

# SDMA Grouping Based on Unsupervised Learning for Multi-User MIMO Systems

Francisco Hugo Costa Neto, Tarcisio Ferreira Maciel

**Abstract**—In this study, we investigate a spatial division multiple access (SDMA) grouping scheme to maximize the total data rate of a multi-user multiple input multiple output (MU-MIMO) system. Initially, we partition the set of mobile stations (MSs) into subsets according to their spatial compatibility. We explore different clustering algorithms, comparing them in terms of computational complexity and capability to partition MSs properly. In the following, we schedule MSs from the different subsets to build the SDMA group. Since we consider a scenario with a massive arrange of antenna elements and that operates on the mmWave scenario, we employ a hybrid beamforming scheme and analyze its behavior in terms of the total data rate. The analog and digital precoders exploit the channel information obtained from clustering and scheduling, respectively. The simulation results indicate that a proper partition of MSs into clusters can take advantage of the spatial compatibility effectively and reduce the multi-user (MU) interference. The hierarchical clustering (HC) enhances the total data rate 25% compared with the baseline approach, while the density-based spatial clustering of applications with noise (DBSCAN) increases the total data rate 20%.

**Index Terms**—SDMA grouping, multi-user MIMO, hybrid beamforming, unsupervised learning, clustering

## I. INTRODUCTION

The fifth-generation (5G) of wireless communications imposes huge requirements on the efficiency and quality of the offered services. The exponential growth of data rates and massive device connectivity contradicts the imminent spectrum shortage [1]. Several technologies have been developed in the last years to satisfy these challenging requirements. In our work, we investigate how the transmission of multiple data streams over a massive arrange of antenna elements using the same time-frequency resource block can improve spectral efficiency, and consequently, the total data rate [2].

Thus, we investigate some of the fundamental techniques from 5G for enhanced mobile broadband (eMBB) services, namely millimeter-wave, massive multi-user (MU) multiple

input multiple output (MIMO), and hybrid beamforming. Moreover, we evaluate the role of multi-antenna techniques in 5G. We exploit space-division multiple access (SDMA) to allow spatial multiplexing, i.e., the division of user equipments (UEs) in groups to share radio resources. The performance of a SDMA group depends on how efficiently the MU interference is mitigated. This is directly determined by the channel characteristics of the selected UEs. It stands out that the selection of UEs into spatially compatible SDMA groups can improve the total data rate. Otherwise, the signals sent to the UEs may interfere with each other and threaten the system performance. Therefore, the SDMA group composition impacts system operation efficiency [3].

SDMA-orthogonal frequency division multiple access (OFDMA) systems can allocate resources in time, frequency, and space dimensions to different UEs. In this context, the large number of degrees of freedom leads to highly complex radio resource allocation (RRA) problems [3], [4]. In particular, SDMA grouping can be classified as an integer optimization problem, since it involves integer variables, like the number of UEs or of time-frequency resource blocks. Usually, these problems have combinatorial behavior, which implies high complexity and ask for an exhaustive search in order to obtain an optimal solution [3]. The RRA problems can be divided into subproblems (e.g., frequency assignment, power allocation, SDMA grouping) to reduce the overall complexity. In this case, each subproblem considers a different dimension of the RRA problem. However, the resulting subproblems still hold a considerably high level of complexity. Thus, suboptimal strategies are often proposed in the literature to solve them [4].

In particular, for the SDMA grouping problem, two methodologies stand out: i) iterative SDMA grouping; and ii) clustering followed by UE scheduling. In [3], [5], there is a direct iterative formation of groups, i.e., each UE is included sequentially in the group, according to the spatial compatibility between the candidate UE and the UEs already admitted to the SDMA group. Another approach, used in [6], [7], first partitions all UEs of the system into subsets, also called clusters, and afterwards schedules UE from different clusters to compose the SDMA group.

In [8], the authors explore the K-means clustering (KMAC), a classical unsupervised machine learning algorithm, to split UEs according to second-order channel statistics computed according to a given mathematical expression. Moreover, it is also considered a scheduling scheme that combines load balancing and precoding design considering proportional fairness. In our work, we also consider the KMAC algorithm, but in a more realistic system-level simulation to determine the

Francisco Hugo Costa Neto and Tarcisio Ferreira Maciel are with the Wireless Telecommunications Research Group, Department of Teleinformatics Engineering, Center of Technology, Federal University of Ceará, 60020-181, Fortaleza, Brazil (e-mail: hugo@gtel.ufc.br, maciel@gtel.ufc.br). (Corresponding author: Francisco Hugo Costa Neto.)

This work was supported by the Ericsson Research, Sweden, and Ericsson Innovation Center, Brazil, under UFC.47 Technical Cooperation Contract Ericsson/UFC. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. Francisco Hugo Costa Neto was supported by CAPES under the grant 88882.183549/2018-01. T. F. Maciel would like to acknowledge CNPq for its financial support under the grants 426385/2016-0 and 308621/2018-2.

A preliminary version of this paper was presented in XXXVII Brazilian Symposium on Telecommunications and Signal Processing (SBTr'2019), Petrópolis, RJ, Brazil, September, 2019 [12].

Digital Object Identifier: 10.14209/jcis.2020.13

channel matrices and calculate covariance matrices. In [9], the authors evaluate other unsupervised machine learning algorithms to partition UEs. They evaluate the agglomerative hierarchical clustering (AHC) and K-medoids clustering (KMDC) considering different similarity measures, such as weighted likelihood, subspace projection, and Fubini-Study. In our work, we evaluate the AHC algorithm in a much more challenging scenario since we consider a larger number of UEs in a random arrangement within the cell. In [10], the authors investigate a SDMA grouping problem in MU MIMO in the context of 5G networks. They evaluate a partitioning process based on KMAC and a scheduling algorithm based on the branch and bound algorithm to select UEs from the clusters to compose a SDMA group that supports multiple spatial streams per cluster.

We consider a MU-MIMO system based on joint spatial division and multiplexing (JSDM). Initially proposed in [11], the JSDM scheme has as the main idea the partition of UEs into clusters according to second-order channel statistics and serving these groups by using a downlink hybrid beamforming scheme. It is a two-stage beamforming that combines an analog precoder, defined according to the partitioning process, with a digital precoder, which is a function of the effective channels of the UEs that compose the SDMA group. The JSDM framework has been extended in [6], where practical issues are addressed, such as the evaluation of the regime of a realistic setting where UEs have different angles of arrival and angular spreads.

In [12], we investigated the maximization of the total sum rate of an MU-MIMO system based on an SDMA scheme combined with hybrid beamforming. The simulation results indicated that the combination of a proper partition of UEs into clusters based on KMAC, and suitable scheduling based on the best-fit provided a technique able to exploit spatial compatibility more effectively and reduced the MU interference. The present work takes further steps in this research topic. We explore different clustering algorithms, comparing them in terms of computational complexity and capability to partition UEs properly.

Hybrid beamforming is a cost-effective alternative to implement massive MIMO in the millimeter-wave spectrum. The hybrid beamforming design requires the optimization of system performance under several hardware constraints, such as the number of radio frequency (RF) chains and the dimension of the phase shifter network [13]. As mentioned previously, the hybrid beamforming has a digital and an analog component. On the one hand, the digital component can be performed for each UE on each sub-carrier. On the other hand, several UEs and sub-carriers share the analog component since it is a post-inverse fast Fourier transform (IFFT) operation. The structure of the analog component influences the hardware efficiency and has a notable influence on the spectral efficiency. [13].

In the proposed hybrid beamforming scheme, we define the analog precoder according to the characteristics of each cluster. Therefore, the design of the clustering algorithm has a critical impact on the achievable performance of the MU-MIMO system. Based on that, we evaluate different unsupervised learning algorithms, analyzing their operating principles, and

their main advantages and disadvantages.

Motivated by the above discussion, we investigate a SDMA grouping algorithm in a millimeter wave MU-MIMO scenario. The main contributions of the work can be summarized as follows:

- 1) evaluation of different unsupervised machine learning algorithms to partition UEs into proper subsets;
- 2) evaluation of a hybrid precoding scheme based on JSDM and its impact on the total data rate.

The remainder of this work is organized as follows. Section II describes assumed system model. Section III discusses the main design aspects of the evaluated SDMA grouping solution. In Section IV, we discuss the proposed hybrid precoding scheme. Finally, performance results are shown in Section V, and the main conclusions are drawn in Section VI.

*Notation:* bold lowercase and uppercase letters represent column vectors and matrices, respectively.  $(\cdot)^T$  and  $(\cdot)^H$  stand for transpose and Hermitian of a matrix, respectively.  $|\cdot|$ ,  $\|\cdot\|_2$ ,  $\|\cdot\|_F$  denote the absolute value, the Euclidean norm, and the Frobenius norm, respectively. Calligraphic upper-case letters denote sets, and  $|\cdot|$  denotes set cardinality.  $\mathbb{E}\{\cdot\}$  denotes expectation operator.

## II. SYSTEM MODEL

We consider the downlink of a multi-user multiple input multiple output (MU-MIMO) system based on OFDMA. The system is composed of one base station (BS) and a set  $\mathcal{J}$  containing  $J$  UEs. The transmitter uses  $M$  antennas to send  $S_j$  data streams to  $N_j$  antennas at the  $j$ th receiver. Before transmission, for a given resource block (RB) and transmission time interval (TTI), the symbol vector  $\mathbf{x}_j \in \mathbb{C}^{S_j \times 1}$  is filtered by the precoding matrix  $\mathbf{F}_j \in \mathbb{C}^{M \times S_j}$ . The filtered symbols are then transmitted through the channel associated with the RB which response is represented by  $\mathbf{H}_j \in \mathbb{C}^{N_j \times M}$ . Thus, the prior-filtering receive vector  $\mathbf{y}_j \in \mathbb{C}^{N_j \times 1}$  at the  $j$ th receiver is given by

$$\mathbf{y}_j = \mathbf{H}_j \mathbf{F}_j \sqrt{\mathbf{P}_j} \mathbf{x}_j + \sum_{i \neq j} \mathbf{H}_j \mathbf{F}_i \sqrt{\mathbf{P}_i} \mathbf{x}_i + \mathbf{z}_j, \quad (1)$$

where  $\mathbf{P}_j \in \mathbb{R}^{S_j \times S_j}$  is the power matrix, given by  $\alpha \mathbf{I}_{S_j}$ , where  $\alpha$  is the transmit power allocated to each stream associated to the  $j$ th UE and  $\mathbf{I}_{S_j}$  is the  $S_j \times S_j$  identity matrix; the second term on the right-hand side of (1) represents the MU interference, also known as intracell interference, caused by any UEs sharing the same RB;  $\mathbf{z}_j \in \mathbb{C}^{N_j \times 1}$  is the additive Gaussian noise vector, whose elements are independent and identically distributed (IID) as  $\mathcal{CN}(0, \sigma_z^2)$ .

The input symbol vector is normalized so that  $\mathbb{E}\{\mathbf{x}_j \mathbf{x}_j^H\} = \mathbf{I}_{S_j}$ . The channel coefficient of a given RB corresponds to that associated with the middle sub-carrier and the first transmitted orthogonal frequency division multiplexing (OFDM) symbol in a TTI or a sub-frame. Thus, we consider that the channel remains constant during resource allocation in a TTI and over an RB. Moreover, we assume that the required channel state information (CSI) is available at the transmitter and receivers.

At the receiver, the vector  $\mathbf{y}_j$  is filtered by the decoding matrix  $\mathbf{G}_j \in \mathbb{C}^{S_j \times N_j}$ . Therefore, the post-filtering receive vector  $\hat{\mathbf{y}}_j \in \mathbb{C}^{S_j \times 1}$  is given by

$$\hat{\mathbf{y}}_j = \mathbf{G}_j \mathbf{H}_j \mathbf{F}_j \sqrt{\mathbf{P}_j} \mathbf{x}_j + \mathbf{G}_j \sum_{i \neq j} \mathbf{H}_i \mathbf{F}_i \sqrt{\mathbf{P}_i} \mathbf{x}_i + \mathbf{G}_j \mathbf{z}_j. \quad (2)$$

We consider a multipath wireless channel model [14] to describe the response channel matrix  $\mathbf{H}_j$ . It can be expressed in terms of underlying physical paths as

$$\mathbf{H}_j = \sum_{p=1}^{N_p} \varphi_p \chi \mathbf{v}_r(\phi_{p,r}) \mathbf{v}_t^H(\phi_{p,t}) e^{-j2\pi\tau_p f} e^{j2\pi\nu_p t}, \quad (3)$$

where, for each propagation path  $p$ ,  $\varphi_p$  is the path gain;  $\chi$  is the shadowing modeled log-normal random variable with standard deviation  $\sigma_{sh}$ ;  $\phi_{p,r}$  is the angle of arrival (AoA) at the receiver;  $\phi_{p,t}$  is the angle of departure (AoD) at the transmitter;  $\tau_p$  is the relative delay;  $f$  is the frequency of the central sub-carrier of the RB;  $\nu_p$  is the Doppler shift;  $\mathbf{v}_r(\phi_{p,r}) \in \mathbb{C}^{N_r \times 1}$  is the response vector and  $\mathbf{v}_t(\phi_{p,t}) \in \mathbb{C}^{N_t \times 1}$  is the steering vector. The multipath wireless channel model is represented in Fig. 1.

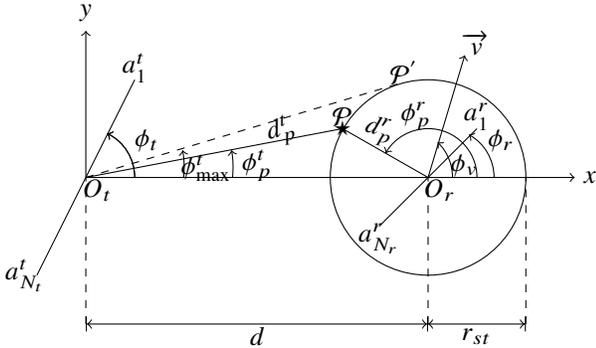


Fig. 1: Multipath wireless channel model for an  $N_r \times N_t$  MIMO channel with local scatters around the UE.

### III. SDMA GROUPING DESIGN

In our study, the SDMA grouping problem is performed into two steps. In the first step, we partition the set  $\mathcal{J}$  of all UEs of the system into a set of clusters  $\mathcal{C}$  according to the spatial compability of their channels. In the second step, we select one UE from each cluster to compose the SDMA group  $\mathcal{G}$ . The following subsections detail the most relevant design aspect of these steps.

#### A. Clustering in MU-MIMO Systems

Clustering is a class of algorithms from the unsupervised machine learning which aims to partition entities according to the similarity among them, i.e., the determination of subsets of entities with similar characteristics without any external intervention [15]–[17]. Therefore, we employ clustering algorithms to divide the population  $\mathcal{J}$  of UEs into a set of clusters  $\mathcal{C} = \{C_1, \dots, C_K\}$ . The UEs belonging to a same cluster are

quite similar, while UEs from distinct clusters are somewhat different.

In our work, we characterize each UE according to its channel covariance eigenspace. Given the channel matrix of the  $j$ th UE,  $\mathbf{H}_j \in \mathbb{C}^{N_j \times M}$ , the sample transmit covariance matrix  $\tilde{\mathbf{H}}_j \in \mathbb{C}^{M \times M}$  is given by

$$\tilde{\mathbf{H}}_j = \frac{1}{W} \sum_{w=1}^W \mathbf{H}_{w,j}^H \mathbf{H}_{w,j}, \quad (4)$$

where  $W$ , called TTI window size, indicates the number of channel matrix samples considered in the averaging process and  $\mathbf{H}_{w,j}^H$  is the conjugate transpose of the channel matrix.

The eigendecomposition of  $\tilde{\mathbf{H}}_j$  can be written as

$$\tilde{\mathbf{H}}_j = \mathbf{D}_j \mathbf{\Lambda}_j \mathbf{D}_j^{-1}, \quad (5)$$

where  $\mathbf{D}_j \in \mathbb{C}^{M \times M}$  defines the matrix composed of eigenvectors and  $\mathbf{\Lambda}_j \in \mathbb{C}^{M \times M}$  is the diagonal matrix of eigenvalues.

Hence, the similarity among entities is defined according to the dominant eigenvector  $\mathbf{d}_j \in \mathbb{C}^{M \times 1}$  of each UE, i.e., the eigenvector associated to the highest eigenvalue of  $\tilde{\mathbf{H}}_j$ .

According to [18], the number of possible ways of partitioning  $J$  UEs into a set of  $K$  clusters (non-empty and disjoint subsets of UEs) is given by the Stirling number of the second kind, which can be written as

$$\frac{1}{K!} \sum_{k=0}^K (-1)^k \binom{K}{k} (K-k)^J. \quad (6)$$

The exhaustive search for suitable clusters consists in the evaluation of all possibilities of partitioning the population of UEs into the set of clusters. In practice, this brute-force approach is an unfeasible task with a prohibitive computational complexity  $O(K^J/K!)$  [18]. Consequently, we evaluate algorithms that represent the main practical paradigms of the cluster analysis, namely K-means clustering (KMAC), agglomerative hierarchical clustering (AHC), and density-based spatial clustering of applications with noise (DBSCAN) [19]–[21].

1) *K-Means Clustering*: initially proposed in [22], it employs a greedy iterative approach to find a partition that minimizes the distance between entities that belong to each cluster to the cluster average value, called centroid. In our model, the centroid is a vector  $\mathbf{v}_k \in \mathbb{C}^{M \times 1}$  that describes the central characteristic of the  $k$ th cluster  $C_k$ , which is given by

$$\mathbf{v}_k = \frac{1}{|C_k|} \sum_{i \in C_k} \mathbf{d}_i, \quad (7)$$

where  $i$  indexes the UEs that currently belong to cluster  $C_k$ .

The pseudo-code of the KMAC is presented in Algorithm 1. The first step of the algorithm is the initialization of centroids. In our study, we randomly select the dominant eigenvector of  $K$  out of the  $J$  UEs as initial cluster centroids. For more details on the impact of centroid initialization on the KMAC algorithm please refer to [23]. Each iteration of the algorithm consists in a clustering assignment followed by a centroid update. Given the group of centroids  $\mathbf{v}_k, \{k = 1, \dots, K\}$ , provided at iteration  $t = 0$ , in the group assignment step, each UE  $j \in \mathcal{J}$  is assigned to the cluster  $C_{k^*}$  with closest mean as

$$C_{k^*} = \arg \min_k \|\mathbf{d}_j - \mathbf{v}_k\|_2^2. \quad (8)$$

---

**Algorithm 1** K-Means Clustering

---

```

1:  $t \leftarrow 0$ 
2: for all  $k \in \{1, \dots, K\}$  do
3:   initialize centroid  $\mathbf{v}_k \in \mathbb{C}^{M \times 1}$ 
4: repeat
5:    $t \leftarrow t + 1$ 
6:   for all  $k \in \{1, \dots, K\}$  do
7:     initialize cluster  $C_k \leftarrow \emptyset$ 
8:     for all  $j \in \{1, \dots, J\}$  do
9:       for all  $k \in \{1, \dots, K\}$  do
10:        assign the  $j$ th UE to the closest cluster  $C_{k^*}$ 
        according to Eq.(8)
11:    for all  $k \in \{1, \dots, K\}$  do
12:      update centroid  $\mathbf{v}_k^{(t)} \in \mathbb{C}^{M \times 1}$  according to Eq.(7).
13: until convergence condition Eq.(9) holds or the maximum
    number of iterations is achieved.

```

---

We define a threshold  $\epsilon > 0$  and test at every iteration  $t$  if there is no significant change of the centroids in comparison to the previous iteration, i.e.,

$$\sum_{k=1}^K \|\mathbf{v}_k^{(t)} - \mathbf{v}_k^{(t-1)}\|_2^2 \leq \epsilon. \quad (9)$$

In the centroid update step, new centroids values  $\mathbf{v}_k$  are computed for each cluster from the UEs in  $C_k$  using Eq.(7). The assignment and centroid update steps are carried out until convergence is reached. The output of the algorithm is a clustering of the UEs into  $K$  disjoint clusters  $C_k, k = \{1, \dots, K\}$ , and a set of vectors  $\{\mathbf{v}_1^{(t)}, \mathbf{v}_2^{(t)}, \dots, \mathbf{v}_K^{(t)}\}$  obtained as the centroids of the clusters. Based on these outputs, we determine the analog precoder, as described in Section IV. The computational complexity of the KMAC algorithm based on the pseudo-code Algorithm 1 is  $O(TJKM)$ , where  $T$  is the number of iterations. More details are provided in Appendix A.

2) *Agglomerative Hierarchical Clustering*: it performs a hierarchical clustering employing a bottom-up strategy, i.e., each UE starts as a cluster  $C_j = \{j\}$ , and the most similar pairs of clusters are successively merged until the target number of clusters  $K$  is reached. The merging of clusters represents the creation of a new cluster in a higher hierarchical level. The linkage rule determines the merge strategy. In our work, we evaluate the complete, average, and Ward linkage rules.

The complete linkage defines the distance between the clusters  $C_i$  and  $C_j$  as the maximum distance between a UE on cluster  $C_i$  and a UE on cluster  $C_j$ , which can be written as

$$\delta_{\text{complete}}(C_i, C_j) = \max \{ \|\mathbf{d}_i - \mathbf{d}_j\|_2 \mid i \in C_i, j \in C_j \}. \quad (10)$$

The average linkage is defined as the average pairwise distance between the UEs from different clusters. It can be written as

$$\delta_{\text{average}}(C_i, C_j) = \frac{1}{|C_i||C_j|} \sum_{i \in C_i} \sum_{j \in C_j} \|\mathbf{d}_i - \mathbf{d}_j\|_2. \quad (11)$$

The Ward linkage defines the distance between the clusters  $C_i$  and  $C_j$  as the increase in the sum squared errors (SSE)

---

**Algorithm 2** Agglomerative Hierarchical Clustering

---

```

1:  $t \leftarrow 0$ 
2: initialize each UE as a cluster  $C_i = \{i\}$  for  $j \in \{1, \dots, J\}$ 
3: compute the distance between clusters  $\delta\{C_i, C_j\}$ 
4: repeat
5:    $t \leftarrow t + 1$ 
6:   determine the closest pair of clusters  $\{C_i, C_j\}$ 
7:   merge the clusters  $C_{ij} \leftarrow C_i \cup C_j$ 
8:   update the set of clusters  $C^{(t)} = C^{(t-1)} \setminus \{C_i \cup C_j\} \cup C_{ij}$ 
9:   update the distance among clusters
10: until  $|C| = K$ 

```

---

when the two clusters are merged. The SSE of the cluster  $C_i$  is defined as

$$\text{SSE}_i = \sum_{j \in C_i} \|\mathbf{d}_j - \mathbf{v}_i\|_2 \quad (12)$$

where  $\mathbf{v}_i$  represents the centroid of the  $i$ th cluster, calculated as  $\frac{1}{|C_i|} \sum_{j \in C_i} \mathbf{d}_j$ .

The SSE for a set of clusters  $C = \{C_1, \dots, C_K\}$  can be written as  $\text{SSE} = \sum_{k=1}^K \text{SSE}_k$ .

The Ward linkage defines the distance between the clusters  $C_i$  and  $C_j$  as the net change in the SSE value when these clusters are merged. It can be written as

$$\delta_{\text{Ward}}(C_i, C_j) = \frac{|C_i||C_j|}{|C_i| + |C_j|} \|\mathbf{v}_i - \mathbf{v}_j\|_2^2. \quad (13)$$

Algorithm 2 describes the pseudo-code of the AHC. Given a set of clusters  $C^{(t)} = \{C_1, \dots, C_J\}$ , at each iteration we determine the most similar pair of clusters  $C_i$  and  $C_j$  according to the linkage rule. These clusters are merged into a new cluster  $C_{ij}$ . We update the set of clusters,  $C^{(t)} = C^{(t-1)} \setminus \{C_i \cup C_j\} \cup C_{ij}$ . The process is repeated until the set of clusters  $C$  reaches the desired size, i.e.,  $|C| = K$ .

Based on the pseudo-code Algorithm 2, the computational complexity of the AHC algorithm is  $O(J^3KM)$ . More details are provided in Appendix B.

3) *Density-Based Spatial Clustering of Applications with Noise*: initially proposed in [24], it is a clustering algorithm that exploits the local density of entities to determine clusters, instead of using the distance among them.

The  $\epsilon$ -neighborhood of the  $j$ th UE is defined as the set of UEs whose Euclidean distance is smaller than  $\epsilon$ . It can be written as

$$\mathcal{N}(j, \epsilon) = \{i \mid \|\mathbf{d}_i - \mathbf{d}_j\|_2 \leq \epsilon\}. \quad (14)$$

The  $j$ th UE is classified a core point if there are at least  $\chi$  UEs in its  $\epsilon$ -neighborhood, i.e., if  $|\mathcal{N}_\epsilon(j)| \geq \chi$ . The parameter  $\chi$  defines the local density, i.e., the frequency threshold that allows the classification of a UE as a core point. If the UE does not meet the frequency threshold, but still belongs to the neighborhood of a core point, it is classified as a border point. If the  $j$ th UE does not meet any of the previously described criteria, it is called a noise point [25].

The  $j$ th UE is directly density reachable from the  $i$ th core point UE if  $j \in \mathcal{N}(i, \epsilon)$ . The  $j$ th UE is called density reachable

from the  $i$ th UE if there is a set of core points leading from  $i$  to  $j$ . Furthermore, two UEs are defined as density connected if they are simultaneously density reachable from the same core point. A density-based cluster is defined as the maximal set of density connected points [18].

The pseudo-code for the DBSCAN is shown in Algorithm 3. Initially, we compute the  $\epsilon$ -neighborhood for each UE, and check if it is a core point. Then, we update the set of core points  $C_{\text{core}}$ . In the following, commands between lines 7 and 15 perform the assignment of UEs to clusters. Some border UE may be reachable from core UEs in more than one cluster. We arbitrarily assign to one of the clusters since overlapping clusters are allowed. Those UEs that do not belong to any cluster are treated as outliers or noise. The computational complexity of the DBSCAN algorithm is  $O(J^2M)$ . More details are provided in Appendix C.

---

**Algorithm 3** Density-Based Spatial Clustering
 

---

```

1: initialize set of core points  $C_{\text{core}} \leftarrow \emptyset$ 
2: for all  $j \in \mathcal{J}$  do
3:   compute the  $\epsilon$ -neighborhood  $\mathcal{N}(j, \epsilon)$  according to
   Eq. (14)
4:   initialize the label of  $j$ th UE  $L_j \leftarrow \emptyset$ 
5:   if  $|\mathcal{N}(j, \epsilon)| \geq \chi$  then
6:     update the set of core points  $C_{\text{core}} \leftarrow C_{\text{core}} \cup \{j\}$ 
7:    $k \leftarrow 0$ 
8:   for all  $j \in C_{\text{core}}$  do
9:     if  $L_j = \emptyset$  then
10:       $k \leftarrow k + 1$ 
11:      assign  $j$ th UE to the  $k$ th cluster  $L_j \leftarrow k$ 
12:      for all  $i \in \mathcal{N}(j, \epsilon)$  do
13:        assign the  $i$ th UE to the  $k$ th cluster  $L_i \leftarrow k$ 
14:      if  $i \in C_{\text{core}}$  then
15:        return to line 12
16: update the set of clusters  $C \leftarrow \{C_1 \cup C_2 \cup \dots \cup C_k\}$ 
17: update the set of noise points  $C_{\text{noise}} \leftarrow \{j \in \mathcal{J} | L_j = \emptyset\}$ 
18: update the set of border points  $C_{\text{border}} \leftarrow \mathcal{J} \setminus \{C_{\text{core}} \cup C_{\text{noise}}\}$ 
    
```

---

### B. Scheduling Algorithm

In the following, given the partition of UEs into a set of clusters  $C$ , we select one UE  $g_k = j^* \in C_k$  from each cluster to create an SDMA group  $\mathcal{G}$  containing  $K$  UEs (or streams). We employ a random scheduling algorithm. That is, the BS randomly selects one UE from each cluster  $C_k$  to compose  $\mathcal{G}$ . Therefore, the selection of the components of the SDMA group occurs without the assessment of any compatibility or suitability criteria.

## IV. HYBRID BEAMFORMING DESIGN

In our study, we adopt a system with a hybrid precoding scheme based on JSDM. Thus, we consider a two-stage precoder composed of an analog and a digital component. In the following, we detail our main design assumptions.

### A. Analog Precoder Design

The analog precoder  $\mathbf{F}_{\text{RF}} \in \mathbb{R}^{M \times K}$ , is defined from the centroids of the clusters  $C = \{C_1, \dots, C_K\}$  and can be written as

$$\mathbf{F}_{\text{RF}} = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \dots \quad \mathbf{v}_K]. \quad (15)$$

where  $\mathbf{v}_k \in \mathbb{C}^{M \times 1}$  is calculated according to Eq. (7).

The analog precoder is assumed to have elements of equal magnitude, i.e., only phase shifting is performed in the analog domain.

### B. Digital Precoder Design

For each selected UE in the SDMA group  $g \in \mathcal{G}$ , we consider that a matched filter on the dominant receive eigenmode of the UE's channel will be employed, i.e., at the receiver side,  $\mathbf{G}_g = \mathbf{u}_{1,g}^H \in \mathbb{C}^{1 \times N_g}$ , where

$$\mathbf{H}_g = \mathbf{U}_g \boldsymbol{\Sigma}_g \mathbf{V}_g^H = [\mathbf{u}_{1,g} \quad \dots \quad \mathbf{u}_{N_g,g}] \boldsymbol{\Sigma}_g \mathbf{V}_g^H \quad (16)$$

is the singular value decomposition (SVD) of the channel  $\mathbf{H}_g \in \mathbb{C}^{N_g \times M}$  of the  $g$ th UE in the SDMA group  $\mathcal{G}$  and  $\mathbf{u}_1$  is its dominant left-singular vector.

Given the total number  $N_{\mathcal{G}} = \sum_{g \in \mathcal{G}} N_g$  of receive antennas in the SDMA group  $\mathcal{G}$ , we define the group channel matrix  $\mathbf{H}_{\mathcal{G}} \in \mathbb{C}^{N_{\mathcal{G}} \times M}$  and the decoder matrix  $\mathbf{G} \in \mathbb{C}^{C \times N_{\mathcal{G}}}$  based on the channel matrices of the selected UEs, so that can write

$$\mathbf{H}_{\mathcal{G}} = [\mathbf{H}_1^T \quad \mathbf{H}_2^T \quad \dots \quad \mathbf{H}_C^T]^T, \quad \text{and} \quad (17)$$

$$\mathbf{G} = \text{diag}(\mathbf{u}_{1,1}^H, \mathbf{u}_{1,2}^H, \dots, \mathbf{u}_{1,C}^H). \quad (18)$$

Therefore, given the block diagonal decoder matrix  $\mathbf{G}$ , the group channel matrix  $\mathbf{H}_{\mathcal{G}}$  and the analog precoder  $\mathbf{F}_{\text{RF}}$ , the equivalent channel matrix  $\mathbf{H}_{\text{eq}} \in \mathbb{C}^{C \times C}$  is given by

$$\mathbf{H}_{\text{eq}} = \mathbf{G} \mathbf{H}_{\mathcal{G}} \mathbf{F}_{\text{RF}}. \quad (19)$$

There are different precoding techniques that either totally or partially suppress spatial interference or ignore it. We evaluate the zero-forcing (ZF) filter as digital precoder. The ZF precoding is conceived to decorrelate the transmit signals so that the signal at every receiver output is free of interference. The precoding matrix is defined as

$$\mathbf{F}_{\text{BB}} = \frac{\mathbf{H}_{\text{eq}}^H (\mathbf{H}_{\text{eq}} \mathbf{H}_{\text{eq}}^H)^{-1}}{\|\mathbf{H}_{\text{eq}}^H (\mathbf{H}_{\text{eq}} \mathbf{H}_{\text{eq}}^H)^{-1}\|_F} \quad (20)$$

The total power constraint is enforced by normalizing the digital and analog filters, such that  $\|\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}} \sqrt{\mathbf{P}_{\mathcal{G}}}\|_F^2 = p_{\text{RB}}$ , where  $\mathbf{P}_{\mathcal{G}} \in \mathbb{R}^{C \times C}$  is the block diagonal power matrix resulting of the combination of the power matrices of each UE belonging to the SDMA group and  $p_{\text{RB}}$  is the power for a given RB. We consider that the number of clusters is equal to the number of RF chains and streams, i.e.,  $S = C$ . Therefore, the dimensions of  $\mathbf{F}_{\text{RF}}$  and  $\mathbf{F}_{\text{BB}}$  are compatible with the dimension of  $\mathbf{F}$ , so that  $\mathbf{F} = \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}} \in \mathbb{C}^{M \times C}$ .

The post-filtering receive vector of the group  $\hat{\mathbf{y}}_{\mathcal{G}} \in \mathbb{C}^{C \times 1}$  is given by

$$\hat{\mathbf{y}}_{\mathcal{G}} = \mathbf{G} \mathbf{H}_{\mathcal{G}} \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}} \sqrt{\mathbf{P}_{\mathcal{G}}} \mathbf{x}_{\mathcal{G}} + \mathbf{G} \mathbf{z}_{\mathcal{G}}, \quad (21)$$

where  $\mathbf{x}_G \in \mathbb{C}^{C \times 1}$  is the group symbol vector and the  $\mathbf{z}_G \in \mathbb{C}^{N_G \times 1}$  is the group noise vector.

Defining  $\mathbf{Q} = \mathbf{G}\mathbf{H}_G\mathbf{F}\sqrt{\mathbf{P}_G} \in \mathbb{C}^{C \times C}$ , the average signal to interference-plus-noise ratio (SINR) perceived by the stream  $i$  can be calculated as

$$\Gamma_i = \frac{|\mathbf{Q}_{ii}|^2}{\sum_{j \neq i} |\mathbf{Q}_{ij}|^2 + \sigma_z^2}, \quad (22)$$

where  $\sigma_z^2$  is the average noise power. The data rate of stream  $i$  is calculated according to Shannon capacity formula [26] and is given by

$$\Phi_i = B \log_2(1 + \Gamma_i), \quad (23)$$

where  $B$  is the bandwidth of the RB.

## V. PERFORMANCE EVALUATION

We consider a single cell system with a carrier frequency of 28 GHz and a bandwidth of 100 MHz. Consequently, according to 5G numerology [27, Table 2], this carrier frequency implies a set of 125 RBs, each one composed of 12 sub-carriers equally spaced of 60 kHz. Furthermore, the number of sub-frames per frame is 10, each sub-frame has 14 symbols and the TTI duration is 0.25 ms. We assume the one ring scattering channel [28], with the propagation effects modeled according to the urban micro (UMi) street canyon deployment proposed in [29]. Therefore, the path loss parameters are  $a = 19.8$ ,  $b = 32.4$  and  $c = 20.0$ . The AoDs are uniformly distributed in the range  $[5^\circ, 15^\circ]$ . The total power is fixed for all precoders with equal power allocation (EPA) among RBs. Other relevant simulation parameters are listed in Table I.

Initially, we define the appropriate number of clusters to partition the set of UEs in our simulation scenario. The determination of the optimal number of clusters is out of the scope of this work. For more details regarding this issue, the reader may refer to [30]–[32] and references therein.

In our study, we define the Calinski-Harabasz index (CHI) as the evaluation metric of clustering performance. Also called variance ratio criterion, it measures the cluster validity according to the average between- and within-cluster dispersion [33]. The CHI of the partition of  $J$  UEs into  $K$  clusters can be written as

$$\Omega(J, K) = \frac{\alpha(J)/(K-1)}{\beta(K)/(J-K)} \quad (24)$$

where  $\alpha(J)$  is the inter-cluster variance, and  $\beta(K)$  is the intra-cluster variance.

TABLE I: Simulation Parameters

| Parameter                    | Symbol        | Value | Unit   |
|------------------------------|---------------|-------|--------|
| Simulation time              | $T$           | 25    | ms     |
| Number of simulation rounds  | –             | 100   | –      |
| Cell radius                  | $R_c$         | 100   | m      |
| Total transmit power         | $P_t$         | 35    | dBm    |
| Noise figure                 | $\zeta$       | 9     | dB     |
| Noise spectral density       | $\xi$         | -174  | dBm/Hz |
| Number of paths              | $N_p$         | 7     | –      |
| Shadowing standard deviation | $\sigma_{sh}$ | 3.1   | dB     |

We consider a MU-MIMO scenario, where the BS is equipped with a uniform linear array (ULA) with  $M = 64$  antennas separated of half-wavelength and each UE is equipped with a ULA with  $N = 2$  antennas also separated of half-wavelength. We assume a set of  $J = 100$  UEs arranged inside a sector of  $120^\circ$  of the cell according to a random uniform distribution.

Figure 2 indicates the average value of the  $\Omega(J, K)$  considering partitions defined by the clustering algorithms KMAC and AHC - Ward linkage, for a variable number of clusters  $K = [2, 10]$ . On one hand, KMAC algorithm presents a  $\Omega(J, K)$  almost constant for all values of  $K$ . On the other hand, the AHC presents a maximum value at  $K = 3$ , and remains constant at a lowest level for all other values. Moreover, AHC achieves absolute values of  $\Omega(J, K)$  higher than the KMAC algorithm. Hence, the AHC can establish a partition of UEs into more dense and well separated clusters than KMAC.

The evaluation of the  $\Omega(J, K)$  values define to our system a range of the most appropriate number of clusters,  $K \geq 3$  due to AHC or  $K \geq 8$  due to KMAC. Since the number of clusters is directly related to the number of RF chains, a design variable that impacts the energy efficiency and the hardware cost of the network, we define the most appropriate number of clusters to our scenario as  $K = 3$ .

In the following, we compare the clustering algorithms described in Section III considering a set of  $J = 100$  UEs to be partitioned into  $K = 3$  clusters. We define as baseline the random clustering (RC) algorithm. It defines a random distribution of UEs into subsets without any evaluation of similarity among the elements of the same cluster. The UEs are arranged inside a sector of  $120^\circ$  of the cell according to a random uniform distribution.

Figure 3 shows the cumulative distribution function (CDF) of the total data rate for our system considering the effect of the different SDMA algorithms. After the definition of the components of the clusters, we randomly select one UE from each subset to compose the SDMA group. Tables II, III, and IV resume the main advantages of the evaluated algorithms in comparison with RC. To simplify the notation, we abbreviate AHC with the Ward linkage rule to AHC-W, AHC with the

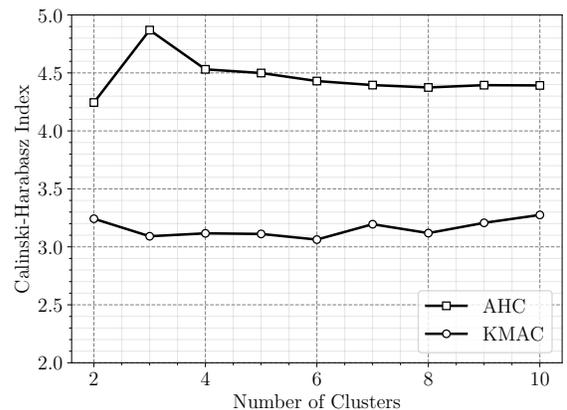
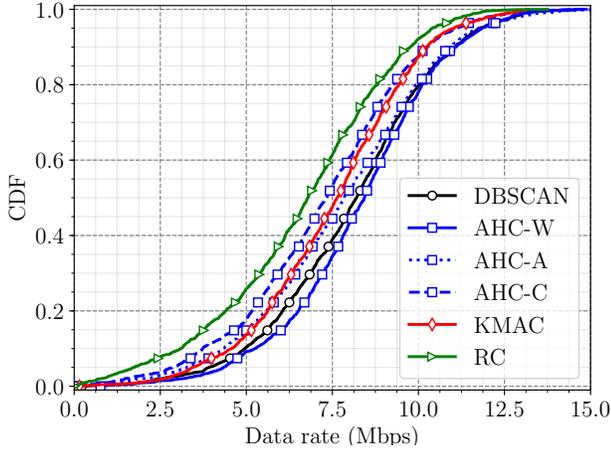
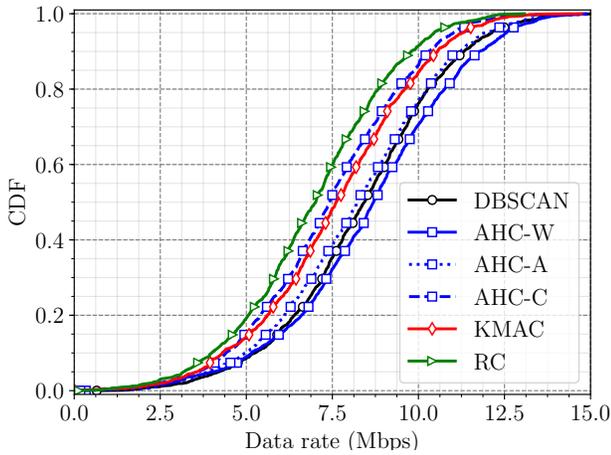


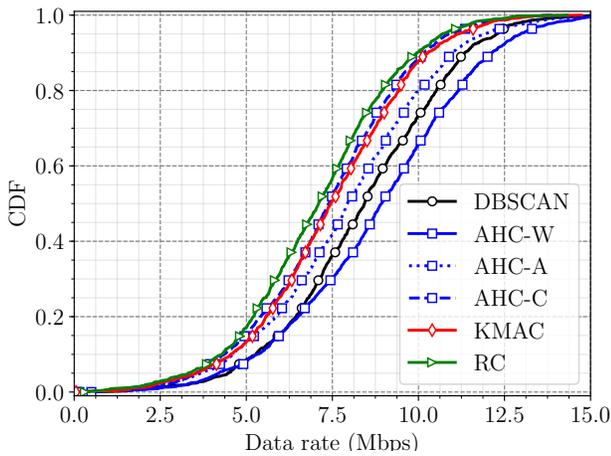
Fig. 2: Behavior of the CHI for KMAC and AHC algorithms in a MU-MIMO system with  $J = 100$  UEs randomly distributed.



(a) BS with  $M = 16$  antennas.



(b) BS with  $M = 32$  antennas.



(c) BS with  $M = 64$  antennas.

Fig. 3: Evaluation of a MU-MIMO scenario considering different clustering algorithms.

TABLE II: Behavior of the average total data rate - MU-MIMO ( $M = 16 / N = 2$ )

| Algorithm | Total Data Rate (Mbps) | Variation (%) |
|-----------|------------------------|---------------|
| RC        | 6.78                   | 0             |
| KMAC      | 7.63                   | +12.53        |
| AHC-C     | 7.34                   | +08.25        |
| AHC-A     | 7.84                   | +15.63        |
| AHC-W     | 8.42                   | +24.19        |
| DBSCAN    | 8.20                   | +20.09        |

TABLE III: Behavior of the average total data rate - MU-MIMO ( $M = 32 / N = 2$ )

| Algorithm | Total Data Rate (Mbps) | Variation (%) |
|-----------|------------------------|---------------|
| RC        | 7.01                   | 0             |
| KMAC      | 7.63                   | +08.84        |
| AHC-C     | 7.42                   | +05.85        |
| AHC-A     | 8.21                   | +17.11        |
| AHC-W     | 8.69                   | +23.96        |
| DBSCAN    | 8.42                   | +20.11        |

complete linkage rule to AHC-C, and AHC with average linkage rule to AHC-A. It can be observed that all clustering algorithms easily overcome the RC since the selection of UEs spatially compatible reduces the MU interference. The AHC algorithm employs a successful strategy of clustering when it considers the Ward linkage rule. It increases the total data rate in approximately 25% in all MU-MIMO scenarios.

The DBSCAN algorithm does not require the number of clusters, but it requires the setting of two parameters ( $\epsilon$  and  $\chi$ ). We establish an iterative algorithm to update ( $\epsilon$  and  $\chi$ ) until the number of clusters equals to  $K$ . Besides this additional setting step, the DBSCAN algorithm still have lower computational complexity than AHC. Moreover, it achieves total data rates close to those obtained by AHC-W.

TABLE IV: Behavior of the average total data rate - MU-MIMO ( $M = 64 / N = 2$ )

| Algorithm | Total Data Rate (Mbps) | Variation (%) |
|-----------|------------------------|---------------|
| RC        | 7.10                   | 0             |
| KMAC      | 7.44                   | +6.13         |
| AHC-C     | 7.39                   | +4.08         |
| AHC-A     | 7.93                   | +11.70        |
| AHC-W     | 8.93                   | +25.77        |
| DBSCAN    | 8.43                   | +18.73        |

Based on the description of the clustering algorithms, the increase in the number of UEs impacts the quality of the partitioning, since the UEs tend to be closer to each other and the channel characteristics become more similar. Therefore, the partitioning of UEs into distinct clusters becomes more difficult. In this case, it is mandatory for the scheduling algorithm to consider the correlation among UEs to compose the SDMA group since the different clusters do not ensure near orthogonal channel matrices. As in our work, we consider random scheduling, the increase of the system load rises the MU interference, reducing the total data rate. For more details, please refer to [34], where the authors provide a detailed

evaluation of the impact of the number of UEs on the total data rate achieved by an SDMA grouping solution combined with JSMD.

## VI. CONCLUSIONS

In this study, we investigated a SDMA grouping scheme to maximize the total data rate of a MU-MIMO system. The main step is the partition of UEs into subsets according to their spatial compatibility. We explored different clustering algorithms, namely KMAC, AHC, and DBSCAN comparing them in terms of computational complexity and capability to partition UEs properly. Since we consider a scenario with a massive arrange of antenna elements and that operates on the mmWave scenario, we employ a hybrid beamforming scheme and analyze its behavior in terms of the total data rate. The analog and digital precoders exploit the channel information obtained from clustering and scheduling, respectively. The simulation results indicate that the combination of a proper partition of MSs into clusters can exploit the spatial compatibility effectively and to reduce inter-cell interference. On one hand, the hierarchical clustering (HC) enhances the total data rate 25% compared with the baseline approach, but has significant computational complexity. On the other hand, the density-based spatial clustering of applications with noise (DBSCAN) has lower computational complexity and increases the total data rate 20%. As a major disadvantage, it requires an additional parameter determination process.

In our work, we consider a conventional CSI feedback. That is, each UE reports the channel covariance matrix to the BS. In scenarios with a massive number of antenna elements at the BS and a high number of UEs, this feedback scheme implies a large signaling overhead. Practical communication systems must consider the shortage of resources and prioritize carrying data instead of signaling. Thus, the impact of CSI acquisition strategies with reduced signaling on the proposed SDMA grouping solution is a relevant research question, but is out of the scope of this work and is left for future studies.

### APPENDIX A

#### COMPUTATIONAL COMPLEXITY K-MEANS CLUSTERING ALGORITHM

According to the pseudo-code Algorithm 1, the initialization of the  $k$ th centroid defined on line 3 has a computational complexity  $O(M)$ . This command determines the copy of the dominant eigenvector  $\mathbf{d}_j \in \mathbb{C}^{M \times 1}$  to the centroid vector  $\mathbf{v}_k$ . This process occurs inside a loop of  $K$  iterations defined in line 2. Therefore, the initialization of centroids has computational complexity  $O(KM)$ . The assignment step determines the association of each UE to a specific cluster. Line 10 uses Eq. 8 determines the association of each UE to a specific cluster. Since it involves the minimization of a difference of vectors  $\mathbf{d}_j \in \mathbb{C}^{M \times 1}$  and  $\mathbf{v}_k$ , it has computational complexity  $O(M)$ . This step is performed inside two loops of  $K$  and  $J$  iterations. Therefore, the assignment step has computational complexity  $O(JKM)$ . The centroid update is defined according to Eq. 7. This operation involves vectors of dimension  $M \times 1$  inside a loop of  $K$  iterations. Thus, the computational complexity of

this step is  $O(KM)$ . Therefore, the computational complexity of the KMAC in a process with a total of  $T$  iterations is  $O(KM) + O(TJKM) + O(TKM) = O(TJKM)$ .

### APPENDIX B

#### COMPUTATIONAL COMPLEXITY

##### AGGLOMERATIVE HIERARCHICAL CLUSTERING ALGORITHM

According to the pseudo-code Algorithm 2, the computation of the distance between the clusters  $C_j$  has computational complexity  $O(J^2M)$  since it requires the computation of Euclidean distances between vectors with dimension  $M \times 1$  inside two loops with  $J$  iterations. The determination of the closest pair of clusters has computational complexity  $O(|C|^2)$  since it requires the evaluation of a data structure with  $|C|^2$  elements. Initially, the set of clusters has  $|C| = J$  elements, decreasing by 1 at each iteration. Therefore, this operation has computational complexity  $O(J^2)$ . After the merge step, the distances from the merged cluster to the other clusters have to be recomputed, whereas the distance between the other clusters remain the same. Therefore, we compute  $J-t$  distances at the  $t$ th step. This operation has computational complexity  $(J-t)O(J^2M) = O(J^3M)$ . The merging process is repeated until the desired number of clusters is reached. Therefore, the computational complexity of the AHC in a process with a total of  $J-K$  iterations is  $O(J^2M) + (J-K)(O(J^2) + O(J^3M)) = O(J^3KM)$ .

### APPENDIX C

#### COMPUTATIONAL COMPLEXITY

##### DENSITY-BASED SPATIAL CLUSTERING OF APPLICATIONS WITH NOISE

The determination of the  $\epsilon$ -neighborhood has a computational complexity  $O(M)$ , since it requires the Euclidean distance of a vector with dimension  $M \times 1$ . This operation is repeated  $J$  times for each UE and occurs inside a loop of  $J$  iterations. Therefore, the computational complexity to determine the  $\epsilon$ -neighborhood of all UEs is  $O(J^2M)$ . Since this operation represents the main computational effort of the algorithm, its computational complexity is  $O(J^2M)$ . For more details regarding the computational complexity of the DBSCAN algorithm, see [25], [35].

### REFERENCES

- [1] M. Agiwal, A. Roy, and N. Saxena, "Next Generation 5G Wireless Networks: A Comprehensive Survey," *IEEE Communications Surveys and Tutorials*, vol. 18, no. 3, pp. 1617–1655, 2016. doi: 10.1109/COMST.2016.2532458.
- [2] E. Castaneda, A. Silva, A. Gameiro, and M. Kountouris, "An Overview on Resource Allocation Techniques for Multi-User MIMO Systems," *IEEE Communications Surveys and Tutorials*, vol. 19, no. 1, pp. 239–284, 2017. doi: 10.1109/COMST.2016.2618870.
- [3] T. F. Maciel and A. Klein, "On the Performance, Complexity, and Fairness of Suboptimal Resource Allocation for Multiuser MIMO OFDMA Systems," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 1, pp. 406–419, Jan. 2010. doi: 10.1109/TVT.2009.2029438.
- [4] A. Zubow, J. Marotzke, D. Camps-Mur, and X. Perez-Costa, "sGGSA: An SDMA-OFDMA Scheduling Solution," in *European Wireless Conference*, Apr. 2012. [Online]. Available: <http://ieeexplore.ieee.org/document/6216823/>.

- [5] Y. J. Zhang and K. B. Letaief, "An Efficient Resource-Allocation Scheme for Spatial Multiuser Access in MIMO/OFDM Systems," *IEEE Transactions on Communications*, vol. 53, no. 1, pp. 107–116, Jan. 2005. doi: 10.1109/TCOMM.2004.840666.
- [6] J. Nam, A. Adhikary, J.-Y. Ahn, and G. Caire, "Joint Spatial Division and Multiplexing: Opportunistic Beamforming, User Grouping and Simplified Downlink Scheduling," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 876–890, Oct. 2014. doi: 10.1109/JSTSP.2014.2313808.
- [7] M. Moretti, L. Sanguinetti, and X. Wang, "Resource Allocation for Power Minimization in the Downlink of THP-Based Spatial Multiplexing MIMO-OFDMA Systems," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 1, pp. 405–411, Jan. 2015. doi: 10.1109/TVT.2014.2320587.
- [8] Y. Xu, G. Yue, N. Prasad, S. Rangarajan, and S. Mao, "User Grouping and Scheduling for Large Scale MIMO Systems with Two-Stage Precoding," in *2014 IEEE International Conference on Communications (ICC)*, 2014, pp. 5197–5202. doi: 10.1109/ICC.2014.6884146.
- [9] Y. Xu and G. Yue and S. Mao, "User Grouping for Massive MIMO in FDD Systems: New Design Methods and Analysis," *IEEE Access*, vol. 2, pp. 947–959, Aug. 2014. doi: 10.1109/ACCESS.2014.2353297.
- [10] W. V. F. Mauricio, D. C. Araujo, F. H. C. Neto, F. R. M. Lima, and T. F. Maciel, "A Low Complexity Solution for Resource Allocation and SDMA Grouping in Massive MIMO Systems," 2018, pp. 1–6.
- [11] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint Spatial Division and Multiplexing—The Large-Scale Array Regime," *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6441–6463, Oct. 2013. doi: 10.1109/TIT.2013.2269476.
- [12] F. H. C. Neto and T. F. Maciel, "SDMA Grouping for 5G Systems," in *Brazilian Symposium on Telecommunications and Signal Processing*, Sep. 2019.
- [13] J. Zhang, X. Yu, and K. B. Letaief, "Hybrid Beamforming for 5G and Beyond Millimeter-Wave Systems: A Holistic View," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 77–91, 2020, issn: 2644-125X. doi: 10.1109/OJCOMS.2019.2959595.
- [14] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed Channel Sensing: A New Approach to Estimating Sparse Multipath Channels," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1058–1076, Jun. 2010. doi: 10.1109/JPROC.2010.2042415.
- [15] Richard O. Duda, Peter E. Hart, David G. Stork, *Pattern Classification*, 2nd. Wiley, Nov. 2000, ISBN: 978-0-471-05669-0.
- [16] C. Jiang, H. Zhang, Y. Ren, Z. Han, K. Chen, and L. Hanzo, "Machine Learning Paradigms for Next-Generation Wireless Networks," *IEEE Wireless Communications*, vol. 24, no. 2, pp. 98–105, Apr. 2017, issn: 1558-0687. doi: 10.1109/MWC.2016.1500356WC.
- [17] M. E. Morcho-Cayamcela, H. Lee, and W. Lim, "Machine Learning for 5G/B5G Mobile and Wireless Communications: Potential, Limitations, and Future Directions," *IEEE Access*, vol. 7, pp. 137 184–137 206, 2019, issn: 2169-3536. doi: 10.1109/ACCESS.2019.2942390.
- [18] M. J. Zaki and W. M. Jr, *Data Mining and Analysis: Fundamental Concepts and Algorithms*. Cambridge University Press, 2014.
- [19] D. Xu and Y. Tian, "A comprehensive survey of clustering algorithms," *Annals of Data Science*, vol. 2, pp. 165–193, 2015. doi: 10.1007/s40745-015-0040-1.
- [20] Christian Hennig and Marina Meila and Fionn Murtagh and Roberto Rocci, Ed., *Handbook of Cluster Analysis*. Taylor & Francis Group, 2016, ISBN: 13: 978-1-4665-5189-3.
- [21] A. Saxena, M. Prasad, A. Gupta, N. Bharill, O. P. Patel, A. Tiwari, M. J. Er, W. Ding, and C.-J. Lin, "A Review of Clustering Techniques and Developments," *Neurocomputing*, vol. 267, pp. 664–681, 2017. doi: 10.1016/j.neucom.2017.06.053.
- [22] S. Lloyd, "Least Squares Quantization in PCM," *IEEE Transactions on Information Theory*, vol. 28, no. 2, pp. 129–137, 1982, issn: 1557-9654. doi: 10.1109/TIT.1982.1056489.
- [23] M. Golasowski and J. Martinovic and K. Slaninova, "Comparison of K-means Clustering Initialization Approaches with Brute-Force Initialization," *Advanced Computing and Systems for Security*, vol. 567, pp. 103–114, Mar. 2017. doi: 10.1007/978-981-10-3409-1\_7.
- [24] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A Density-Based Algorithm for Discovering Clusters a Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise," in *International Conference on Knowledge Discovery and Data Mining*, Portland, Oregon, 1996, 226–231.
- [25] E. Schubert, and J. Sande M. Ester and H. P. Kriegel and X. Xu, "DBSCAN Revisited, Revisited: Why and How You Should (Still) Use DBSCAN," *ACM Transactions Database Systems*, vol. 42, no. 3, Jul. 2017, issn: 0362-5915. doi: 10.1145/3068335.
- [26] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd, J. W. Sons, Ed. John Wiley & Sons, 2006, ISBN: 978-0471241959.
- [27] A. A. Zaidi, R. Baldemair, H. Tullberg, H. Bjorkegren, L. Sundstrom, J. Medbo, C. Kilinc, and I. D. Silva, "Waveform and Numerology to Support 5G Services and Requirements," *IEEE Communications Magazine*, vol. 54, no. 11, pp. 90–98, Nov. 2016. doi: 10.1109/MCOM.2016.1600336CM.
- [28] M. Patzold, *Mobile Radio Channels*, 2nd. John Wiley & Sons, 2012, ISBN: 978-0-470-51747-5.
- [29] M. Peter and K. Sakaguchi, "Measurement Campaigns and Initial Channel Models for Preferred Suitable Frequency Ranges," Millimetre-Wave Based Mobile Radio Access Network for Fifth Generation Integrated Communications (mmMAGIC), Deliverable 2.1, Mar. 2016. doi: H2020-ICT-671650-mmMAGIC/D2.1. [Online]. Available: [https://bscw.5g-mmmagic.eu/pub/bscw.cgi/d94832/mmMAGIC\\_D2-1.pdf](https://bscw.5g-mmmagic.eu/pub/bscw.cgi/d94832/mmMAGIC_D2-1.pdf).
- [30] M. P. Tan and C. A. Floudas, "Encyclopedia of Optimization," in C. A. Floudas and P. M. Pardalos, Ed. Boston, MA: Springer US, 2009, ch. Determining the Optimal Number of Clusters, pp. 687–694, ISBN: 979-0-387-74759-0. doi: 10.1007/978-0-387-74759-0\_123.
- [31] S. Zhou, Z. Xu, and F. Liu, "Method for Determining the Optimal Number of Clusters Based on Agglomerative Hierarchical Clustering," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 12, pp. 3007–3017, Dec. 2017, issn: 2162-2388. doi: 10.1109/TNNLS.2016.2608001.
- [32] C. Patil and I. Baidari, "Estimating the Optimal Number of Clusters k in a Dataset Using Data Depth," *Data Science and Engineering*, vol. 4, pp. 132–140, Jun. 2019. doi: 10.1007/s41019-019-0091-y.
- [33] T. Calinski and J. Harabasz, "A Dendrite Method for Cluster Analysis," *Communications in Statistics*, vol. 3, no. 1, pp. 1–27, 1974. doi: 10.1080/03610927408827101.
- [34] F. H. Costa Neto and D. C. Araujo and T. F. Maciel, "Hybrid Beamforming Design based on Unsupervised Machine Learning for Millimeter Wave Systems," *International Journal of Communication Systems*, vol. 33, no. 5, Jan. 2020. doi: 10.1002/dac.4276.
- [35] J. Gan and Y. Tao, "DBSCAN Revisited: Mis-Claim, Un-Fixability, and Approximation," in *International Conference on Management of Data*, Melbourne, Victoria, Australia, May 2015, 519–530, ISBN: 9781450327589. doi: 10.1145/2723372.2737792.



**Francisco Hugo Costa Neto** received the B.Sc. degree in Electrical Engineering in 2014 and M.Sc. degree in Teleinformatics Engineering in 2016, both from Federal University of Ceará (UFC), Fortaleza, Brazil. He is currently a Ph.D. student and researcher at the Wireless Telecommunications Research Group (GTEL), UFC, where he works on projects in cooperation with Ericsson Research. Some topics of his research interests include radio resource management, hybrid beamforming, numerical optimization, and machine intelligence.



**Tarcisio Ferreira Maciel** received the B.Sc. and M.Sc. degrees in electrical engineering from the Federal University of Ceará (UFC), in 2002 and 2004, respectively, and the Dr.Ing. degree in electrical engineering from the Technische Universität Darmstadt (TUD), Germany, in 2008. Since 2001, he has actively participated in several projects in a technical and scientific cooperation between the Wireless Telecom Research Group (GTEL), UFC, and Ericsson Research. From 2005 to 2008, he was a Research Assistant with the Communications Engineering Laboratory, TUD. Since 2008, he has been a member of the Postgraduation Program in Teleinformatics Engineering, UFC. In 2009, he was a Professor of computer engineering with UFC-Sobral. Since 2010, he has been a Professor with the Center of Technology, UFC. His research interests include radio resource management, numerical optimization, and multiuser/multiantenna communications.