Learning and Nonlinear Models - Journal of the Brazilian Society on Computational Intelligence (SBIC), Vol. 23, Iss. 1, pp. 4-18, 2025
https://doi.org/10.21528/lnlm-vol22-no2-art1

SUM-RATE MAXIMIZATION IN MASSIVE MIMO WITH JOINT ANTENNA SELECTION AND POWER ALLOCATION

Jonathan N. Gois [1,2], **Matheus S. P. Silva** [1], **Vinicius N. H. Silva** [2], **Tadeu N. Ferreira** [2]
[1] Federal Center for Technological Education Celso Suckow da Fonseca (Cefet/RJ), Brazil,
[2] Fluminense Federal University (UFF), Brazil
jonathan.gois@cefet-rj.br, matheus.perdigao@aluno.cefet-rj.br, {viniciusnhs, tadeu_ferreira}@id.uff.br

Abstract – Telecommunications improvements in the last decades have led to an increase in the available data rate and reliability, and to reduce latency. Massive MIMO (mMIMO) has emerged as a promising solution to replace simple antenna setups. This arrangement involves a large number of transmitter antennas compared to the served terminals in the covered area, enabling higher data rates through the key characteristics of mMIMO. Despite its advantages, the utilization of such large arrays is energy-intensive. With fewer antennas, it is possible to achieve high data rates with reduced power consumption. Antenna selection is crucial for energy efficiency while maintaining computational efficiency. Additionally, power allocation in the downlink for each served terminal is an important step to consider. This work aims to propose a joint algorithm for selecting antennas and distributing energy among terminals, simplifying the process through low-complexity algorithms. The results show the penalty in sum-rate capacity for each proposed method. The two proposed algorithms demonstrate an adequate approach to the target value asymptotically as the number of transmitter antennas increases. This implies the possibility of achieving the joint process with significantly lower computational costs compared to perform antenna selection and power allocation separately.

Keywords – Massive MIMO; Antenna Selection; Power Allocation; Sum-Rate

1 INTRODUCTION

The evolution of 5G systems entails various techniques aiming at consistently enhancing transmission rates, reliability, while reducing latency. Massive Multiple-Input and Multiple-Output (mMIMO) stands out as a pivotal element, which unlocks numerous benefits [1–3]. Among an extensive list of benefits associated with increasing the number of transmitter antennas, two notable advantages stand out: the possibility of a higher data throughput and the simplicity in signal processing. A mMIMO arrangement enables the achievement of higher spatial multiplexing and energy comsumption efficiency under certain conditions.

mMIMO is usually designed with two different setups [4]. A co-located mMIMO aims to deploy the large array in a compact area, leading to a lower cost backhaul. Another scheme is distributed mMIMO [5], which places the antennas across a wide space to ensure broader user coverage. On the other hand, there is an increased cost of the backhaul and all hardware required to deploy the antennas in different locations.

In mMIMO schemes, an increment of antennas in the transmitter array improves the spectral efficiency of the overall system by exploiting spatial diversity and multiplexing. The growing number of transmitter antennas relative to users in a coverage area increases data throughput, reduces relative power consumption for transmission, and enables effective power control policies to ensure good service for all users in the cell [6].

Considering the mMIMO arrangement and its spatial multiplexing achieved for the complete antenna array, some antennas may perform worse than others in relation to channel quality. As a result, the usage of these antennas can be inefficient or even unproductive. Therefore, an increase in the number of antennas leads to higher energy expenditure compared to a selected subset of the best antennas in the array. The full array can be prohibitive in terms of energy cost for commercial purposes. Some options are proposed in the literature to reduce the energy consumption and hardware cost in mMIMO. Firstly, each independent antenna is connected to a Radio Frequency (RF) chain, which converts digital signals into analog electromagnetic waves. As the number of transmitter antennas grows, the number of dedicated RF chains should also be increased. The transceiver hardware may then become cost-prohibitive and power-hungry [7]. The first branch of solutions can be explored for this reason. By reducing the number of transmitter antennas, it is possible to reduce the number of RF chains. This solution is not adequate, since the benefits of spatial resolution are diminished when antennas are removed. Therefore, the sub-array RF architecture aims to connect a set of antennas to the same RF chain. This solution maintains spatial coverage, while it reduces signal diversity. The other possible solution branch is the Full-Array Antenna Selection (FAS) Architecture. The purpose of this approach is to maintain signal diversity and implement more transmitter antennas to RF chains. In the system hardware, it is possible to switch the connections between antennas and RF chains. This approach enables the implementation of a number of antennas greater than the amount of RF chains, ensuring spatial resolution without increasing the cost of RF chains.

Some criteria for selecting antennas have been explored in the literature. This can be accomplished using an error metric, such as the average bit-error rate (BER) among the served users. Several papers [8, 9] have exploited the minimization of BER. Other works focus on the optimization of a capacity-oriented criterion [10, 11]. Information capacity refers to the maximum spectral efficiency achieved for a given arrangement, that is, the maximum data rate that can be supported by the system. This class of works involves two main parameters: the selected antennas in the array and the power allocated to each antenna. Most works optimize either the selection of antennas in the array or the allocation of power to each antenna to enhance capacity.

A common challenge for all objectives is the computational complexity involved in the selection and power allocation algorithms [12]. The environment undergoes constant changes, as do the optimization parameters. Attempting to execute all steps in the downlink transmission within a time-limited window can be impractical. The energy consumption of complex algorithms can be considered either infeasible or impracticable.

1.1 LITERATURE REVIEW

Marzetta *et al.* proposed the concept of mMIMO in [1], by using setups for a very large number of antennas. The antenna selection algorithm was conceived before the advent of mMIMO, within the framework of a traditional MIMO environment.

Recalling the benefits of mMIMO, the spatial multiplexing achieved by this scheme makes it possible to utilize millimeter waves without line-of-sight between a transmitter and a receiver. A wider bandwidth increases the channel capacity to transmit data, thereby enabling higher data rates to be achieved. Nevertheless, this bandwidth is affected by path loss, scattering, and penetration loss [13]. To overcome this obstacle, a large antenna array improves beam directionality. One more advantage of antenna selection can be exploited in this context. Antenna selection algorithms can be used to enhance beam directionality, which saves the energy spent on sidelobes.

Antenna selection serves as a low-cost low-complexity alternative to harness many advantages of mMIMO. Sanayei *et al*. [10] review numerous articles on this topic, providing an insightful overview. The two methods proposed in that work are based on maximizing either spatial multiplexing or channel capacity. These methods are compared with full antenna array and an optimal solution. In summary, several research branches are addressed, with the most significant point being that antenna selection algorithms still require further investigation.

For reducing computational complexity in antenna selection (AS) algorithms, the QR decomposition was presented as an option [14]. Disregarding the power in each antenna, the quest of maximizing channel capacity was addressed by decomposing the MIMO channel matrix using QR decomposition. They achieved near-optimal capacity using this approach and reduce complexity compared to full-search AS.

The channel capacity was maximized in [11] through convex optimization. The results of this method were very close to the optimal solution based on exhaustive search, with the advantage of a lower complexity algorithm. The authors tested various experimental setups to measure the channel capacity. Both linear and cylindrical arrays were employed, with both arrays operating at a center frequency of 2.6 GHz and a bandwidth of 50 MHz. In their results, the comparison of the number of RF chains for the total number of antennas was made. With a similar experimental setup, the same research group [15] proposed a method to maximize sum-rate. In that work, they investigated the possibility of reducing computational complexity and hardware energy consumption. They included additional experimental results varying the number of users and scenarios, and tested more AS methods.

Some works attempted to address the computational complexity problem using different approaches. Amadori *et al*. [16] proposed replacing channel capacity maximization with constructive interference maximization. The idea was to select antennas that maximize the power received at the terminals. Chen *et al*. [17] converted the problem of maximizing channel capacity into a decision-making problem solved by a Monte Carlo Tree Search algorithm.

To reduce the complexity of channel capacity convex optimization, other works have proposed some bio-inspired methods, such as, Tabu search [18], Quantum-inspired [19], and Hybrid Sea Lion-Whale Algorithm [20]. The idea of applying bio-inspired methods is to reduce complexity by replacing convex optimization, and to reduce convergence time.

Other works have discussed the implementation of channel capacity maximization as a parallel problem. Mendonça *et al*. [9] proposed an application of greedy AS algorithms. Matching Pursuit Antenna Selection (MPAS) was employed by replacing the channel capacity expression with a simpler optimization problem solved using the matching pursuit technique.

An interesting strategy was outlined in [21]. By separating the whole set of antennas into groups, it becomes possible to select the best antenna group for each terminal. When the number of antennas increases in each group, the selection algorithm was simplified. Their algorithm aimed to maximize the Signal-to-Interference-plus-Noise Ratio (SINR) which is related to maximizing channel capacity. Other works [22, 23] have also proposed the same approach using a greedy algorithm.

Despite the various methods proposed in the literature for antenna selection and power allocation problems [24], an optimization following a joint approach of AS and power allocation is often avoided. As elucidated by Hao Li *et al*. [25], the interrelated variables involved in joint optimization make it impractical to directly derive a closed-form solution for this problem. The authors propose an iterative solution to reduce optimization complexity. Throughout the iterations, some antennas are selected, and power allocation is solved alternately until convergence. Another iterative solution was proposed in the literature. A Quantum-inspired backtracking [26] search was recently proposed to reduce the number of iterations. In spite of this, both methods have high computational complexity, creating a barrier to real-time implementation.

Regarding power efficiency and secure data transfer, the work of Gharagezlou *et al*. [27, 28] presents a joint solution for AS and power allocation with cell division. This paper aims to perform optimal power allocation considering a corrupted scenario and a minimum downlink rate. An iterative solution is derived using Lagrangian multipliers. This article disregards the complexity of those algorithms.

1.2 CONTRIBUTIONS OF THIS PAPER

Due to the advances in mMIMO technologies and the definitions of the standards of the 5G evolution, this work aims to enhance the algorithms responsible for jointly selecting the active antennas and allocating power to each one of them. To achieve this objective, MPAS and its derivations are presented as the solution to an enhanced antenna selection scheme. It incorporates an additional step to improve channel capacity by allocating energy more effectively than equal power allocation, without requiring additional computational resources. Additionally, this article proposes a simplified particle swarm optimization for a joint AS and power allocation.

The main contributions of this work are then listed:

- A mMIMO model environment is considered, by incorporating both antenna selection and power allocation procedures, and its analysis is conducted in terms of channel capacity and sum-rate capacity;
- Power allocation using a simple version of PSO, which is extended to facilitate the joint operation;
- An extension of MPAS is proposed, along with the joint algorithm incorporating it;
- A comprehensive comparison of all methods is conducted, incorporating each step, one by one, and the results are analyzed.

1.3 ORGANIZATION

This paper is organized as follows. In this section, we have provided the introduction of this article. Section 2 presents the mMIMO system model, discusses some restrictions for the proposed solution, and formulates the problem to be solved. The proposed methods are described in Section 3, followed by their results in Section 4. Finally, we draw the main conclusions in Section 5.

2 SYSTEM MODEL AND PROBLEM FORMULATION

Consider an mMIMO environment in the downlink in a single cell. The base station (BS) is associated to a coverage cell and has M transmitter antennas, operating in a full-array scheme. This BS serves K single-antenna terminals. To meet the mMIMO requirements, we consider $M \gg K$, and we assume that neighboring cells do not interfere in this model. The choice of a single-cell model without interference from neighboring cells is widely adopted in order to investigate some key aspects of massive MIMO, avoiding external influences. Especially for power allocation, the multi-cell scheme adds inter-cell interference in addition to intra-cell interference. Other aspects need to be further investigated, such as pilot allocation, pilot reuse, channel estimation, and cooperation between neighboring cells. The single-cell assumption considers only intra-cell interference, excluding external factors for the time being.

As a premise [29], we consider a time division duplex (TDD) scheme, where the time-frequency resources are equally shared among the users' equipments (UE). The MIMO downlink channel between the antennas and the UEs needs to be estimated. The channel state information (CSI) is used as this estimation. By applying TDD, channel reciprocity can be assumed. This assumption helps reduce the overhead in the coherence interval for uplink and downlink pilot sequences. It can be considered when there are no imperfections in the hardware on either side [30].

To further simplify this problem, we assume a known CSI at the base station. All the methods tested in this article require an estimate of the channel, then considering knowledge of the channel do not favor any algorithm in the comparison in this paper.

The mMIMO channel $\mathbf{G}$, in a given time-frequency resource, is modeled as [29]:

$$\mathbf{G} = \mathbf{HBD}^{1/2}. \tag{1}$$

In equation (1), $\mathbf{H} \in \mathbb{C}^{M\times K}$ is the small-scale fading matrix, and $\mathbf{D} \in (\mathbf{R}_{+}^{*})^{K\times K}$ denotes the large-scale fading components as a diagonal matrix where the element $d_{kk} = \beta_k > 0,\ 1 \leq k \leq K$, is the path loss for the k-th UE. The matrix $\mathbf{B} \in \mathbb{C}^{K\times K}$ represents a phase shift that is dependent of UE positioning.

Assuming an uniform rectangular array (URA), it is possible to model [31,32] the small-scale fading components, considering the spatial correlation of these antennas, as:

$$h_{mk} = e^{-\frac{i2\pi d_{mk}}{\lambda}},\ 1 \leq m \leq M \text{ and } 1 \leq k \leq K, \tag{2}$$

where λ is the carrier wavelength and d_{mk} is a distance between the antenna m and the user k modeled as [32]:

$$d_{mk} \approx d_k - \sin(\theta_{mk})\left((p_x-1)\delta\cos(\phi_{mk}) + (p_y-1)\delta\sin(\phi_{mk})\right). \tag{3}$$

The distance approximation d_{mk} in equation (3) is a function of a radial distance d_k, the azimuth angle position defined by ϕ_{mk} and the zenith angle θ_{mk}. Each d_{kl} depends on (p_x, p_y), that is the m-antenna position. The antenna spacing δ are considered in the simulation results, since it directly affects the spatial correlation [33].

Learning and Nonlinear Models - Journal of the Brazilian Society on Computational Intelligence (SBIC), Vol. 23, Iss. 1, pp. 4-18, 2025
https://doi.org/10.21528/lnlm-vol22-no2-art1

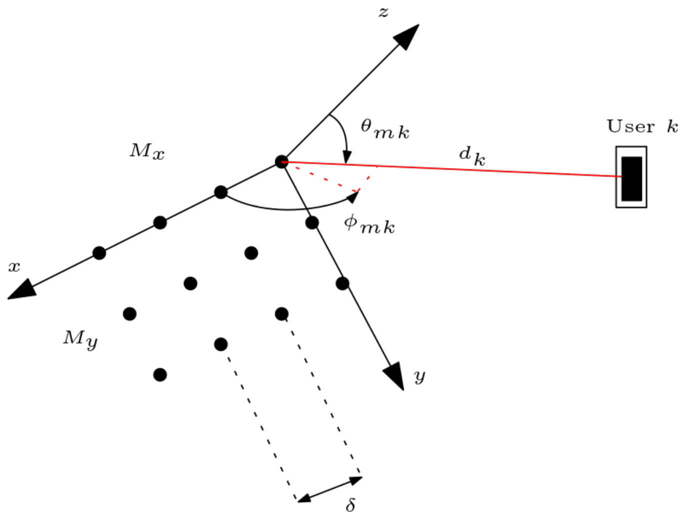


Figure 1: Illustration of URA antenna model.

Considering a free-space path loss environment, the digital terms corresponding to large-scale fading components are given by:

$$\beta_{mk} = \left(\frac{\lambda}{4\pi}\right)^2 \frac{1}{d_{mk}^2}. \tag{4}$$

To mitigate the distortions caused by the channel, before sending the signal to UEs, the BS performs precoding. We apply a zero-forcing precoding [34], where the precoding matrix is given by $\mathbf{P} = \mathbf{G}(\mathbf{G}^{\mathrm{H}}\mathbf{G})^{-1}$. The received signal at each UE is:

$$y_k = \underbrace{\sqrt{\alpha\rho_{dl}\eta_k}\mathbf{g_k}^{\mathrm{H}}\mathbf{p_k}q_k}_{\text{desired signal}} + \underbrace{\sum_{i=1, i\neq k}^{K} \sqrt{\alpha\rho_{dl}\eta_k}\mathbf{g_k}^{\mathrm{H}}\mathbf{p_i}q_i}_{\text{interference}} + \omega_k. \tag{5}$$

In Equation (5), vector $\mathbf{g}_k$ corresponds to the k-th column of $\mathbf{G}$. This vector represents the channel between the M antennas and the UE k. Besides that, $\mathbf{p}_k$ is the k-th column of $\mathbf{P}$. The message to be sent to UE is represented by the symbol q_k.

The channel is corrupted by additive Gaussian noise, depicted as ω_k. The power allocation for each terminal is represented by η_k where $0 < \eta_k < 1$. This factor represents the proportion of the total power spent in the downlink transmission, denoted by ρ_{dl} . Variable $\alpha = 1/\mathbb{E}\{\mathrm{tr}(\mathbf{P}^{\mathrm{H}}\mathbf{P})\}$ is a normalization factor as described in [25]. With the addition of this variable, the precoding matrix has a unitary norm, and the power allocation occurs solely based on the factor η_k. We can aggregate all received symbols as:

$$\mathbf{y} = \sqrt{\alpha}\rho_{dl}\boldsymbol{\eta}\mathbf{G}^{\mathrm{H}}\mathbf{P}\mathbf{q} + \boldsymbol{\omega}, \tag{6}$$

where $\mathbf{q}$ is a message vector for all users, $\boldsymbol{\omega}$ is the noise vector and the power allocation matrix is given by:

$$\boldsymbol{\eta} = \begin{bmatrix} \sqrt{\eta_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\eta_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\eta_K} \end{bmatrix}. \tag{7}$$

When the number of transmitter antennas is much larger than number of terminals, a phenomenon occurs known as "channel hardening" [35], which is expressed mathematically as

$$\frac{||\mathbf{g}_k||_2^2}{\mathbb{E}\{||\mathbf{g}_k||_2^2\}} = 1, \text{ when } M \to \infty. \tag{8}$$

A more audacious assumption for large arrays is known as favorable propagation [2], which is related to the distribution of small-scale fading. When these coefficients are assumed independent and identically distributed with zero mean, it leads to $\mathbf{H}^{\mathrm{H}}\mathbf{H} = \mathbf{I}_M$, where $\mathbf{I}_M$ is the $M \times M$ identity matrix. Therefore [2],

$$\frac{1}{M}\mathbf{G}^{\mathrm{H}}\mathbf{G} = \frac{1}{M}\mathbf{D}^{1/2}\mathbf{H}^{\mathrm{H}}\mathbf{H}\mathbf{D}^{1/2} \approx \mathbf{D}. \tag{9}$$

The assumptions of favorable propagation and channel hardening for millimeter wave have been investigated in the literature based on measured propagation scenarios [35]. These studies indicate that these hypotheses may be valid under certain conditions.

2.1 Channel Capacity

The capacity of an mMIMO channel is obtained through the generalization of the channel capacity concept [36], which represents the average amount of information that can be recovered after transmission by observing the received signal. Consider a known mMIMO CSI, where this information is attenuated or corrupted. Moreover, consider a point-to-point MIMO link, the channel capacity which, which is the maximum value of mutual information, between input and output can be described as [2]:

$$C_{\text{total}} = \max_{\boldsymbol{\eta}} \log_2 \det \left(\mathbf{I}_K + \rho_{dl} (\mathbf{G}\boldsymbol{\eta})(\mathbf{G}\boldsymbol{\eta})^{\mathrm{H}} \right). \quad (10)$$

Capacity C_{total} represents the amount of information that can be transmitted per second, considering the available bandwidth. Commonly, matrix $\boldsymbol{\eta}$ is omitted from Equation (10) by considering the optimal solution to power allocation achieved by the water-filling method, or by considering equal power allocation, where $\boldsymbol{\eta}$ is an identity matrix. The mMIMO channel capacity is expressed as a function of SINR as:

$$C_{\text{total}} = \sum_{k=1}^{K} \log_2(1 + \Gamma_k), \quad (11)$$

where Γ_k is the SINR referred to at terminal k.

Channel capacity can be interpreted as the upper limit for the sum of achievable downlink rates R_k:

$$C_{\text{total}} \geq \sum_{k=1}^{K} R_k = \sum_{k=1}^{K} \log_2(1 + \Gamma_k). \quad (12)$$

For a given arrangement of power allocation per user. This quantity is known as the sum-rate, which is an approximation of the channel capacity when $\boldsymbol{\eta}$ is maximized.

Insights may be derived from the hypothesis of channel hardening and favorable propagation. The sum-rate capacity C_{SR} equation can be reformulated as follows [37]:

$$C_{SR} = \sum_{k=1}^{K} \log_2(1 + \Gamma_k) = \sum_{k=1}^{K} \log_2 \det \left(\mathbf{I}_K + \frac{\rho_{dl}}{M} (\mathbf{G}\boldsymbol{\eta})(\mathbf{G}\boldsymbol{\eta})^{\mathrm{H}} \right), \quad (13)$$

$$C_{SR} \approx \sum_{k=1}^{K} \log_2(1 + \rho_{dl}\eta_k ||g_{kk}||^2). \quad (14)$$

To ascertain the fitness of this approximation, one evaluates the difference between the sum-rate approximation and the channel capacity [37]:

$$\Delta_{\text{capacity}} = \frac{\sum_{k=1}^{K} \log_2(\mathbf{I}_K + \rho_{dl}\eta_k ||g_{kk}||^2) - \log_2 \det(\mathbf{I} + \rho_{dl}(\mathbf{G}\boldsymbol{\eta})(\mathbf{G}\boldsymbol{\eta})^{\mathrm{H}}))}{\log_2 \det(\mathbf{I} + \rho_{dl}(\mathbf{G}\boldsymbol{\eta})(\mathbf{G}\boldsymbol{\eta})^{\mathrm{H}}))}, \quad (15)$$

for a given power allocation.

Sum-rate is a means to finely tune the transmission parameters to achieve better transmission. It is used in this work as qualitative metric to the proposed algorithms. To prove the possibility of switching the metrics, we conducted a simulation experiment [38] to understand the distance between the metrics using the gap to capacity. That study shows that, when the number of antennas increases for the channel used in this work, the gap asymptotically approaches zero.

2.2 Antenna Selection

Antenna selection solution offers the potential to reduce costs by saving power and producing cheaper hardware without a harsh penalty to the spatial multiplexing achieved through the use of multiple antennas. We introduce an additional adjustable variable to our problem by using the following equation:

$$\mathbf{G}_s = \mathrm{rem}(\mathbf{Z}\mathbf{G}), \quad (16)$$

where $\mathbf{Z} \in \{0, 1\}^{M \times M}$ is a diagonal matrix. The value z_{mm} is defined as the Boolean value 1, when the m-th antenna is activated, or 0 when it is not. The operator $\mathrm{rem}(\cdot)$ is responsible for removing the null rows or columns. Considering S activated antennas, the main diagonal of $\mathbf{Z}$, denoted as $\mathbf{z}$, has a ℓ_0-norm equal to S [9, 11].

Vector $\mathbf{z} \in \{0, 1\}^M$ is the output of an antenna selection algorithm that typically aims to maximize the capacity or minimize the bit error rate (BER) in downlink transmission.

When considering the downlink, all previous equations can be adjusted by replacing the matrix $\mathbf{G}$ by $\mathbf{G}_s$, taking into account changes in dimensions and some adjustments to the precoding matrix.

The antenna selection optimization aims to minimize the degradation of channel capacity, while reducing the number of activated antennas. Disregarding the power allocation, the general problem can be described as [11]:

$$\begin{aligned} \text{maximize} \quad & \log_2 \det(\mathbf{I}_K + \frac{\rho_{dl}}{M}\mathbf{G}_s\mathbf{G}_s^{\mathrm{H}}), \\ \text{subject to} \quad & \mathbf{G}_s = \mathbf{Z}\mathbf{G}, \\ & ||\mathbf{z}||_0 = S, \\ & z_i \in \{0,1\} \text{ for i} = 1,\dots,\text{M}. \end{aligned} \tag{P1}$$

Problem (P1) can be simplified under the favorable propagation hypothesis and by considering the sum-rate metric as follows:

$$\begin{aligned} \text{maximize} \quad & \sum_{k=1}^{K} \log_2(\mathbf{I}_K + \frac{\rho_{dl}}{K}||[\mathbf{G}_s]_{kk}||^2), \\ \text{subject to} \quad & ||\mathbf{z}||_0 = S, \\ & z_i \in \{0,1\} \text{ for i} = 1,\dots,\text{M}. \end{aligned} \tag{P2}$$

where we consider an equal power allocation to compute the sum-rate and the operator $[\cdot]$ extracts the elements from the main diagonal of its operand.

2.3 Power Allocation

Assuming the antenna selection problem is resolved, a power allocation optimization is conducted to maximize the sum-rate as follows:

$$\begin{aligned} \text{maximize} \quad & \log_2 \det(\mathbf{I}_K + \rho_{dl}\mathbf{G}_s\boldsymbol{\eta}^2\mathbf{G}_s^{\mathrm{H}}), \\ \text{subject to} \quad & 0 < \eta_k < 1 \text{ for k} = 1,\dots,\text{K}, \\ & \sum_{k}^{K} \eta_k = 1. \end{aligned} \tag{P3}$$

The problem of antenna selection and power allocation is a non-separable optimization problem. Nevertheless, some works present a solution for each step, disregarding the correlation between the variables. The joint antenna selection and power allocation, embracing the mMIMO advantages, is described as:

$$\begin{aligned} \text{maximize} \quad & \sum_{k=1}^{K} \log_2(\mathbf{I}_K + \rho_{dl}\eta_k^2||[\mathbf{G}_s]_{kk}||^2), \\ \text{subject to} \quad & ||\mathbf{z}||_0 = S, \\ & z_i \in \{0,1\} \text{ for i} = 1,\dots,\text{M}, \\ & 0 < \eta_k < 1 \text{ for k} = 1,\dots,\text{K}, \\ & \sum_{k}^{K} \eta_k = 1. \end{aligned} \tag{P4}$$

A cautious approach is performed when the optimization uses the small-scale fading channel components. This channel capacity approximation leverages the benefits of mMIMO and reduces the computational complexity required to solve the antenna selection, power allocation, and the joint problem.

3 PROPOSED ALGORITHMS

To implement the proposed solution, we divide the problem as presented in the literature. Firstly, we propose using antenna selection based on matching pursuit, considering the full problem without the key properties of mMIMO. After that, we propose a power allocation solution, considering the resolved problem of antenna selection, to maximize the sum-rate. In this proposed method, we incorporate modifications to the particle swarm optimization (PSO) algorithm to further reduce computational complexity. The proposed Particle Swarm Optimization with Halving (PSOH) aims to select the initial particles effectively and iteratively remove half of them. Lastly, we propose a joint antenna selection and power allocation algorithm. The consideration of the problem's complexity is taken into account and subsequently compared.

3.1 Hierarchical Matching Pursuit Antenna Selection

Hierarchical Matching Pursuit Antenna Selection (HMPAS) represents an enhanced version of the Matching Pursuit Antenna Selection proposed by Mendonça *et al*. [9]. Our research group has previously proposed simplifications to this problem, namely the Subarray MPAS (SMPAS) [22], to achieve lower computational complexity. By introducing MPAS, we can rewrite the optimization problem using a greedy algorithm as follows:

$$\begin{aligned} &\text{minimize} && ||\mathfrak{D}\mathbf{z} - \mathbf{b}||_2^2 \\ &\text{subject to} && ||\mathbf{z}||_0 = S. \\ & && z_i \in \{0,1\} \text{ for } i = 1, \ldots, M. \end{aligned} \tag{17}$$

where $\mathfrak{D}$ is a dictionary matrix and $\mathbf{b}$ is the target vector to approximate. The combinations of columns of $\mathfrak{D}$ that construct $\mathbf{b}$ are represented by the Boolean vector $\mathbf{z}$. In other words, we transform the maximization problem (P1) into a minimization problem as described by Equation (17).

By ignoring channel hardening temporarily, one can describe the dictionary for channel-level antenna selection in accordance with [9] as:

$$\mathfrak{D} = \begin{bmatrix} \text{vec}(\mathbf{p_1}\mathbf{p_1}^{\text{H}}) & \text{vec}(\mathbf{p_2}\mathbf{p_2}^{\text{H}}) & \ldots & \text{vec}(\mathbf{p_M}\mathbf{p_M}^{\text{H}}) \end{bmatrix} \in \mathbb{C}^{K^2 \times M}, \tag{18}$$

where operator $\text{vec}(\cdot)$ represents the column stacking operation used to transform the matrix $\mathbf{p_m}\mathbf{p_m}^{\text{H}} \in \mathbb{C}^{K\times K}$ $(0 < m \leq M)$ into a vector. The vectors $\mathbf{p_i}$ are the column vectors from the precoder $\mathbf{P}$.

Another important variable in Equation (17) is the target vector, which represents the full-antenna channel [9]:

$$\mathbf{b} = \text{vec}(\mathbf{P}\mathbf{P}^{\text{H}}) \in \mathbb{C}^{\text{K}^2}. \tag{19}$$

Following the described scenario, the MPAS are solved using an iterative method, which is detailed in Algorithm 1.

Algorithm 1 MPAS algorithm

---

1. Input: dictionary $\mathfrak{D}$ and the target vector $\mathbf{b}$;
2. Creating an index set $\mathcal{I} = \{1, \ldots, M\}$;
3. Initialize the counter $i \leftarrow 1$, the residue vector $\mathbf{r}_1 \leftarrow \mathbf{b}$ and $\mathbf{z} \leftarrow \mathbf{0}_M$;
4. Iterative loop: while $i \leq S$ do:
   a. Find the best approximation to $\mathbf{r}_i$ in the dictionary $\mathfrak{D}$:
   $l \leftarrow \arg\max_{i\in\mathcal{I}} \text{tr}(\mathbf{r}_i^{\text{T}}\mathfrak{D})$.
   b. Make $z_l \leftarrow 1$;
   c. Update the residue $\mathbf{r}_{i+1} \leftarrow \mathbf{r}_i - \max_{i\in\mathcal{I}} \text{tr}(\mathbf{r}_i^{\text{T}}\mathbf{D})$;
   d. Remove the used element in index set $\mathcal{I}$;
   e. Increment counter: $i \leftarrow i + 1$.
5. Output $\mathbf{z}$.

The original method suggest a factor γ, where $0 < \gamma \leq 1$, to control the residue energy decay that can be implemented in step (c). Without loss of generality, this factor is set as $\gamma = 1$.

HMPAS emerges with the concept that antennas are physically allocated in an array, and that closely situated antennas could have similar channel representations. We can directly derive the solution by assuming that channel indexes are organized according to their physical placement in the array, or we can search for similarities in the channel representation. The first approach is denominated Subarray Matching Pursuit Antenna Selection (SMPAS) [22] and the second one is the HMPAS [23].

To overcome the barrier of searching for the best index within the dictionary, we propose a two-step search process. Firstly, we identify the best correspondence in a representative of all subdictionaries. Search is performed within the subdictionary to find the best match for HHMPAS or in the entire subdictionary to simplify the process for SMPAS. To achieve the best results to maximize channel capacity, we proceed with the HMPAS, which achieves results that are similar to the MPAS but with reduced complexity.

With Algorithm 2, we drastically reduce the computational complexity in accordance to the increasing number of subdictionaries C as demonstrated in [23]. In addition to this, we demonstrate here that it is possible to include power allocation to reduce the number of activated antennas without loss of channel capacity. By adjusting the remaining power, it is feasible to achieve the same channel capacity with a larger arrangement.

Algorithm 2 HMPAS algorithm

1. Input: dictionary $\mathfrak{D}$ and the target vector $\mathbf{b}$;
2. Split the dictionary into C subdictionaries, $\mathfrak{D} = \{\mathfrak{D}_1, \mathfrak{D}_2, \ldots, \mathfrak{D}_C\}$. For each subdictionary, select a representative candidate $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_C\}$
3. Initialize the counter $i \leftarrow 1$, the residue $\mathbf{r}_1 \leftarrow \mathbf{b}$ and $\mathbf{z} \leftarrow \mathbf{0}$;
4. Iterative loop: while $i \leq S$ do:
   a. Perform the MPAS to choose the most representative j-th subdictionay $\mathfrak{D}_j$, the vector $\mathbf{c}_j$;
   b. Using $\mathfrak{D}_j$, perform the MPAS to select one antenna using $\mathbf{r}_i$ as approximation vector;
   c. Remove the used vector in $\mathfrak{D}_j$. If their cardinality is zero, also remove the subdictionary from $\mathfrak{D}$;
   d. Update the residue: $\mathbf{r}_{i+1} \leftarrow \mathbf{r_i} - \max_{i\in\mathcal{I}} \mathrm{tr}(\mathbf{r}_i^{\mathrm{T}}\mathfrak{D}_j)$;
   e. Increment counter: $i \leftarrow i+1$;
5. Output $\mathbf{z}$.

3.2 Power allocation to Multiple Antenna Arrays

Following the same procedure outlined in the previous section, we can conduct an user power allocation while disregarding the AS algorithm. To understand the power allocation strategies, different types of fairness involved in this process are discussed. The simplest algorithm to distribute the total energy expenditure among all users is the equal power allocation. As the name suggests, this method equally divides the total energy. In this case, the concept of fairness is based on a power constraint.

Based on data rate, the max-min fairness power allocation (MMFPA) is constructed to guarantee egalitarian rate performance for each user. Attempting to achieve equal data rates for all served users, the MMFPA may expend a significant amount of energy on users with poor channel conditions that deliver low data rates. This can result in a reduced overall data rate when bad channels are present. Consequently, the proportional fairness power allocation (PFPA) is proposed to offer non-uniform service but optimize the overall data rate.

We can balance the individual and overall performance by allocating power according to the incremental gain in SINR achieved. In this case, equilibrium is based on equal SINR gain, which is proportional to the data rate. An water-filling power allocation (WFPA) is the solution for this approach, by allocating energy based on the incremental gain in SINR for each user.

Other different metrics can be used here to ensure fair power allocation. In this work, we focus on achieving the MMFPA. We can rewrite the objectives of the MMFPA using the Problem P5:

$$
\begin{aligned}
&\text{maximize} && \min(\Gamma_k) = \rho_{dl}\eta_k^2 ||[\mathbf{G}_s]_{kk}||^2, \\
&\text{subject to} && 0 < \eta_k < 1, \quad k = 1, \ldots, K, \\
& && \sum_k^K \eta_k = 1.
\end{aligned} \tag{P5}
$$

Problem P5 disregards the influence of neighboring cells and represents a simplification of the most common approach used in the literature. To solve the MMFPA problem, we can use the bisection algorithm as shown in Algorithm3. Here, we will apply a modified version of MMFPA to reduce the search interval and ensure convergence, as proposed in [39].

The MMFPA algorithm depends on a previously selected antenna array $\mathbf{G}_s$. Moreover, MMFPA operates based on large-scale fading. The challenge here is to reduce the conjugate complexity of the antenna selection algorithm and the power allocation strategy.

3.3 Joint Antenna Selection and Power Allocation based on Particle Swarm Optimization with Halving

A joint solution begins with a natural optimization approach that guarantees convergence as the number of iterations approaches infinity with an optimal adjustment of parameters settings. Particle swarm optimization (PSO) is a popular strategy for optimizing non-convex functions without a closed-form solution. Its ability to exploit the solution space often yields good solutions within a limited execution time.

To adjust the spent execution time, in this work, an additional step is included that is labeled here as “halving". The traditional PSO algorithm begins with a fixed number of particles. All particles are evaluated within the optimization function. In our case, the sum-rate expression is depicted in Equation (14). The iterations involve two behaviors: finding the local best optimal position and adhering to community consensus. The local best position pertains to the individual behavior of each particle, adding memory

Algorithm 3 MMFPA algorithm using bisection proposed in [39].

1. Input: number of users K, small-scale fading matrix $\mathbf{H}$ and the large-scale fading components $[\mathbf{G}_s]_{kk}, 1 \leq k \leq K$;
2. Define a very small positive tolerance δ as a thresold for convergence criterion;
3. Calculate the column-wise normalized version of small-scale fading matrix: $\tilde{\mathbf{H}} \leftarrow \text{diag}\left(\frac{1}{||\mathbf{h}_1||_2}, \ldots, \frac{1}{||\mathbf{h}_K||_2}\right)\mathbf{H}$;
4. Create the matrix $R \leftarrow |\tilde{\mathbf{H}}|^2 - \mathbf{I}_K$;
5. Initialize a vector corresponding to initial rates: $\boldsymbol{\gamma} \leftarrow [\mathbf{G}_s]_{kk}\text{diag}\left([||\mathbf{h}_1||_2^2, \ldots, ||\mathbf{h}_K||_2^2]\right)^{\mathrm{T}}$;
6. Define an auxiliary variable $\mathbf{s} \leftarrow [1/\gamma_1, \ldots, 1/\gamma_K]^{\mathrm{T}}$
7. Initialize the upper and lower bound rates as $\gamma_{lower} \leftarrow 0$ and $\gamma_{upper} \leftarrow 1/\max(\text{eig}(\text{R}))$ where $\text{eig}(\cdot)$ returns the eigenvalues of the operands;
8. While $|\gamma_{upper} - \gamma_{lower}| > \delta$ do:
   a. Update the vector of rates $\gamma_m \leftarrow (\gamma_{upper} + \gamma_{lower})/2$
   b. Update the vector of power allocation values $\boldsymbol{\eta} \leftarrow (\mathbf{I}_K/\gamma_m - R)^{-1}\mathbf{s}$
   c. If $|\boldsymbol{\eta}|_1 \leq 1$:
      - Update the lower bound: $\gamma_{lower} \leftarrow \gamma_m$;

      Else
      - Update the upper bound: $\gamma_{upper} \leftarrow \gamma_m$;
9. Output $\boldsymbol{\eta}$.

to each individual. The community consensus identifies the best particle, providing guidance for each particle towards the best result found.

The halving step aims to reduce complexity in this entire setup. Starting with a large number of particles, each iteration evaluates the fitness of all individuals and removes the worst-performing individuals.

To model the optimization problem, using the Sylvester Determinant Theorem [40] for channel capacity, we need to adjust the problem formulation as follows:

$$\begin{aligned} &\text{maximize} && \log_2 \det(\mathbf{I}_M + \rho_{dl}\mathbf{G}^{\mathrm{H}}\boldsymbol{\xi}\mathbf{G}), \\ &\text{subject to} && 0 < \xi_m < 1, \;\; m = 1, \ldots, M, \\ &&& \sum_{m}^{M} \xi_m = 1. \end{aligned} \tag{P6}$$

Problem (P6) introduces a novel variable for joint optimization. The variable $\boldsymbol{\xi} \in (\mathbb{R}_+^*)^{M \times M}$ is responsible for indicating the most important antennas, with higher values of ξ_m which corresponds to the value in m-th diagonal of $\boldsymbol{\xi}$, and, indirectly, indicates the power allocation for each user. Firstly, we need to sort the values of $[\boldsymbol{\xi}]_{mm}$ and select the S greater values to perform the antenna selection step. So,

$$\mathbf{z} = [\boldsymbol{\xi}] \geq \text{S-th largest value of } [\boldsymbol{\xi}], \tag{20}$$

and the power allocation matrix $\boldsymbol{\eta}$ is constructed as:

$$\boldsymbol{\eta} = \mathbf{G}^{\mathrm{H}}\boldsymbol{\xi}(\mathbf{G}^{\mathrm{H}})^{\dagger}, \tag{21}$$

where $(\cdot)^{\dagger}$ denotes the Moore-Penrose inverse operator. To ensure a feasible value for $\boldsymbol{\eta}$, the result can be normalized to ensure $\eta_k > 0$ and $\sum_{k=1}^{K} \eta_k = 1$.

3.4 Joint Antenna Selection and Power Allocation based on Matching Pursuit

Bearing in mind the need to reduce the computational load on the base station, a joint algorithm for antenna selection and power allocation is then proposed. HMPAS has proven to be an interesting method for antenna selection when considering both large- and small-scale fading in the channel. Certainly, the entire system description is simplified by considering only the

Algorithm 4 PSO-based proposed algorithm to Power Allocation.

1. Create a initial population $\mathfrak{D} = \{\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_{K^2}\}$ with the same approach as Equation (18), with K^2 particles;
2. Define constants: community consensus cc, individual best position ib and random walk rw;
3. Perform the fitness function, Equation (14), for all particles in initial population;
4. Create a memory of each particle $\mathfrak{D}^*$, initialize this variable using the initial positions;
5. Define $i \leftarrow K^2$;
6. Iterative loop: while $i > 1$;
   a. Define the global best particle $\mathbf{d}_j$;
   b. Update all particles: $\mathbf{d}_k \leftarrow cc.\mathbf{d}_j + ib.\mathbf{d}_k^* + rw.\mathbf{v}$, where $\mathbf{v} \in \mathbb{C}^M$ is a random vector;
   c. Perform the fitness function for all particles. If the procedure increases the fitness value, update the memory population $\mathfrak{D}^*$.
   d. Utilize the halving method, eliminating half of the particles in $\mathfrak{D}$. The criterion for elimination is based on the $i/2$ worst values of fitness;
   e. Update the loop iterator $i \leftarrow i/2$;
7. Using the memory $\mathfrak{D}^*$, localize the best particle over all iterations;
8. Output the best particle as ξ.

large-scale fading. It is important to highlight that Problem P5, which is used for power allocation, is formulated based on this simplification.

When the large-scale fading is equal for the channels between all transmitted antennas and a single terminal, equal power allocation becomes the optimal solution. In most general cases, this assumption is not true. Certainly, in this case, we are considering egalitarian fairness. Small- and large-scale fading components must be taken into account to construct optimal power allocation strategies based on a previously chosen fairness criterion. In this case, we propose an algorithm using a relaxed strategy based on MPAS, referred to here as Continuous MPAS (CMPAS), which is expressed as:

$$\begin{aligned} &\text{minimize} && ||\mathfrak{D}\tilde{\mathbf{z}} - \mathbf{b}||_2^2 \\ &\text{subject to} && ||\tilde{\mathbf{z}}||_0 = S. \\ & && \tilde{z}_i \in [0,1] \text{ for } i = 1, \ldots, M. \end{aligned} \tag{22}$$

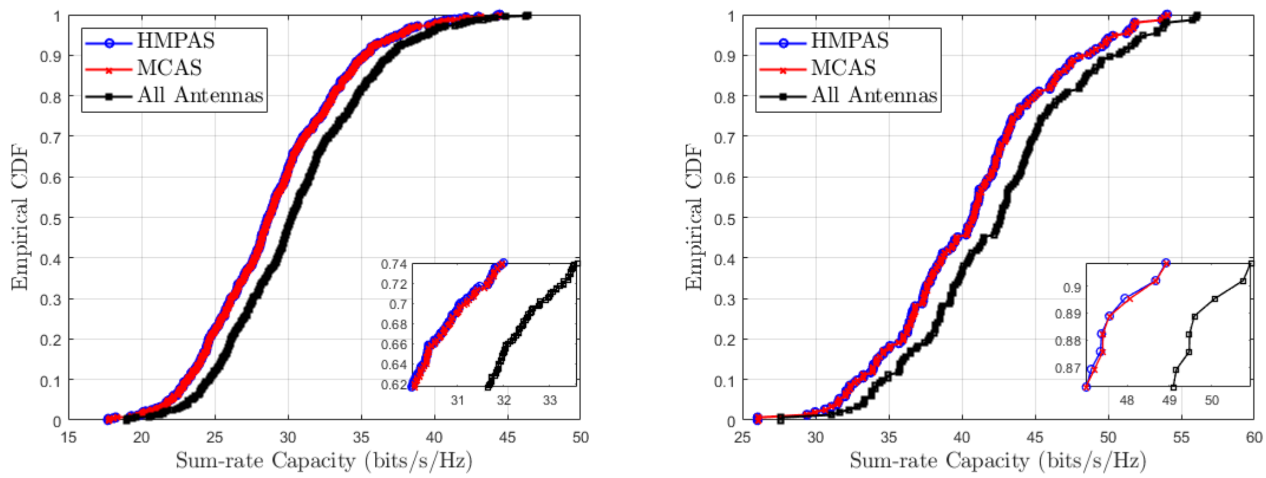
The boolean vector $\mathbf{z}$ is treated here as a continuous variable $\tilde{\mathbf{z}}$ to jointly produce two outcomes: the antenna selection and the power allocation results. This joint result can then be divided into the previously defined variables as shown in Equations (20) and (21). To employ this optimization, we can employ both MPAS and HMPAS. As shown in [23], both algorithms exhibit identical performance, and only the HMPAS version will be used in the results section.

4 RESULTS

To conduct the result analysis, we initially define some variables to establish a fair environment for comparing all the proposed algorithms. We define two different antenna arrangements by varying the total number of antennas M: 128 and 256. In these cases, $k = 16$ users are served. We considered a URA arrangment, with carrier frequency of 28 GHz and the antenna spacing of $\lambda/2$. The users were spread throughout the cell considering a minimum distance of BS equals to 10 m and a maximum distance equals to 100 m. The azimuth angle are randomly chosen in interval of $[-\pi, \pi]$ and the zenith angle in $[0, \pi]$. The cumulative density of probabilities (CDF) of the channel capacity, sum-rate or minimum rate is calculated over 10^3 iterations. We analyze the system using a SNR equal to 30 dB. The channel model employed in this study is proposed by Haneda *et al.* [41] and further elaborated in Figueiredo *et al.* [31]. It considers a free-space non-fading Line-of-Sight (LoS) propagation between the transmitter antennas arranged in a Uniform Linear Array (ULA) and the single-antenna receivers.

At this initial stage, we analyze only the antenna selection method. For comparison, we utilize the algorithm proposed by Gao [11], referred to here as Maximum Capacity Antenna Selection (MCAS). This algorithm approaches the maximization of channel capacity as a convex problem, with resolution employing an interior-point algorithm. As demonstrated in [9, 22, 23], MPAS/HMPAS is faster than MCAS and achieves similar results. To conduct the following experiment, we apply HMPAS using 2 clusters. These clusters are obtained using the classic K-means clustering method. The use of 2 clusters aims to reduce the complexity of MPAS by half, as demonstrated in [23].

https://doi.org/10.21528/lnlm-vol22-no2-art1



(a) 128 transmitter antennas in BS.

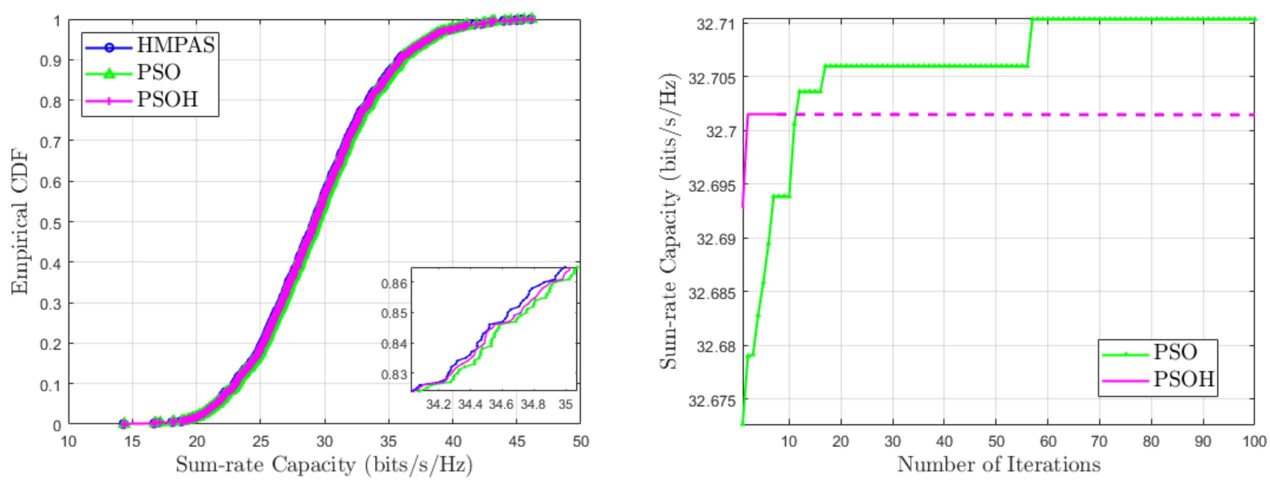
(b) 256 transmitter antennas in BS

Figure 2: Analysis of sum-rate using HMPAS, MCAS, and all antennas activated. In the simulated scenario, 90% of total antennas are selected.

In Figure 2, it is clear that HMPAS and MCAS exhibit similar performance. It is notable that the differences among methods diminish as the number of antennas in the arrangement increases. The major advantage of using the HMPAS is reduced complexity. As demonstrated in [9, 23], the number of operations is decreased by approximately 10^3 operations when compared to MCAS.

Following the analysis, we compare HMPAS, PSO and PSOH for antenna selection, while considering equal power allocation for both algorithms. Considering the case of $M = 128$ and $S = 115$, i.e., 90% selected, the result is exhibited in Figure 3. In these figures, we observe that PSO achieves performance close to that of HMPAS, but the convergence time can be prohibitive.

In the Figure 3b is illustrated the necessary number of iterations of traditional PSO and the proposed here. By restricting the number of iterations in PSO, we can achieve good results with lower time expenditure. To illustrate the reduction in complexity between PSOH and PSO, we compare the number of times the fitness function, that is, sum-rate capacity, needs to be computed to converge to the maximum value. In this simulation, PSO begins with 100 particles over 100 iterations, totalling 10,000 fitness evaluations. On the other hand, PSOH starts with $K^2 = 256$ particles and reduces by half over time, resulting in a total of 510 fitness evaluations. Also, in the Figure fig:2b, it can be concluded that the early stopping applied in PSOH results in a capacity impairment of approximately 0.03%.

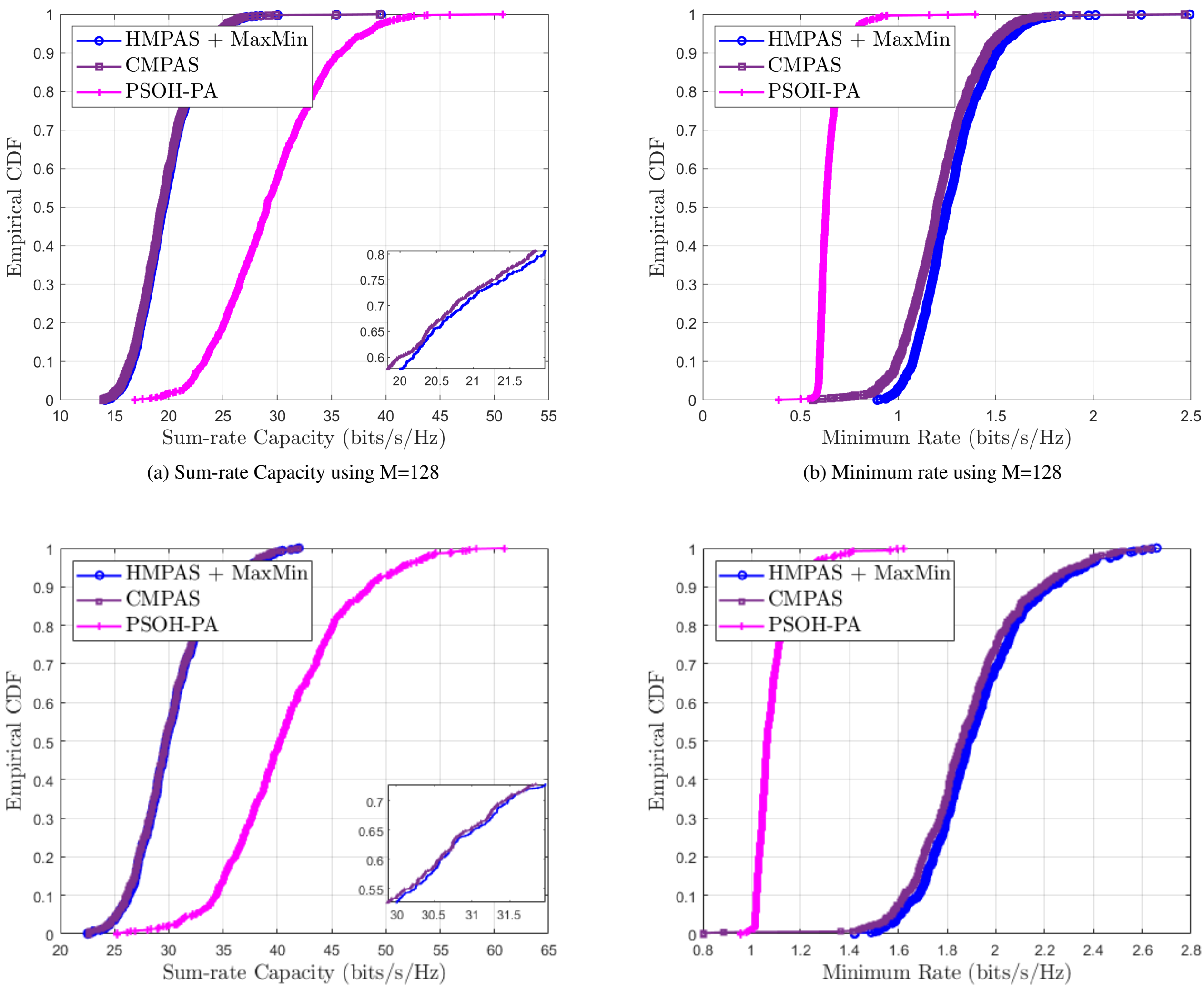


(a) Channel Capacity

(b) Iterations of PSO Algorithm to converge

Figure 3: Result of comparison between MPAS and PSO-based to antenna selection. Subfigure (b) illustrates the difference in convergence time between simple PSO and PSO with the Halving operation. The magenta dashed line illustrates the sum-rate capacity of PSOH after convergence.

To compare joint power allocation and antenna selection, the following results are evaluated. The proposed PSOH, the HMPAS with Max-min optimization, and the CMPAS are tested. The achieved results are as in Fig. 4.



(a) Sum-rate Capacity using M=128 (b) Minimum rate using M=128

(c) Sum-rate Capacity using M=256 (d) Minimum rate using using M=256

Figure 4: Analysis of the sum-rate capacity of HMPAS, MCAS, and all antennas activated. In these graphs, 90% of total antennas were selected.

As depicted in Figure 4, the MMFPA criterion is applied with HMPAS as the benchmark. In other words, we aim to achieve the maximum minimum rate for each user. To illustrate this, Figures 4a and 4c present the sum-rate capacity of all methods proposed here. CMPAS closely follows the results of HMPAS+MaxMin in both arrangements (M=128 and 256 antennas). PSOH-PA is configured to maximize the sum-rate capacity, as shown in Figures 4a and 4c. On the other hand, when analyzing the minimum rate achieved by each method, as shown in Figures 4b and 4d, it is possible to see how close the proposed method CMPAS is to the traditional approach.

To quantify the computational simplicity of the proposed methods, Table 1 is presented. In this table, we vary the number of antennas at the base station and the number of terminals served. To measure the execution time, we ran HMPAS and CMPAS 10^3 times and present the average result here. For the slower methods, we ran each arrangement for 1 hour in MCAS and allocated sufficient time for each test in PSOH-H.

Comparing the standard method for solving channel capacity optimization (MCAS) with the natural approach presented here, we can conclude that PSOH-PA is a viable solution for small arrangements and a small number of users. However, as the number of terminals increases, the number of initial particles also increases (K^2), making the solution impractical for these cases.

Restricting the analysis to AS methods only, HMPAS outperforms the convex approach by approximately 2,000 times in terms of time, in the best case. The average for the tested parameters is 450 times faster. The results show that it is feasible to perform AS in real time.

As an important note, the proposed HMPAS presented in [23] takes approximately 50% less time than the standard MPAS. Here, we compare the fastest version of MPAS with CMPAS. In Table 1, it is evident that the best case is 2.49 times faster than HMPAS with MaxMin. The worst case occurs with 128 antennas for 64 users, where the proposed method is 17% slower. In general, CMPAS is about 40% faster.

	AS Methods						AS and PA Methods								
	MCAS			HMPAS			HMPAS+MaxMin			CMPAS			PSOH-PA		
K \ M	128	256	512	128	256	512	128	256	512	128	256	512	128	256	512
16	11.498	26.643	58.0285	0.0906	0.2053	0.8719	0.0922	0.2059	0.8727	**0.0370**	**0.1436**	**0.5798**	7.1516	42.653	233.13
32	38.564	82.135	203.203	0.0974	0.4076	1.5932	0.0983	0.4086	1.5944	**0.0753**	**0.3060**	**1.1289**	203.99	816.95	4101.1
64	352.69	616.54	1291.9	0.1745	0.9351	3.7771	**0.1767**	0.9373	3.7799	0.2112	**0.8248**	**3.4475**	5872.5	20022	100444

Table 1: Time spent (in seconds) for each method varying the arrangement size and the number of users.

5 CONCLUSION

The need for higher rates and more robust transmission has led to the adoption of multiple transmitter antennas. Recently, the proposition of mMIMO has significantly influenced the prospects for the next standards of mobile communication. With a large number of transmitter antennas, mMIMO can achieve its key points, such as channel hardening and favorable propagation, which lead to higher achievable data rates.

As the number of antennas increases, the associated hardware costs and energy expenditure can become prohibitive. The solution lies in reducing the number of activated antennas without sacrificing the benefits achieved, thereby ensuring acceptable spectral efficiency. In a real environment, applying complex algorithms for antenna selection implies costly hardware requirements. Therefore, reducing the complexity of these algorithms is necessary. The total power used in the symbols for each served terminal in the cell poses a challenge. Increasing the energy expenditure can lead to better spectral efficiency and, consequently, higher data rates for the specific terminal. The increase in energy for a given terminal results in lower data rates for others. Thus, determining how much energy to allocate is an important consideration for BSs.

This work proposes two options to reduce complexity by implementing a joint algorithm for antenna selection and power allocation. By utilizing the MMFPA, this work aims to maximize the average sum-rate capacity by considering the selected antennas and power allocation strategies.

This work demonstrates the possibility of achieving good rates even with a significant reduction in the complexity of the algorithms utilized. The PSOH and CMPAS algorithms are proposed in this paper to jointly perform both operations.

As future work, the authors suggest applying the joint power allocation and antenna selection to cell-free environments. With certain considerations, scalability in this type of structure can be achieved. Another task that can be incorporated into this algorithm is optimizing downlink transmission, resulting in a more comprehensive algorithm for overall system optimization.

6 ACKNOWLEDGMENT

The authors would like to thank Brazilian agencies CNPq, CAPES, Niterói City office, FAPERJ Projects E-26/210.524/2019, E-26/200.161/2023, E-26/211.184/2019 and RNP/MCTIC, Grant No. 01245.010604/2020-14, for funding.

REFERENCES

[1] T. L. Marzetta. “Noncooperative cellular wireless with unlimited numbers of base station antennas”. *IEEE transactions on wireless communications*, vol. 9, no. 11, pp. 3590–3600, 2010.

[2] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors and F. Tufvesson. “Scaling up MIMO: Opportunities and challenges with very large arrays”. *IEEE signal processing magazine*, vol. 30, no. 1, pp. 40–60, 2012.

[3] D. d. S. Brilhante, J. C. Manjarres, R. Moreira, L. de Oliveira Veiga, J. F. de Rezende, F. Müller, A. Klautau, L. Leonel Mendes and F. A. P. de Figueiredo. “A literature survey on AI-aided beamforming and beam management for 5G and 6G systems”. *Sensors*, vol. 23, no. 9, pp. 4359, 2023.

[4] E. Björnson, L. Sanguinetti, H. Wymeersch, J. Hoydis and T. L. Marzetta. “Massive MIMO is a reality—What is next?: Five promising research directions for antenna arrays”. *Digital Signal Processing*, vol. 94, pp. 3–20, 2019.

[5] S. Zhou, M. Zhao, X. Xu, J. Wang and Y. Yao. “Distributed wireless communication system: a new architecture for future public wireless access”. *IEEE Communications Magazine*, vol. 41, no. 3, pp. 108–113, 2003.

[6] T. L. Marzetta. “Massive MIMO: an introduction”. *Bell Labs Technical Journal*, vol. 20, pp. 11–22, 2015.

[7] Y. Gao, H. Vinck and T. Kaiser. “Massive MIMO antenna selection: Switching architectures, capacity bounds, and optimal antenna selection algorithms”. *IEEE Transactions on signal processing*, vol. 66, no. 5, pp. 1346–1360, 2017.

[8] R. W. Heath, S. Sandhu and A. Paulraj. “Antenna selection for spatial multiplexing systems with linear receivers”. *IEEE Communications letters*, vol. 5, no. 4, pp. 142–144, 2001.

[9] M. O. K. Mendonça, P. S. R. Diniz, T. N. Ferreira and L. Lovisolo. “Antenna Selection in Massive MIMO Based on Greedy Algorithms”. *IEEE Transactions on Wireless Communications*, vol. 19, no. 3, pp. 1868–1881, 2020.

[10] S. Sanayei and A. Nosratinia. "Antenna selection in MIMO systems". *IEEE Communications magazine*, vol. 42, no. 10, pp. 68–73, 2004.

[11] X. Gao, O. Edfors, J. Liu and F. Tufvesson. "Antenna selection in measured massive MIMO channels using convex optimization". In *2013 IEEE globecom workshops (GC Wkshps)*, pp. 129–134. IEEE, 2013.

[12] Y. Xu, E. G. Larsson, E. A. Jorswieck, X. Li, S. Jin and T.-H. Chang. "Distributed Signal Processing for Extremely Large-Scale Antenna Array Systems: State-of-the-Art and Future Directions". *arXiv preprint arXiv:2407.16121*, 2024.

[13] X. Wang, L. Kong, F. Kong, F. Qiu, M. Xia, S. Arnon and G. Chen. "Millimeter wave communication: A comprehensive survey". *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 1616–1653, 2018.

[14] M. Gharavi-Alkhansari and A. B. Gershman. "Fast antenna subset selection in MIMO systems". *IEEE transactions on signal processing*, vol. 52, no. 2, pp. 339–347, 2004.

[15] X. Gao, O. Edfors, F. Tufvesson and E. G. Larsson. "Massive MIMO in real propagation environments: Do all antennas contribute equally?" *IEEE transactions on communications*, vol. 63, no. 11, pp. 3917–3928, 2015.

[16] P. V. Amadori and C. Masouros. "Interference-driven antenna selection for massive multiuser MIMO". *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, pp. 5944–5958, 2015.

[17] J. Chen, S. Chen, Y. Qi and S. Fu. "Intelligent massive MIMO antenna selection using Monte Carlo tree search". *IEEE Transactions on Signal Processing*, vol. 67, no. 20, pp. 5380–5390, 2019.

[18] Z. Abdullah, C. C. Tsimenidis and M. Johnston. "Tabu search vs. bio-inspired algorithms for antenna selection in spatially correlated massive MIMO uplink channels". In *2016 24th European Signal Processing Conference (EUSIPCO)*, pp. 41–45. Eurasip, 2016.

[19] Z. Abdullah, C. C. Tsimenidis and M. Johnston. "Quantum-inspired Tabu Search algorithm for antenna selection in massive MIMO systems". In *2018 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6. IEEE, 2018.

[20] C. Singh and P. C. Kishoreraja. "Optimal transmit antenna selection using hybrid algorithm for massive MIMO technology". *International Journal of Communication Systems*, vol. 36, no. 1, pp. e5366, 2023.

[21] Z. Abdullah, C. C. Tsimenidis, G. Chen, M. Johnston and J. A. Chambers. "Efficient low-complexity antenna selection algorithms in multi-user massive MIMO systems with matched filter precoding". *IEEE transactions on vehicular technology*, vol. 69, no. 3, pp. 2993–3007, 2020.

[22] J. N. Gois and T. Ferreira. "Low-Complexity Algorithm for Full-Array Antenna Selection based on Matching Pursuit". In *XXXIX Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBrT2021)*. SBrT, 2021.

[23] J. N. Gois, V. N. H. Silva and T. Ferreira. "Low Complexity Algorithm for Antenna Selection using Hierarchical Matching Pursuit". In *XLI Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBrT2023)*. SBrT, 2023.

[24] M. Balachandran and N. M. Vali Mohamad. "Joint Design of Antenna Selection and Resource Allocation in Massive MIMO-Based Ultra Dense Heterogeneous Networks". *Journal of Electrical and Computer Engineering*, vol. 2024, no. 1, pp. 3809470, 2024.

[25] H. Li, J. Cheng, Z. Wang and H. Wang. "Joint Antenna Selection and Power Allocation for an Energy-efficient Massive MIMO System". *IEEE Wireless Communications Letters*, vol. 8, no. 1, pp. 257–260, 2019.

[26] H. Gao, Y. Su, S. Zhang, Y. Hou and M. Jo. "Joint Antenna Selection and Power Allocation for Secure Co-Time Co-Frequency Full-Duplex Massive MIMO Systems". *IEEE Transactions on Vehicular Technology*, vol. 70, no. 1, pp. 655–665, 2021.

[27] A. S. Gharagezlou, M. Nangir and N. Imani. "Joint antenna selection and power allocation in massive MIMO systems with cell division technique for MRT and ZF precoding schemes". *Soft Computing*, pp. 1–12, 2024.

[28] N. Mohamadi, M. Dong and S. ShahbazPanahi. "Low-Complexity Joint Antenna Selection and Robust Multi-Group Multicast Beamforming for Massive MIMO". *IEEE Transactions on Signal Processing*, 2024.

[29] T. L. Marzetta, E. G. Larsson and H. Yang. *Fundamentals of massive MIMO*. Cambridge University Press, 2016.

[30] E. Björnson, J. Hoydis, M. Kountouris and M. Debbah. "Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits". *IEEE Transactions on information theory*, vol. 60, no. 11, pp. 7112–7139, 2014.

[31] F. A. P. de Figueiredo, C. F. Dias, E. R. de Lima and G. Fraidenraich. "Capacity Bounds for Dense Massive MIMO in a Line-of-Sight Propagation Environment". *Sensors*, vol. 20, no. 2, 2020.

[32] P. Chandhar, D. Danev and E. G. Larsson. "On ergodic rates and optimal array geometry in line-of-sight massive MIMO". In *2016 IEEE 17th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 1–6. IEEE, 2016.

[33] J. Qian, S. Shen and R. Murch. "Including Antenna Effects Into Capacity Formulations of Line-of-Sight MIMO Channels". *IEEE Wireless Communications Letters*, 2024.

[34] N. Fatema, G. Hua, Y. Xiang, D. Peng and I. Natgunanathan. "Massive MIMO linear precoding: A survey". *IEEE systems journal*, vol. 12, no. 4, pp. 3920–3931, 2017.

[35] B. M. Hochwald, T. L. Marzetta and V. Tarokh. "Multiple-Antenna Channel Hardening and Its Implications for Rate Feedback and Scheduling". *IEEE Transactions in Information Theory*, vol. 50, no. 9, pp. 1893–1909, Sept 2004.

[36] J. R. Hampton. *Introduction to MIMO communications*. Cambridge university press, 2013.

[37] H. Q. Ngo, E. G. Larsson and T. L. Marzetta. "Aspects of favorable propagation in massive MIMO". In *2014 22nd European Signal Processing Conference (EUSIPCO)*, pp. 76–80. IEEE, 2014.

[38] M. S. P. Silva, V. N. H. Silva, J. N. Gois and T. N. Ferreira. "Análise de Endurecimento do Canal e Propagação Favorável em Sistemas MIMO Massivo". In *2024 XXV Congresso Brasileiro de Automática*. SBA, 2024. Submitted.

[39] R. S. Chaves, E. Cetin, M. V. Lima and W. A. Martins. "On the convergence of max-min fairness power allocation in massive MIMO systems". *IEEE Communications Letters*, vol. 24, no. 12, pp. 2873–2877, 2020.

[40] D. A. Harville. "Matrix algebra from a statistician's perspective", 1998.

[41] K. Haneda, J. Zhang, L. Tan, G. Liu, Y. Zheng, H. Asplund, J. Li, Y. Wang, D. Steer, C. Li *et al.*. "5G 3GPP-like channel models for outdoor urban microcellular and macrocellular environments". In *2016 IEEE 83rd vehicular technology conference (VTC spring)*, pp. 1–7. IEEE, 2016.