

# Low Complexity Algorithm for Antenna Selection using Hierarchical Matching Pursuit

Jonathan N. Gois, Vinicius N. H. Silva, Tadeu N. Ferreira

**Abstract**—Massive MIMO systems have grown in popularity due to the implementations of 5G and the prospect of 6G technologies. Massive MIMO enables more efficient use of spectrum resources and larger data rates. When the number of antennas in a base station increases, so do the energy consumption and hardware cost. The high use of energy and processing can be managed through antenna selection. Among the structures proposed in the literature, this article presents hierarchical matching pursuit algorithm for antenna selection (HMPAS) that uses a combination of branch-and-bound with the matching pursuit antenna selection algorithm. In our experiments, the proposed algorithm presented considerable computational complexity reduction resulting in a shorter running time than the benchmark algorithms, while maintaining the error rate.

**Keywords**—Massive MIMO, antenna selection, matching pursuit

## I. INTRODUCTION

Wireless systems based on multiple-input multiple-output (MIMO) have become increasingly popular in recent years. Massive MIMO arrangements are usually employed for better data transmission, and for improving the spatial multiplexing gain and the energy efficiency of the whole transmission system [1]. Notwithstanding, emerging technologies have used a millimeter wave range to achieve higher data throughput. On the other hand, millimeter waves are more affected by path loss effect, scattering, and penetration loss [2]. To overcome this obstacle, the previous large antenna array can be used to reduce the path loss or to increase the beam directionality.

When all the antennas are activated, the total energy spent in the transmission is increased, with a higher spectral efficiency. To save energy, a possible solution is to select a subset of antennas to perform the transmission trying to maintain an interesting spectral efficiency.

There are several potential architectures for building an antenna array at the base station (BS). A fully connected architecture is the simplest, but it is the most expensive one. Each antenna of this type is connected to a radiofrequency (RF) chain. Then, separate signals can be sent to each antenna. A partially connected architecture is the second option. In this case, several antenna sets are supplied by the same RF chain. In the literature, most works discourse about the fully connected architecture.

Jonathan Gois is with Eng. Elétrica, CEFET Campus Nova Friburgo, Nova Friburgo, RJ, Brasil. The authors are with PPGEET, Universidade Federal Fluminense (UFF), Niterói, Brasil. e-mail: jonathan.gois@cefet-rj.br, viniciusnhs@id.uff.br, tadeu\_ferreira@id.uff.br. This work was partially supported by CNPq, FAPERJ Projects E-26/210.524/2019, E-26/200.161/2023, E-26/211.184/2019, and RNP, with resources from MCTIC, Grant No. 01245.010604/2020-14, under the 6G Mobile Communications Systems project of the Radiocommunication Reference Center (CRR) of the National Institute of Telecommunications (Inatel)

The authors in [3] review some articles about antenna selection for multiple antennas in terms of size, power, and hardware characteristics. The main methods presented are based on spatial multiplexing and channel capacity. The work discusses an exhaustive search, a channel-capacity-based method and a norm-based selection. Finishing the review, a greedy method is presented. The incremental selection algorithm starts selecting a row vector representing the channel with the highest norm. In each step, the residue of projection in this vector is approximated on the orthogonal vector and chosen as the one whose projection has the largest norm. The algorithm stops when the  $L$  antennas are selected.

Other works discussed greedy algorithms. In [4], Mendonça et al. use the matching pursuit based algorithm (named MPAS–Matching Pursuit Antenna Selection) to select the antennas. In this work the dictionary and the approximation vector for matching pursuit were proposed.

A family of channel capacity based algorithms is also found in the literature. The channel capacity is maximized using convex optimization in [5]. Amadori et al. [6] suggest algorithms in the same line-of-sight, maximizing the constructive interference between the users.

From another point of view, bio-inspired methods are presented in the literature [7], [8], [9] to maximize the channel capacity of selected antennas. Techniques such as Tabu search, Quantum-inspired, Particle Swarm Optimization, Genetic Algorithm, Artificial Bee Colony, and Hybrid Sea Lion-Whale Algorithm were simulated in those articles. The focus of the simulations is the computational cost.

Some authors suggest the reduced complexity architecture [10]. A Radio-Frequency (RF) chain is connected in each antenna in the fully connected architecture. In the reduced complexity architecture, some antennas are connected in the same RF chain. This architecture reduces the number of possible choices in antenna selection but also reduces the spectral efficiency.

Another way to reduce the complexity in the selection is a method known as Branch-And-Bound (BAB) [11]. In this method, a tree-based selection is used to reduce the number of searches.

Lastly, some methods based on intelligent algorithms are proposed in the literature. Chen et al. [12] show a Monte Carlo Tree Search as an option to select antennas, whereas Abdullah et al. [13] propose a grouping strategy.

We divide the article into the following sections: Section II describes the background of Massive MIMO antenna selection, the system modeling used, and the reduced system model after the antenna selection. Moreover, the matching pursuit antenna selection is introduced and its parameters are defined.

Section III depicts our proposed method, the Hierarchical Matching Pursuit Antenna Selection (HMPAS), and links it to the traditional Matching Pursuit Antenna Selection algorithm. This section ends with the pseudocode of the HMPAS algorithm. Section IV shows the simulation results, as well as the environment description in which the simulations were performed. Finally, Section V presents our conclusion.

## II. SYSTEM MODEL

Consider a downlink model in a single cell,  $M$  antennas at the BS serve  $K$  terminals at the cover cell depicted by the channel matrix  $\mathbf{G} \in \mathbb{C}^{M \times K}$ . The interference between neighboring cells is not considered in this model. In other words, their co-channel interference (CCI) in the interest cell is not taken into account. For a signal message  $\mathbf{x} \in \mathbb{C}^{M \times 1}$ , the received signal in each terminal is described by:

$$\mathbf{y} = \mathbf{G}^T \mathbf{P} \mathbf{x} + \boldsymbol{\omega}, \quad (1)$$

where  $\boldsymbol{\omega}$  is defined as an additive Gaussian noise vector, and  $\mathbf{P} \in \mathbb{C}^{K \times M}$  is a precoding matrix. The precoding matrix has the objective to reduce the channel effects at the sent message  $\mathbf{x}$ . A wireless channel suffers from multipath fading, which generates time dispersion of the transmitted signal and, thus, intersymbol interference. Moreover, some other impairment as non-linearity of the channel is mitigated using a precoder.

When the entire set of available antennas is used, maximum consumption of energy is achieved. Exploiting the redundancy of spatial diversity of a massive MIMO antenna array, we intend to reduce the entire set to  $S$  selected antennas. These antennas should maximize the channel capacity:

$$\mathcal{C} = \log_2 \det(\mathbf{I} - \rho \mathbf{G}^H \mathbf{diag}(\mathbf{z}) \mathbf{G}), \quad (2)$$

where  $\rho$  is the energy spent in each antenna,  $\mathbf{z}$  is a Boolean vector, while 0 represents the unselected antenna and 1, the selected antenna. Eliminating the unused antennas from the full array, we can write the matrix of selected antennas  $\mathbf{G}_S$  as:

$$\mathbf{G}_S = \text{rem}(\mathbf{diag}(\mathbf{z}) \mathbf{G}), \quad (3)$$

where operator  $\text{rem}(\cdot)$  is responsible to remove the null columns formed in the multiplication operation.

Without loss of generality, we consider  $\rho = 1/K$ , with the energy uniformly distributed among the selected antennas. The whole system after the antenna selection can be expressed as:

$$\mathbf{y} = \mathbf{G}_S^T \mathbf{P}_S \mathbf{x} + \boldsymbol{\omega}, \quad (4)$$

where  $\mathbf{P}_S$  represents the channel precoding.

Several criteria can be used to select the activated antennas. For example, we can perform an exhaustive search over all antennas to find the best combination of activated antennas to maximize a channel capacity, as described in Equation (2), or any other desired metric. An optimization of channel capacity can be performed to find the vector  $\mathbf{z}$ , which shows the activated antennas.

Another exploited solution is based on sparse recovery. The main idea in this approach is to create a sparse and overcomplete dictionary and select the best antennas according

to that. Greedy algorithms seek to maximize the channel capacity by performing the optimization on the following cost function [4]:

$$\begin{aligned} & \underset{\mathbf{z} \in \mathbb{C}^{M \times 1}}{\text{minimize}} \quad \|\mathcal{D} \mathbf{z} - \mathbf{b}\|_2^2, \\ & \text{subject to} \quad \|\mathbf{z}\|_0 = S. \end{aligned} \quad (5)$$

The dictionary  $\mathcal{D}$  and the approximation vector  $\mathbf{b}$  are important variables in this application. They need to be carefully chosen to represent the contribution of each antenna in the entire system. In [4], the authors proposed the dictionary and the approximation vector  $\mathbf{b}$  as:

$$\mathcal{D} = \left[ \text{vec}\{\mathbf{p}_1 \mathbf{p}_1^H\} \quad \text{vec}\{\mathbf{p}_2 \mathbf{p}_2^H\} \quad \dots \quad \text{vec}\{\mathbf{p}_M \mathbf{p}_M^H\} \right] \quad (6)$$

$$\mathbf{b} = \text{vec}(\mathbf{P}^T \mathbf{P}) \quad (7)$$

where the operator  $\text{vec}(\cdot)$  is responsible to resize its matrix argument into a vector. In Equation (6), the columns  $\mathbf{p}_i$  are the columns of precoder matrix  $\mathbf{P}$ .

The literature also suggest other pairs of  $\mathbf{b}$  and  $\mathcal{D}$ . For example, Gharavi-Alkhansari et al. [14] suggested the pair:

$$\mathcal{D} = (\mathbf{I}_M - \mathbf{G}^H \mathbf{G})^{-1} \quad (8)$$

$$\mathbf{b}_j = \mathbf{g}_j^H \mathbf{g}_j \quad (9)$$

where  $\mathbf{g}_j$  is the  $j$ -th column of the channel matrix  $\mathbf{G}$ .

In this article, we use the pair  $\mathbf{b}$  and  $\mathcal{D}$  arranged in the Equations (6) and (7) which achieves better performance in terms of bit-error rate, as in [4].

In Equations (6) and (7), the precoder matrix  $\mathbf{P}$  referred to zero-forcing precoding matrix [15], defined as:

$$\mathbf{P}_{ZF} = \hat{\mathbf{G}}^\dagger \quad (10)$$

$$= (\hat{\mathbf{G}}^* \hat{\mathbf{G}})^{-1} \hat{\mathbf{G}}^*, \quad (11)$$

where  $\hat{\mathbf{G}}$  is a estimation of the channel matrix  $\mathbf{G}$  performed by transmitter.

In this work, we use a zero-forcing precoder matrix as the dictionary.

## III. HIERARCHICAL MATCHING PURSUIT ANTENNA SELECTION

Hierarchical Matching Pursuit (HMP) algorithms are an interesting solution for the reduction of computational complexity. The HMP converts the exhaustive search in a dictionary for a tree-based search.

Matching Pursuit represents a vector  $\mathbf{b}$  into components of  $\mathcal{D} = \{\mathbf{d}_1, \dots, \mathbf{d}_M\}$ , where the dictionary is a sparse space. We represent the vector  $\mathbf{b}$  as a sum of projections:

$$\mathbf{b} = \sum_{i \in \mathcal{I}} a_i \mathbf{d}_i. \quad (12)$$

The set  $\mathcal{I}$  contains the indices which form the vector  $\mathbf{b}$  with projection in  $d_i$  equals to  $a_i$ . In other words, the set  $\mathcal{I}$  is composed of all antenna indexes that form  $\mathbf{b}$ .

Hierarchical Matching Pursuit [16] starts with division of the dictionary  $\mathcal{D}$  into  $\mathcal{D}_j$ ,  $j = 0, \dots, C$  where each  $\mathcal{D}_j \in \mathcal{D}$ . In other words,

$$\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2 \cdots \cup \mathcal{D}_C \quad (13)$$

where each  $\mathcal{D}_j$  is a small disjoint sparse subset of  $\mathcal{D}$ . In order to represent each subdictionary  $\mathcal{D}_j$ , we forms another dictionary  $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_C\}$ , with the centroids of each  $\mathcal{D}_j$ . The pair  $\{\mathbf{c}_j, \mathcal{D}_j\}$  will be formed using the  $K$ -means algorithm [17].

---

**Algorithm 1** Hierarchical Matching Pursuit Antenna Selection

---

- 1) Input:  $\mathcal{D}$ ,  $\mathbf{b}$ ,  $C$
  - 2) Performing the  $K$ -means into set of vectors  $\mathcal{D}$ :
    - a) Select  $C$  points as centroids  $c_1, c_2, \dots, c_C$  randomly;
    - b) For each column vector  $d_i$ , do:
      - i) Find the nearest centroid;
      - ii) Assign the point to cluster;
    - c) For each cluster  $j = 1, \dots, C$ , do:
      - i) Compute the new centroids using the mean of points assigned to that cluster;
    - d) Go back step (b) until convergence.
  - 3) Initialize variables  $\mathbf{z} \leftarrow 0_M$ ,  $i \leftarrow 1$ ,  $j \leftarrow 1$  and  $r(1) \leftarrow \mathbf{b}$
  - 4) For  $j = 1$  to  $j = S$ , do:
    - a) Find the closest centroid into cluster set  $\mathcal{C}$  to  $r(j)$ :
      - i)  $max\_ip \leftarrow -\infty$
      - ii)  $closest\_centroid \leftarrow 0$   $\triangleright$  Index of closest centroid
      - iii) For  $i = 1$  to  $i = C$ , do:
        - $ip \leftarrow \langle r(j), c(i) \rangle$
        - If  $ip > max\_ip$ , then:
          - $max\_ip \leftarrow ip$
          - $closest\_centroid \leftarrow i$
        - $i \leftarrow i + 1$
      - iv)  $size\_D \leftarrow size(\mathcal{D}_{closest\_centroid})$
    - b) Find the vector in the reduced dictionary  $\mathcal{D}_i$ 
      - i)  $max\_ip \leftarrow -\infty$
      - ii)  $closest\_vec \leftarrow 0$   $\triangleright$  Index of closest vector in  $\mathcal{D}_i$
      - iii) For  $l = 1$  to  $l = size\_D$ 
        - $ip \leftarrow \langle r(j), d_i(l) \rangle$
        - If  $ip > max\_ip$ , then:
          - $max\_ip \leftarrow ip$
          - $closest\_vec \leftarrow l$
        - $l \leftarrow l + 1$
      - iv) Map  $closest\_vec$  in an index of  $\mathcal{D}$ ,  $idx$
      - v)  $\mathbf{b}(idx) = 1$
      - vi)  $r(j+1) = r(j) - d_{idx}$
- 

We can divide the hierarchical matching pursuit into two steps. The first one is to realize the first approximation in the set  $\mathcal{C}$ . After finding the best approximation in this set, we

will search in the associate dictionary  $\mathcal{D}_j$ . In this subset, we will repeat the search and find the best approximation. Now, we can link the chosen vector  $\mathbf{d}_k$  with the selected antenna. Hierarchical Matching Pursuit Antenna Selection (HMPAS) is described in the Algorithm 1.

## IV. RESULTS

### A. General Parameters

Intending to generate simulated results, some parameters should be defined. The environment was defined with  $M = 256$  antennas serving  $K = 16$  terminals. We desire to select  $S = 128$  antennas from the total  $M$ . A Monte Carlo campaign was performed with 1000 rounds to simulate the HMPAS in contrast to MPAS. The sent data was encoded using Binary Phase Shift-Keying (BPSK). A total of 200 bits, randomly generated, were sent in each round. After the generation and encoding, the data is sent in a Rayleigh channel imposed by Gaussian noise varying the signal-to-noise ratio between -15 dB to 0 dB. The simulation result is analyzed by the bit error rate (BER) average per user. To perform the HMPAS, we defined the number of clusters ranging from 2 up to 32 clusters.

Furthermore, the channel needed to be estimated in the transmitter to perform the zero-forcing precoding. So, in our case, we used a perfect estimation in all simulations, without loss of generality. This assumption does not prejudice the results because all simulations are affected equally by an imperfect Channel State Information (CSI). The time spent on the channel estimation step is disregarded in the simulations.

In all simulations, the software Matlab was used to create the simulation environment and the routines of HMPAS and MPAS. The PC on which the results are obtained is an Intel I5-4460 CPU 3.20 GHz  $\times$  2 and 8 Gb of RAM.

### B. Overall Quality results

Ideally, we would like to select a number  $S < M$  antennas at which bit error rates are maintained or slightly affected. Here, we will start with the MPAS as a benchmark of quality in the transmission because the goal is to reduce the complexity of this algorithm.

In Figure 1 is possible to notice the same BER average per user for the MPAS and the versions of HMPAS with 2, 4, 8, 16, and 32 clusters.

### C. Complexity analysis

Following the results achieved in [4], the number of flops required to perform the MPAS with the ZF precoding is  $\frac{28}{3}K^3(1+L) + 8M^2K + 8MSK^2 + MK^2(6+8L) + SK^2(6+8L) - 4S^2K^2 - 3K^2(1+L) + 8MKL + 5$ . Where  $M$ ,  $S$ , and  $K$  were already defined before as the number of antennas, the number of selected antennas, and the number of served terminals, respectively. The variable  $L$  is the size of the block message, here defined as 200 bits.

As also described in the same work, the number of flops to create the dictionary is  $6MK^2$ , reminding the original dictionary size is  $MK \times M$ . The main loop of the MPAS

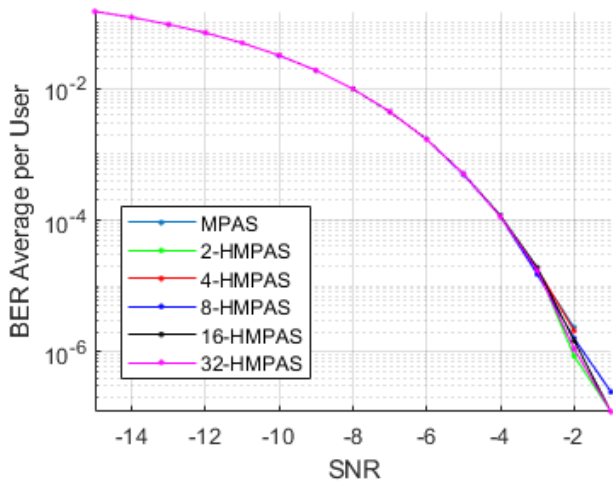


Fig. 1. BER average per user plot for the MPAS and the HMPAS with different number of clusters.

required the following number of flops  $4 + 2SK^2 - 4S^2K^2 + 4SK^2 + 8MK^2S$ .

Upon closer inspection, we can notice the direct dependency of the number of columns in the dictionary ( $MK$ ) with the number of flops in the main loop. Actually, the dependency of the dictionary size is also present in the total of flops with another contributions as the generation of the residue  $r$  in each step of the selection.

Here, we performed the  $K$ -means algorithm to split the dictionary into minor dictionaries. The time complexity of this algorithm [18] is equals to  $\mathcal{O}(n^2)$  where  $n$  is the number of input patterns or, in other words, the size of input data. This is a simplification of the complexity  $\mathcal{O}(CTn)$ , where  $C$  is the number of clusters and  $T$  is the iterations needed to algorithm converge. Considering  $T \propto n$ , the effective cost, in flops, of the  $K$ -means algorithm is  $Cn^2$ .

To execute the HMPAS algorithm, we need to run the  $K$ -means algorithm only once to split the dictionary. So, the number of flops in this operation results in  $CM^2K^2$ . Considering an average size of dictionary  $\bar{D}$  equals to  $MK/C$ , the number of flops of main loop as described as  $(4 + 2SK^2 - 4S^2K^2 + 4SK^2 + 8\frac{MK^2}{C}S) + CMK(1 + MK)$ .

Also, the total number of flops will be changed with this approach. As the reduction in the number of flops is dependent on the average size of the clusters, in this work we show the reduction of number of flops using the time spent in the simulation as described in the Figure 2.

Using the tested setup already described, the 8-HMPAS achieved less than 1/3 of time spent compared with the benchmark MPAS. The other configurations varying the number of clusters in HMPAS also reduce the time spent in the simulation. We can also notice the dependency of number of cluster in the time spent in this algorithm.

To demonstrate the efficiency of HMPAS to reduce the temporal complexity, Figure 3 shows the variation between 1 to 256 clusters using the HMPAS algorithm compared to MPAS. It is possible to note that there is a minimum time in the HMPAS curve.

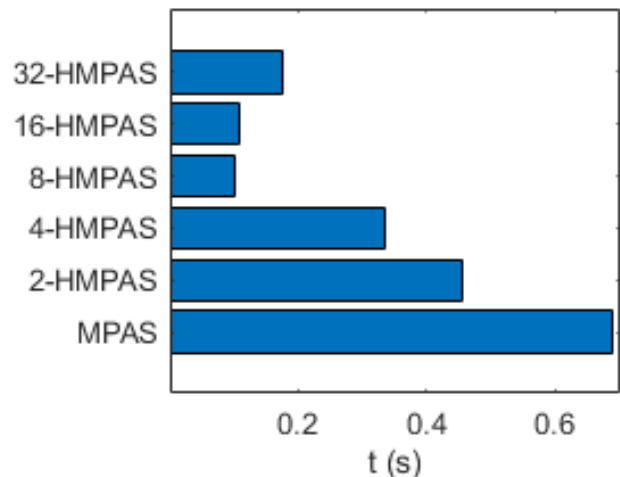


Fig. 2. Total time spent for different number of cluster compared to MPAS.

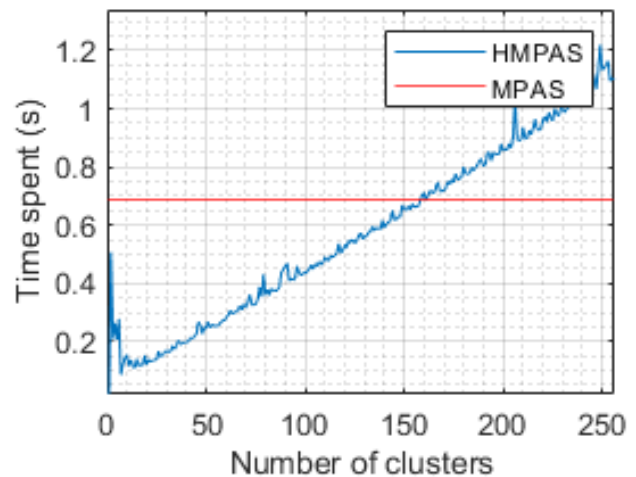


Fig. 3. Variation of time spent in HMPAS with different numbers of clusters.

In order to find this minimum time, a comparison of the total time spent in both algorithms can be performed. The number of clusters that reduce the amount of time spent can be expressed as:

$$C = \frac{4S^2 \left( 1 + \sqrt{1 - \frac{1+MK}{2S^2}} \right)}{1 + MK} \quad (14)$$

Using the described configuration, the minimum time is achieved with  $C \approx 30$  clusters.

## V. CONCLUSION

In this paper, we propose a low complexity algorithm for antenna selection labeled as Hierarchical Matching Pursuit Antenna Selection (HMPAS) to apply in Massive MIMO systems. The selection of antennas in the transmitter is responsible to increase the beam directionality and this can also used for energy saving.

The main idea of this paper is reduced the search space of the MPAS reducing the size of the dictionary with the  $K$ -means clustering algorithm. The MPAS algorithm is executed two times: the first one is to find the best correspondence to the dictionary composed by the cluster vectors and the second time applied in the reduced dictionary.

The proposed algorithm shown the BER results comparable with the benchmark. Moreover, comparing the time complexity, our algorithm shows a significant reduction of time spent.

At last, our work demonstrate the possibility to reduce the total complexity of the MPAS algorithm reducing the sizes of dictionaries without loss of performance. It is possible to apply this algorithm with another types of clustering algorithms as also with another greedy algorithms. This work can be extended for another types of dictionaries (using the Gharavi-Alkhansar [14] proposition, for example) or another types of precoder.

#### ACKNOWLEDGMENTS

The authors thank Marcele O. K. de Mendonça for providing some of the scripts used in this work. 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] S. Khalid, R. Mehmood, W. B. Abbas, F. Khalid, and M. Naem, "Joint transmit antenna selection and precoding for millimeter wave massive mimo systems," *Physical Communication*, vol. 42, p. 101137, 2020.
- [2] 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.
- [3] S. Sanayei and A. Nosratinia, "Antenna selection in mimo systems," *IEEE Communications magazine*, vol. 42, no. 10, pp. 68–73, 2004.
- [4] M. O. Mendonça, P. S. 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, 2019.
- [5] 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.
- [6] 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.
- [7] 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)*. IEEE, 2016, pp. 41–45.
- [8] —, "Quantum-inspired tabu search algorithm for antenna selection in massive mimo systems," in *2018 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2018, pp. 1–6.
- [9] 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, p. e5366, 2023.
- [10] A. Garcia-Rodriguez, C. Masouros, and P. Rulikowski, "Reduced switching connectivity for large scale antenna selection," *Ieee transactions on communications*, vol. 65, no. 5, pp. 2250–2263, 2017.
- [11] M. Tian, W. Sun, P. Zhang, L. Huang, and Q. Li, "Joint beamforming design and receive antenna selection for large-scale mimo wiretap channels," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 3, pp. 2716–2730, 2020.
- [12] 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.
- [13] 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.
- [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] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna gaussian broadcast channel," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1691–1706, 2003.
- [16] H. G. Feichtinger, A. Turk, and T. Strohmer, "Hierarchical parallel matching pursuit," in *Image Reconstruction and Restoration*, vol. 2302. SPIE, 1994, pp. 222–232.
- [17] S. Lloyd, "Least squares quantization in pcm," *IEEE Transactions on Information Theory*, vol. 28, no. 2, pp. 129–137, 1982.
- [18] M. K. Pakhira, "A linear time-complexity k-means algorithm using cluster shifting," in *2014 International Conference on Computational Intelligence and Communication Networks*, 2014, pp. 1047–1051.