot K \sigma_w^2}{\text{tr}(\hat{\mathbf{G}} \hat{\mathbf{G}}^H)}, \quad (58)$$

where

$$\sigma_w^2 = T_0 \times k_B \times B \times N F(W), \quad (59)$$

$T_0 = 290 \text{ (Kelvin)}$ is the noise temperature, $k_B = 1.381 \times 10^{-23} \text{ (Joule per Kelvin)}$ is the Boltzmann constant, $B = 20 \text{ MHz}$ is the bandwidth and $N F = 9 \text{ dB}$ is the noise figure. Therefore, the SNR expression is

$$\text{SNR} = \frac{\rho_f \mathbb{E}[\|\hat{\mathbf{G}}\|_F^2]}{\text{tr}(\mathbf{C}_w)} = \frac{\rho_f \text{tr}(\hat{\mathbf{G}} \hat{\mathbf{G}}^H)}{K \sigma_w^2}. \quad (60)$$

Since we are comparing different precoding designs in this work, an abstract SNR expression is being considered. Note that the expression does not take in account the beamforming gain. Thus, the SINR can be higher than the SNR in the numerical results [33].

In the experiments we will present next, we combined different precoding, power allocation and APS techniques. Therefore, when describing a technique we will use the following notation:

- Precoding + Power Allocation + APS

For each category, we have the following methods:

- Precoding: CB, ZF and the Proposed MMSE
- Power Allocation: OPA, APA and UPA
- APS: ES-APS and LS-APS

In the first experiment, we present a comparison of the proposed techniques and the centralized scheme from [17]. In this comparison, we consider MMSE+UPA with NS and LS-APS. From [17], we considered the P-MMSE technique and the developed clustering method, which implies APS. The precoding scheme is actually a centralized MMSE precoder combined with a heuristic power allocation strategy. For this simulation, we employ the same distance matrix, with coefficients $d_{m,k}$, so that each method is able to calculate its own APS and precoding vectors. We perform 150 channel realizations under perfect CSI conditions. Additionally, we consider in both scenarios an SNR variation between 0 and 25 dB, which is described by (60), given $\sigma_w^2 = 1$. As shown in Fig. 3, gains up to 6 dB can be obtained by the MMSE + UPA schemes over the P-MMSE technique. For a sum-rate equal to 20, differences around 3 dB can be observed when evaluating the compared techniques.

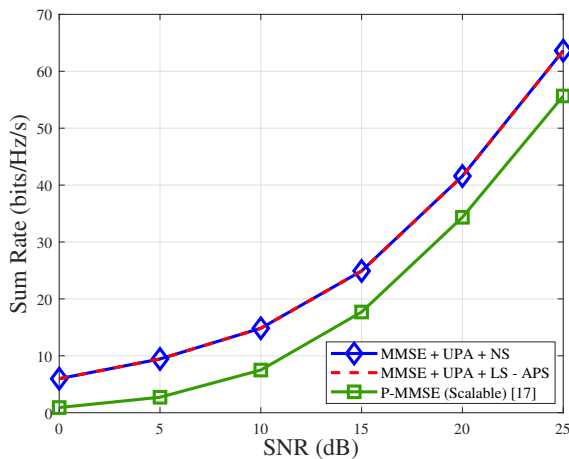


Fig. 3: Sum-Rate vs. SNR with $M = 128$, $K = 16$, $n = 1$ and 150 channel realizations.

Second, we demonstrate the effects of APS in sum-rate performance. With the decrease in the number of selected APs, a performance degradation is observed. In Fig. 4, we notice that with 50% of selected APs, the performance is comparable to the case with no selection (NS). However, when less than 10% of the available APs are selected for each user, a significant performance degradation can be perceived. This phenomenon occurs because the channel matrices are sparse, $\mathbf{G}, \hat{\mathbf{G}}, \tilde{\mathbf{G}}$, which means that only the APs close to the user are efficiently transmitting signals to it. In combination with the per antenna power constraint one can conclude that some channels do not contribute to the received signal and therefore can be discarded through an APS scheme. Although a weak channel does not necessarily mean a small contribution to the received signal, its association with the per antenna power constraint leads to this effect. When selecting half of the available APs, most of the available benefits are obtained since the remaining channels contribute substantially less to performance. Nevertheless, when less APs are selected, a larger percentage of links that significantly contribute to performance improvement can be discarded, leading to performance degradation.

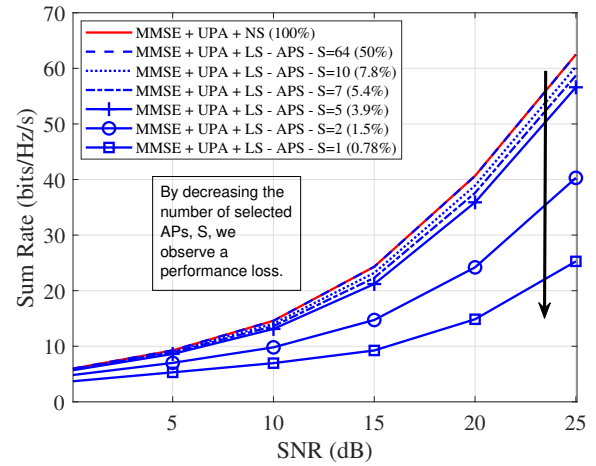


Fig. 4: Sum-Rate vs. SNR with $L = 128$, $N = 1$, $K = 16$, $n = 1$, 120 channel realizations and $E_{tr} = M\rho_f$.

In Fig. 5, we compare the proposed iterative MMSE precoder with OPA and UPA with the CB and ZF precoders from [9, 13] in terms of sum-rate vs. SNR. In the first case, both ES-APS and LS-APS are compared to NS. Since ES-APS has high computational complexity, a very small system is considered in Fig. 5 and Fig. 6 with $L = 5$ single-antenna APs, $S = 3$ selected APs and $K = 2$ users only. In Fig. 6, we explore the same scenario, but instead of looking at OPA, we consider APA.

As shown in Fig. 5, MMSE + OPA is the scheme with the best performance. It is also visible that MMSE + UPA achieves higher rates than ZF + OPA for lower SNR values and it is better than ZF + UPA in the whole experiment. Additionally, we can see that the application of ES-APS/LS-APS generates comparable or even improved results for MMSE, ZF and CB + OPA. In the case of CB + UPA, there is a small degradation in performance when applying APS techniques. We remark that both selection schemes are comparable and are shown together for aesthetic purposes. The reason for this is that the cell-free channel matrix is sparse, which means that most of the APs far from the user are not efficiently transmitting signals to it. Therefore, by APS methods, also called clustering methods, we discard unnecessary transmissions. By using the LS-APS technique, which chooses the strongest channels to transmit, we can approach the optimal technique, which tests all arrangements. As a result, we conclude that the suboptimal scheme presented is an effective replacement for the optimal technique but with lower computational complexity.

Fig. 6 provide us some insight on the performance of the MMSE + APA scheme when compared to MMSE + UPA, and ZF. The uniform and adaptive techniques have a better performance than ZF + OPA for lower values of SNR. Moreover, we can see that the MMSE+APA scheme performs better than MMSE+UPA. In all combined techniques, OPA performs better than UPA and APA. In both figures, the MMSE precoder from [26] shows a degraded performance due to its lack of appropriate power allocation.

In the second experiment, we explore LS-APS in a large system with $L = 128$ single-antenna APs, $S = 64$ selected APs and $K = 16$ users, in terms of minimum SINR and sum-rate. As we did previously, in Fig. 7 and Fig. 9 we compare the CB and ZF precoders from [9, 13] with OPA and UPA and in Fig. 8 and Fig. 10 we substitute OPA for APA.

The results in Fig. 7 and Fig. 8 validate the functionality of the max-min fairness power allocation algorithm, where our main objective was that the OPA algorithms have at least the same minimum SINR as UPA ones, if not higher. As shown in Fig. 7, in higher SINR values, we can see the considerable improvement provided by the OPA algorithm, when compared to UPA. On the other hand, in Fig. 8, we study APA instead of OPA and we are able to observe that its performance is comparable to the one of the UPA scheme. We can

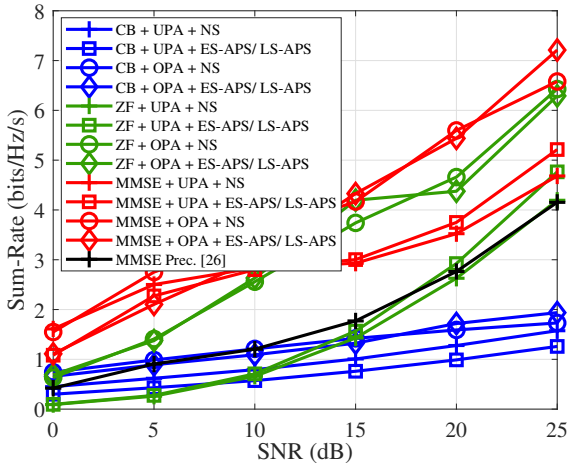


Fig. 5: Sum-Rate vs. SNR with $L = 5, N = 1, S = 3, K = 2, n = 0.99$, 120 channel realizations and $E_{tr} = M\rho_f$.

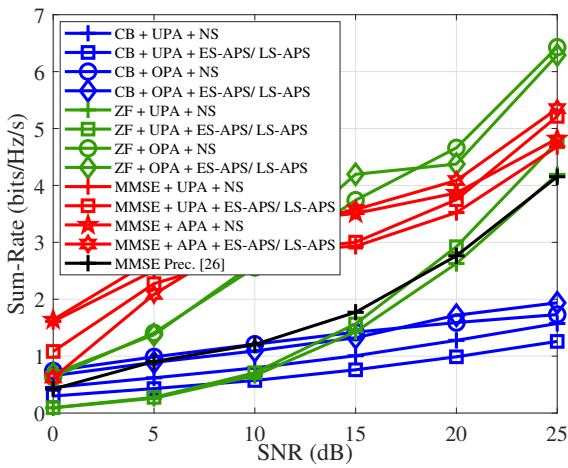


Fig. 6: Sum-Rate vs. SNR with $L = 5, N = 1, S = 3, K = 2, n = 0.99$, 120 channel realizations and $E_{tr} = M\rho_f$.

conclude that the OPA technique improves the minimum SINR, as opposed to APA and UPA, where the goal is to improve the overall performance of the system.

In Fig. 9, the MMSE + OPA scheme has the best performance compared to the other schemes. Moreover, the MMSE + UPA technique achieves higher rates than ZF + UPA. We also note that performance is maintained when applying LS-APS for large systems, except for CB + OPA, when performance is improved. For all precoders, OPA provides significantly better rates than UPA. In a larger system, the performance of the precoder from [26] is also not as good as MMSE + UPA due to its inappropriate design for cell-free systems.

We notice in Fig. 10 that MMSE + APA can provide an improvement in performance when compared to MMSE + UPA, in terms of sum-rate. Therefore, it is an attractive solution in comparison with the remaining precoders when combined with UPA.

In Fig. 11, we assess the sum-rate vs. SNR for APs with different numbers of antenna elements. We compare MMSE+UPA+LS-APS under scenarios with a fixed total number of antennas, $M = 256$, but with different numbers of antennas per AP. Fig. 11 verifies that the best sum-rate performance is achieved with a single-antenna AP scenario. As pointed out in [8], although adding more antennas to an AP leads to more channel hardening, it also brings losses in macro-diversity, resulting in lower average rates. Therefore, the best settings for the proposed scenarios are presented in Figs. 5 to 10.

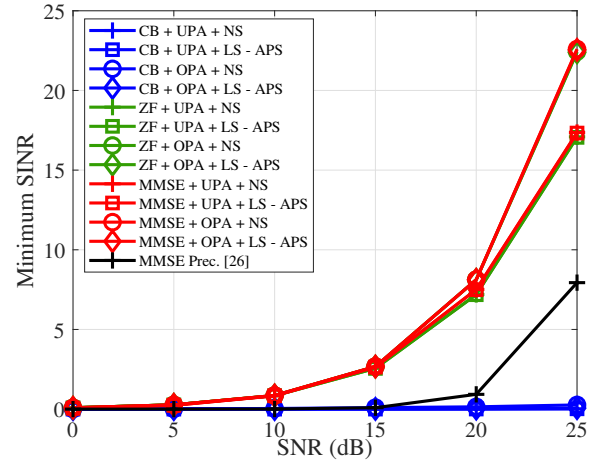


Fig. 7: Minimum SINR vs. SNR with $L = 128, N = 1, S = 64, K = 16, n = 0.99$, 120 channel realizations and $E_{tr} = M\rho_f$.

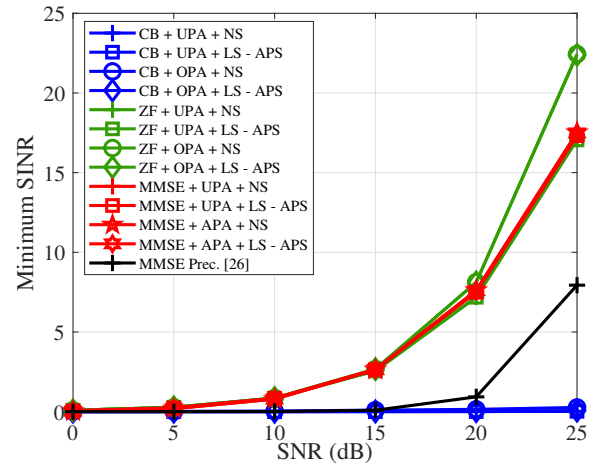


Fig. 8: Minimum SINR vs. SNR with $L = 128, N = 1, S = 64, K = 16, n = 0.99$, 120 channel realizations and $E_{tr} = M\rho_f$.

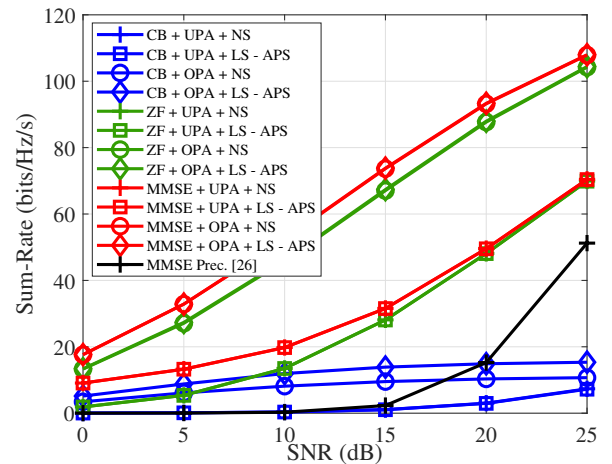


Fig. 9: Sum-Rate vs. SNR with $L = 128, N = 1, S = 64, K = 16, n = 0.99$, 120 channel realizations and $E_{tr} = M\rho_f$.

The last experiment also considers a large system, but now in terms of BER vs. SNR using multiple-antenna APs. We assume perfect CSI ($n = 1$) and QPSK modulation. We consider LS-APS

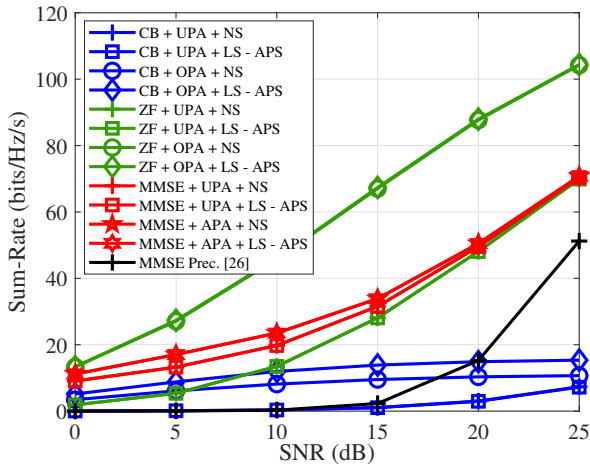


Fig. 10: Sum-Rate vs. SNR with $L = 128$, $N = 1$, $S = 64$, $K = 16$, $n = 0.99$, 120 channel realizations and $E_{tr} = M\rho_f$.

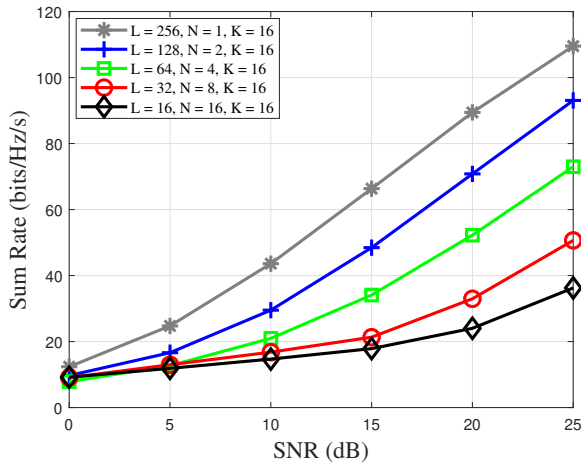


Fig. 11: Sum-Rate vs. SNR with $M = 256$, $S = 128$, $K = 16$, $n = 0.99$, 120 channel realizations and $E_{tr} = M\rho_f$. MMSE+UPA+LS-APS scheme used.

with an antenna array of $N = 4$ elements each, $L = 24$ APs (total of $M = 96$ antennas), $S = 12$ selected APs (total of 48 selected antennas) and $K = 8$ users.

As in experiment 2, LS-APS causes no degradation in performance, with the benefit of reducing the computational complexity by half. The insight provided by Fig. 12 and Fig. 13 is the same as before. MMSE + OPA has the best performance when compared to other precoders and OPA performs better when applied to all precoders. We also emphasize here that MMSE + APA is a promising solution against UPA and improves performance. Additionally, we remark that in terms of BER, for higher SNR values, MMSE + APA is comparable to ZF + OPA.

6 Conclusions

We have presented iterative APS, MMSE precoding and power allocation techniques for the downlink of a Cell-Free Massive MIMO system with single and multiple-antenna APs, in the presence of perfect and imperfect CSI. Two APS schemes were presented, one based on ES, which takes the experiment to its optimal scenario and the other, less complex but with comparable performance, based on the largest LS coefficients. A linear MMSE precoder has been developed by taking into account a power allocation matrix in its derivation. Then, three power control algorithms are proposed, OPA,

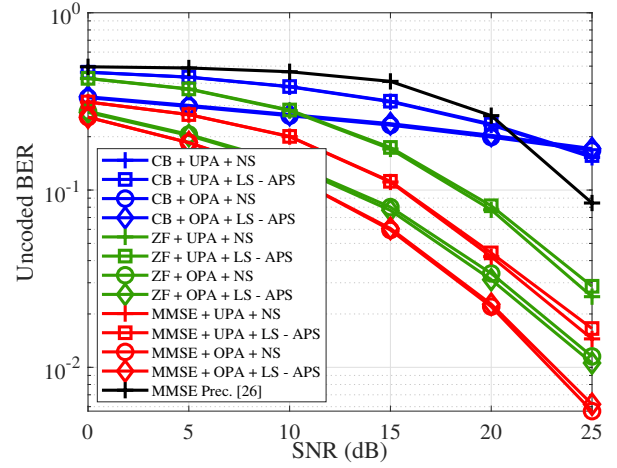


Fig. 12: BER vs. SNR with $L = 24$, $N = 4$, $S = 12$, $K = 8$, $n = 1$, 120 channel realizations, 100 symbols per packet and $E_{tr} = M\rho_f$.

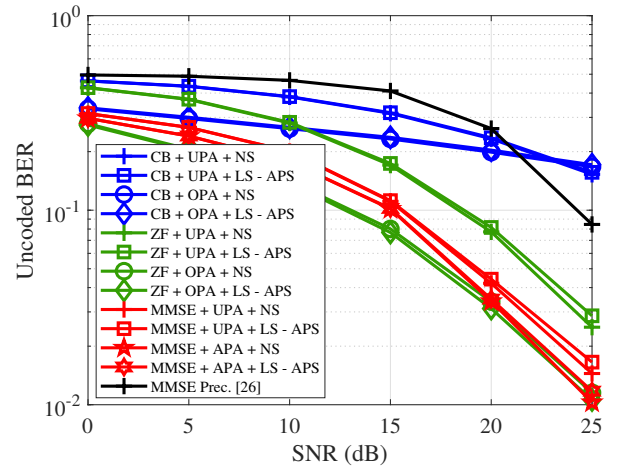


Fig. 13: BER vs. SNR with $L = 24$, $N = 4$, $S = 12$, $K = 8$, $n = 1$, 120 channel realizations, 100 symbols per packet and $E_{tr} = M\rho_f$.

APA and UPA, with different performances, criteria and computational complexities. We have then derived sum-rate expressions for the proposed scheme along with a study of the computational cost of all proposed and existing techniques. Simulations show that the proposed techniques outperform existing approaches and can reduce computational cost. The scalability of the proposed system as well as a robust version of the proposed schemes will be analysed in future works.

7 References

- [1] Marzetta, T. L.: 'Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas', IEEE Trans. Wireless Commun., 2010, 9, (11), pp. 3590-3600.
- [2] Liu, Z., Dai, L.: 'A comparative study of downlink MIMO cellular networks with co-located and distributed base-station antennas', IEEE Trans. Wireless Commun., 2014, 13, (11), pp. 6259-6274.
- [3] Truong, K. T., Heath, R. W.: 'The Viability of Distributed Antennas for Massive MIMO Systems', Proc. Asilomar Conf. Signals, Syst. Comput., Pacific Grove, CA, Nov. 2013, pp. 1318-1323.
- [4] Joung, J., Chia, Y. K., Sun, S.: 'Energy-Efficient, Large-Scale Distributed-Antenna System (L-DAS) for Multiple Users', IEEE J. Sel. Topics Signal Process., 2014, 8, (5), pp. 954-965.
- [5] Ngo, H. Q., Ashikhmin, A., Yang, H., et al.: 'Cell-Free Massive MIMO: Uniformly Great Service For Everyone', Proc. 16th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC), Stockholm, Jun. 2015, pp. 201-205.

- [6] Yang, H., Marzetta, T. L.: 'Energy efficiency of massive MIMO: Cell Free vs. Cellular', Proc. IEEE 87th Veh. Technol. Conf. (VTC Spring), Porto, Jun. 2018, pp. 1–5.
- [7] Marzetta, T. L., Larsson, E. G., Yang, H., *et al*: 'Fundamentals of Massive MIMO' (Cambridge University Press, Cambridge, United Kingdom, 2016)
- [8] Chen, Z., Björnson, E.: 'Channel hardening and favorable propagation in cell-free massive MIMO with stochastic geometry', IEEE Trans. Commun., 2018, **66**, (11), pp. 5205–5219.
- [9] Ngo, H. Q., Ashikhmin, A., Yang, H., *et al*: 'Cell-free massive MIMO versus small cells', IEEE Trans. Wireless Commun., 2017, **16**, (3), pp. 1834–1850.
- [10] Björnson, E., Sanguinetti, L.: 'Making cell-free massive MIMO competitive with MMSE processing and centralized implementation', IEEE Trans. Wireless Commun., 2020, **19**, (1), pp. 77–90.
- [11] Nayeri, E., Ashikhmin, A., Marzetta, T. L., *et al*: 'Performance of Cell-Free Massive MIMO Systems with MMSE and LSFD Receivers', Proc. 50th Asilomar Conf. Signals, Syst. Comput., Pacific Grove, CA, Nov. 2016, pp. 203–207.
- [12] Chen, S., Zhang, J., Björnson, E., *et al*: 'Structured Massive Access for Scalable Cell-Free Massive MIMO Systems', IEEE J. Sel. Areas in Comm., 2020.
- [13] Nayeri, E., Ashikhmin, A., Marzetta, T., *et al*: 'Precoding and Power Optimization in Cell-Free Massive MIMO Systems', IEEE Trans. Wireless Commun., 2017, **16**, (7), pp. 4445–4459.
- [14] Ngo, H. Q., Tran, L., Duong, T. Q., *et al*: 'Energy Efficiency Optimization for Cell-Free Massive MIMO', Proc. 18th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC), Jul. 2017.
- [15] Nguyen, L. D., Duong, T. Q., Ngo, H. Q., *et al*: 'Energy Efficiency in Cell-Free Massive MIMO with Zero-Forcing Precoding Design', IEEE Commun. Letters, 2017, **21**, (8), pp. 1871–1874.
- [16] Interdonato, G., Karlsson, M., Björnson, E.: 'Local Partial Zero-Forcing Precoding for Cell-Free Massive MIMO', IEEE Trans. Wireless Commun., 2020, **19**, (7), pp. 4758–4774.
- [17] Björnson, E., Sanguinetti, L.: 'Scalable Cell-Free Massive MIMO Systems', IEEE Trans. Commun., 2020, **68**, (7), pp. 4247–4261.
- [18] Rezaei, F., Tellambura, C., Tadaion, A. A.: 'Rate Analysis of Cell-Free Massive MIMO-NOMA With Three Linear Precoders', IEEE Trans. Commun., 2020, **68**, (6), pp. 3480–3494.
- [19] Bashar, M., Cumanan, K., Burr, A. G., *et al*: 'On the Uplink Max–Min SINR of Cell-Free Massive MIMO Systems', IEEE Trans. Wireless Commun., 2019, **18**, (4), pp. 2021–2036.
- [20] Buzzi, S., D'Andrea, C., Zappone, A., *et al*: 'User-Centric 5G Cellular Networks: Resource Allocation and Comparison With the Cell-Free Massive MIMO Approach', IEEE Trans. Wireless Commun., 2020, **19**, (2), pp. 1250–1264.
- [21] Ngo, H. Q., Tran, L., Duong, T. Q., *et al*: 'On the total energy efficiency of cell-free massive MIMO', IEEE Transactions on Green Communications and Networking, 2018, **2**, (1), pp. 25–39.
- [22] Dao, H. T., Kim, S.: 'Effective Channel Gain-Based Access Point Selection in Cell-Free Massive MIMO Systems', IEEE Access, 2020, **68**, pp. 108127–108132.
- [23] Mai, T. C., Ngo, H. Q., Duong, T. Q.: 'Downlink Spectral Efficiency of Cell-Free Massive MIMO Systems With Multi-Antenna Users', IEEE Trans. Commun., 2020, **68**, (8), pp. 4803–4815.
- [24] Ibrahim, A. A. I., Ashikhmin, A., Marzetta, T. L.: 'Cell-Free Massive MIMO Systems Utilizing Multi-Antenna Access Points', Proc. 51st Asilomar Conf. Signals, Syst. Comput., Pacific Grove, CA, Nov. 2017, pp. 1517–1521.
- [25] Wang, J., Dai, L.: 'Asymptotic Rate Analysis of Downlink Multi-User Systems With Co-Located and Distributed Antennas', IEEE Trans. Wireless Commun., 2015, **14**, (6), pp. 3046–3058.
- [26] Joham, M., Utschick, W., Nossek, A.: 'Linear transmit processing in MIMO communications systems', IEEE Trans. Signal Processing, 2005, **53**, (8), pp. 2700–2712.
- [27] Clarke, P., de Lamare, R. C.: 'Transmit diversity and relay selection algorithms for multirelay cooperative MIMO systems', IEEE Trans. Veh. Technol., 2012, **61**, (3), pp. 1084–1098.
- [28] Boyd, S., Vandenberghe, L.: 'Convex Optimization' (Cambridge University Press, Cambridge, United Kingdom, 2019)
- [29] Watkins, D. S., Jones, M.: 'Fundamentals of Matrix Computations' (Wiley, 2014)
- [30] Fukuda, R. M., Abrão, T.: 'Linear, Quadratic, and Semidefinite Programming Massive MIMO Detectors: Reliability and Complexity', IEEE Access, 2019, **7**, pp. 29506–29519.
- [31] Tang, A., Sun, J., Gong, K.: 'Mobile propagation loss with a lowbase station antenna for NLOS street microcells in urban area'. Proc. IEEE Veh. Technol. Conf. (VTC), Rhodes, Greece, May 2001, pp. 333–336
- [32] Brown, T., Kyritsi, P., de Carvalho, E.: 'Practical Guide to the MIMO Radio Channel with Matlab Examples' (Wiley, 2012)
- [33] Tse, D., Viswanath, P.: 'Fundamentals of Wireless Communication' (Cambridge Univ. Press, 2005)
- [34] Hassibi, B., Hochwald, B. M.: 'How much training is needed in multiple-antenna wireless links?', IEEE Trans. Inf. Theory, 2003, **49**, (4), pp. 951–963.

Appendices

In the following the derivation of $\text{SINR}_{k,\text{MMSE}}$ is explained. The received signal by user k is written as

$$\begin{aligned}
 y_k &= \sqrt{\rho_f} \mathbf{g}_k^T \mathbf{P}_{\text{MMSE}} \mathbf{N}_{\text{MMSE}} \mathbf{s} + w_k \\
 &= \sqrt{\rho_f} \left(\hat{\mathbf{g}}_k' + \tilde{\mathbf{g}}_k' \right)^T \frac{f_{\text{MMSE}}}{\sqrt{\rho_f}} \left(\hat{\mathbf{G}}'^* \hat{\mathbf{G}}'^T + \frac{K\sigma_w^2}{E_{tr}} \mathbf{I}_M \right)^{-1} \hat{\mathbf{G}}'^* \mathbf{N}^{-1} \\
 &\quad \mathbf{N}_{\text{MMSE}} \mathbf{s} + w_k \\
 &= \underbrace{\sqrt{\rho_f} \hat{\mathbf{g}}_k'^T \frac{f_{\text{MMSE}}}{\sqrt{\rho_f}} \left(\hat{\mathbf{G}}'^* \hat{\mathbf{G}}'^T + \frac{K\sigma_w^2}{E_{tr}} \mathbf{I}_M \right)^{-1} \hat{\mathbf{G}}'^* \mathbf{N}^{-1} \mathbf{N}_{\text{MMSE}} \mathbf{s}}_{\text{desired signal + interference}} \\
 &\quad + \underbrace{\sqrt{\rho_f} \tilde{\mathbf{g}}_k'^T \frac{f_{\text{MMSE}}}{\sqrt{\rho_f}} \left(\hat{\mathbf{G}}'^* \hat{\mathbf{G}}'^T + \frac{K\sigma_w^2}{E_{tr}} \mathbf{I}_M \right)^{-1} \hat{\mathbf{G}}'^* \mathbf{N}^{-1} \mathbf{N}_{\text{MMSE}} \mathbf{s}}_{\text{CSI error}} \\
 &\quad + w_k.
 \end{aligned} \tag{61}$$

In (61), we know that channel coefficients, data symbols and noise are mutually independent, which allows us to show that the desired signal, the interference and the channel estimation error are mutually uncorrelated. As stated by [34] and based on the worst case uncorrelated additive noise, the achievable rate is given by $\log_2(1 + \text{SINR}_{k,\text{MMSE}})$, where

$$\text{SINR}_{k,\text{MMSE}} = \frac{\mathbb{E} [|A_1|^2]}{\sigma_w^2 + \sum_{i=1, i \neq k}^K \mathbb{E} [|A_{2,i}|^2] + \mathbb{E} [|A_3|^2]}. \tag{62}$$

The quantities in (62) are given by

$$A_1 = \sqrt{\rho_f} \hat{\mathbf{g}}_k'^T \mathbf{P}_k \sqrt{\eta_k} s_k, \tag{63}$$

which is the desired signal,

$$A_{2,i} = \sqrt{\rho_f} \hat{\mathbf{g}}_k'^T \mathbf{P}_i \sqrt{\eta_i} s_i, \tag{64}$$

which is the interference caused by user i and

$$A_3 = \sqrt{\rho_f} \tilde{\mathbf{g}}_k'^T \mathbf{P}_{\text{MMSE}} \mathbf{N}_{\text{MMSE}} \mathbf{s} \tag{65}$$

which refers to CSI error. The mean-square values of A_1 , $A_{2,i}$ and A_3 can be cast as

$$\begin{aligned}
 \mathbb{E} [|A_1|^2] &= \mathbb{E} \left[\left(\sqrt{\rho_f} \hat{\mathbf{g}}_k'^T \mathbf{P}_k \sqrt{\eta_k} s_k \right)^* \left(\sqrt{\rho_f} \hat{\mathbf{g}}_k'^T \mathbf{P}_k \sqrt{\eta_k} s_k \right) \right] \\
 &= \mathbb{E} \left[\rho_f \eta_k s_k^* \mathbf{P}_k^H \hat{\mathbf{g}}_k'^* \hat{\mathbf{g}}_k' \mathbf{P}_k s_k \right] \\
 &= \rho_f \text{tr} \left(\mathbb{E} \left[\eta_k \mathbf{P}_k^H \hat{\mathbf{g}}_k'^* \hat{\mathbf{g}}_k' \mathbf{P}_k s_k s_k^* \right] \right) \\
 &= \rho_f \text{tr} \left(\eta_k \mathbf{P}_k^H \hat{\mathbf{g}}_k'^* \hat{\mathbf{g}}_k' \mathbf{P}_k \right) \\
 &= \rho_f \eta_k \psi_k,
 \end{aligned} \tag{66}$$

$$\begin{aligned}
 \mathbb{E} [|A_{2,i}|^2] &= \mathbb{E} \left[\left(\sqrt{\rho_f} \hat{\mathbf{g}}_k'^T \mathbf{P}_i \sqrt{\eta_i} s_i \right)^* \left(\sqrt{\rho_f} \hat{\mathbf{g}}_k'^T \mathbf{P}_i \sqrt{\eta_i} s_i \right) \right] \\
 &= \mathbb{E} \left[\rho_f \eta_i s_i^* \mathbf{P}_i^H \hat{\mathbf{g}}_k'^* \hat{\mathbf{g}}_k' \mathbf{P}_i s_i \right] \\
 &= \rho_f \text{tr} \left(\mathbb{E} \left[\eta_i \mathbf{P}_i^H \hat{\mathbf{g}}_k'^* \hat{\mathbf{g}}_k' \mathbf{P}_i s_i s_i^* \right] \right) \\
 &= \rho_f \text{tr} \left(\eta_i \mathbf{P}_i^H \hat{\mathbf{g}}_k'^* \hat{\mathbf{g}}_k' \mathbf{P}_i \right) \\
 &= \rho_f \eta_i \phi_{k,i},
 \end{aligned} \tag{67}$$

and

$$\begin{aligned}
\mathbb{E} \left[|A_3|^2 \right] &= \mathbb{E} \left[\left| \sqrt{\rho_f} \tilde{\mathbf{g}}_k'^T \mathbf{P}_{\text{MMSE}} \mathbf{N}_{\text{MMSE}} \mathbf{s} \right|^2 \right] \\
&= \mathbb{E} \left[\rho_f \mathbf{s}^H \mathbf{N}_{\text{MMSE}} \mathbf{P}_{\text{MMSE}}^H \tilde{\mathbf{g}}_k'^* \tilde{\mathbf{g}}_k'^T \mathbf{P}_{\text{MMSE}} \mathbf{N}_{\text{MMSE}} \mathbf{s} \right] \\
&= \rho_f \text{tr} \left(\mathbb{E} \left[\mathbf{N}_{\text{MMSE}} \mathbf{P}_{\text{MMSE}}^H \tilde{\mathbf{g}}_k'^* \tilde{\mathbf{g}}_k'^T \mathbf{P}_{\text{MMSE}} \mathbf{N}_{\text{MMSE}} \mathbf{s} \mathbf{s}^H \right] \right) \\
&= \rho_f \text{tr} \left(\mathbf{N}_{\text{MMSE}}^2 \mathbf{P}_{\text{MMSE}}^H \mathbb{E} \left[\tilde{\mathbf{g}}_k'^* \tilde{\mathbf{g}}_k'^T \right] \mathbf{P}_{\text{MMSE}} \right) \\
&= \rho_f \sum_{i=1}^K \eta_i \gamma_{k,i}.
\end{aligned} \tag{68}$$

The final expressions form above include the following definitions

$$\psi_k = \mathbf{p}_k^H \tilde{\mathbf{g}}_k'^* \tilde{\mathbf{g}}_k'^T \mathbf{p}_k, \text{ for } k = 1, \dots, K, \tag{69}$$

$$\phi_{k,i} = \mathbf{p}_i^H \tilde{\mathbf{g}}_k'^* \tilde{\mathbf{g}}_k'^T \mathbf{p}_i, \text{ for } i \neq k, i = 1, \dots, K, \tag{70}$$

$$\gamma_k = \text{diag} \left\{ \mathbf{P}_{\text{MMSE}}^H \mathbb{E} \left[\tilde{\mathbf{g}}_k'^* \tilde{\mathbf{g}}_k'^T \right] \mathbf{P}_{\text{MMSE}} \right\}, \tag{71}$$

where $\mathbf{p}_k = [p_{1,k}, \dots, p_{M,k}]^T$ is the column k of matrix \mathbf{P}_{MMSE} , ψ_k is the k th element of vector $\boldsymbol{\psi}$, $\phi_{k,i}$ is the i th element of vector $\boldsymbol{\phi}_k$, $\gamma_{k,i}$ is the i th element of vector $\boldsymbol{\gamma}_k$, and $\mathbb{E} \left[\tilde{\mathbf{g}}_k'^* \tilde{\mathbf{g}}_k'^T \right]$ is a diagonal matrix with $\left((1 - n_{m,k}) \beta_{mk}' \right)$ on its m th diagonal element. After substituting (66), (67) and (68) in the $\text{SINR}_{k,\text{MMSE}}$ expression we obtain

$$\text{SINR}_{k,\text{MMSE}} = \frac{\rho_f \eta_k \psi_k}{\sigma_w^2 + \rho_f \sum_{i=1, i \neq k}^K \eta_i \phi_{k,i} + \rho_f \sum_{i=1}^K \eta_i \gamma_{k,i}}. \tag{72}$$