# Adaptive Precoding with DBSCAN Grouping for Efficient Massive MIMO Beamforming

Kanwalinderjit Kaur, Elias Farmer, Priyanshu Luhar
*Department of Computer and Electrical Engineering and Computer Science*
*California State University, Bakersfield*
Bakersfield, United States
kgagnej@csub.edu, efarmer2@csub.edu, pluhar@csub.edu

***Abstract*—Massive Multiple-Input Multiple-Output (M-MIMO) is a key technology for 5G networks that allows multiple transmitters to communicate with multiple receivers simultaneously. This paper proposes a hybrid beamforming technique that uses the Density-Based Spatial Clustering or Application with Noise (DBSCAN) clustering algorithm to assign users into groups. A round-robin scheduling is used to transmit data to one user in each subgroup at a time to maximize the angular separation of each signal to reduce interference. The digital baseband switches between Zero-Forcing (ZF) and Minimum Mean Square Error (MMSE) to minimize the Bit Error Rate (BER), Latency, and improve the throughput. With this approach, overall gain in BER is 8%, latency is 20%, and throughput is 1% in comparison to Voronoi, K-Means, and Tabu grouping approaches.**

***Keywords—Beamforming, Massive MIMO, Precoding, 5G***

## I. INTRODUCTION

The need for fast and reliable wireless communication has increased significantly in recent years, driven by the rapid growth of personal devices and the Internet of Things (IoT) [1]. The current wireless standard, 5G, developed by 3GPP, is designed to deliver high-speed and low-latency connectivity through the extensive use of Massive MIMO technology [2]. Massive MIMO allows multiple transmitter antennas to simultaneously transmit data to multiple users on the same frequency band. Many MIMO techniques exist and are used to varying degrees, such as Spatial Multiplexing, Ray tracing, and Hybrid Beamforming. These approaches allow Massive MIMO to achieve high network throughput and improved performance, even in dense network environments [3].

Spatial Multiplexing, a Massive MIMO technique, enables the transmitter to transmit multiple independent data streams, with each being able to work independently or with the other [4]. To model the channel in a realistic indoor scenario and evaluate its performance, raytracing is used to simulate how the environment results in copies of the signal due to reflection, refraction, and diffraction. This model measures the Power Delay, Path Loss, and Bit Error Rate (BER), making it easier to optimize Massive MIMO systems [5].

In beamforming, the transmitter antennas create electronic beams that are steered toward user equipment by adjusting the phase and amplitude of the antennas [2][6]. The three main beamforming techniques are Analog, Digital, and Hybrid. With analog beamforming, a singular Radio Frequency (RF) transceiver phase-shifts and amplifies its output before being sent to an antenna element. On the other hand, with digital beamforming, each antenna has its own transceiver. Hybrid Beamforming combines both methods by forming the analog beam on the RF stage and performing digital beamforming on the baseband level, saving power [6].

The objective of this research is to write a new precoding technique and apply a grouping mechanism to improve the performance of the new beamforming technique. For objective one, a new adaptive precoding technique is implemented that dynamically switches between ZF and MMSE to take advantage of their performance differences at varying SNR. For the second objective, the DBSCAN grouping mechanism is applied for user grouping. The two objectives are combined to improve the performance of the implemented technique and are compared with [7], [8], and [9].

This research introduces a novel approach to beamforming using adaptive precoding, which automatically switches between ZF and MMSE precoding on the digital baseband to help maximize performance. A new beamforming technique is implemented that makes use of DBSCAN clustering to put the users in groups who are ready to receive data. To test the effectiveness of the proposed beamforming technique, simulations are performed where users have a direct line of sight (LoS) with the transmitter.

The paper is structured as follows. Section II reviews existing studies on MIMO techniques and highlights research gaps that inspire the proposed beamforming method. Section III details the design, simulation setup, and implementation process, while Section IV compares adaptive precoding results with Cheng et al. and evaluates the performance of DBSCAN, Voronoi, K-Means, and Tabu clustering approaches.

## II. LITERATURE REVIEW

The integration of Tabu-Voronoi clustering heuristics with key management schemes has shown significant improvements in energy efficiency and communication security for heterogeneous sensor networks, particularly through dynamic cluster boundary adaptation using Tabu optimization techniques [5]. Similarly, a QoS-based heuristic clustering approach has been proposed to optimize latency, reliability, and energy utilization in two-tier sensor networks, achieving enhanced throughput and service performance under varying network conditions [6].

Paper [10] proposes a grouping technique for an mmWave-NOMA system to address the problem where user signals may be highly correlated with each other because multiple users may either try to pick the same beam or very similar beams from a Discrete Fourier Transform (DFT) codebook. They propose a novel agglomerative nesting (AGNES) grouping algorithm that takes advantage of the signal, initially assigning users into their own group, before repeatedly merging groups based on the correlation of their channels. With D-AGNES, each group is assigned a beam that only they can use before it is deleted from the DFT codebook, with D-AGNES prioritizing groups that achieve the highest beam gain. To address the problem of picking similar beams, S-AGNES uses a similar approach to D-AGNES, where once a group selects the best beam from the DFT Codebook, instead of removing the beam, the remaining users have their channels modified so that they do not interfere with the selected beam. Paper [11] proposes an mmWave-NOMA system that creates user groups by using K-Means clustering. In their system, each group's representative is updated to be the one with the lowest channel correlation, and this is repeated until the representatives of every group no longer change.

Kumuda et al's work [6] is focused on the development of a basic Hybrid Beamforming (HBF), where they raised concerns about digital beamforming (DBF) and its higher hardware and signal processing costs compared to analog beamforming (ABF). Their simulations compared all three beamforming techniques where three users were served at various distances; results indicated that DBF had the highest capacity and ABF had the lowest interference. Their proposed HBF is a good trade-off between the two techniques. While this is important towards the advancement of HBF, it does not consider the higher number of users, different modulations, and adaptive precoding.

Several precoding techniques exist to help with network performance. ZF precoding is used to help cancel out inter-user interference, though it can lead to performance degradation when users have serious fading channels due to the channel noise being amplified [12]. MMSE precoding addresses the problems of ZF precoding by balancing interference suppression and noise amplification, improving overall network throughput. Beulah and Markkandan's paper [13] compared the performance of MMSE precoding with two other precoding techniques, Tomlinson-Harashima and Block Diagonalization. The paper performed multiple simulations using a variety of transmitter antennas and considering whether users have the perfect Channel State Information (CSI), before comparing the resulting BER of each precoding technique, with MMSE achieving the lowest BER. While the paper sufficiently provided the implementation details of each precoding technique and their performance, it opens a potential possibility of switching between precoding methods. Should the noise in a user's channel be low due to their SNR being high, the performance between MMSE and ZF precoding is less pronounced, with their performance [12].

Overall, most of these works have a great individualistic approach to Massive MIMO via beamforming a clustering method, or a precoding technique optimization. However, there are clear gaps that need to be addressed, such as the need for an adaptive precoding to decide between MMSE and ZF. DBSCAN has not yet been implemented on a beamforming topology with a higher number of users, and the different modulations needed to provide relevance in accordance with 3GPP standards. This research work aims to fill these important research gaps and give a unified approach to adaptive precoding change with the optimization of DBSCAN.

## III. METHODOLOGY

### *A. System Model*

In this research work, the transmitter has 64 transmitter antennas ($N_T$) on a Uniform Linear Array (ULA). Each transmitter antenna is equipped with a 5G C band, allowing it to transmit signals with a frequency of 3.5 GHz. The height of the antenna array is 45m above the ground. The transmitter is equipped with 4 RF Chains ($N_{RF}$) to transmit the signals at the speed of light (c). Let $\lambda$ be the speed of light (c) divided by the frequency. Each transmitter antenna is separated by $\lambda/2$ meters.

In the given environment, there are 300 users on the ground level randomly positioned in a range of 250m and 500m radius around the transmitter, with each user having 4 receiver antennas (NR) on a ULA. In the simulation, 20 random users (K) are ready to receive data from the transmitter. For this system model, perfect CSI is assumed. While practical systems are subject to channel estimation errors, this assumption allows us to isolate and analyze the specific performance gains of adaptive precoding and DBSCAN grouping.

The transmitter antennas transmit 100,000 ($N_s$) symbols to each served user with a symbol rate ($s_r$) of 1,000,000 symbols per second. To calculate the effectiveness of the technique, the BER, Throughput, and Latency are all calculated on an SNR range between –5 to 20 dB. The transmitted signals are modulated using the 64QAM and 256QAM modulation schemes, as specified by 3GPP technical standards Release 19 [14].

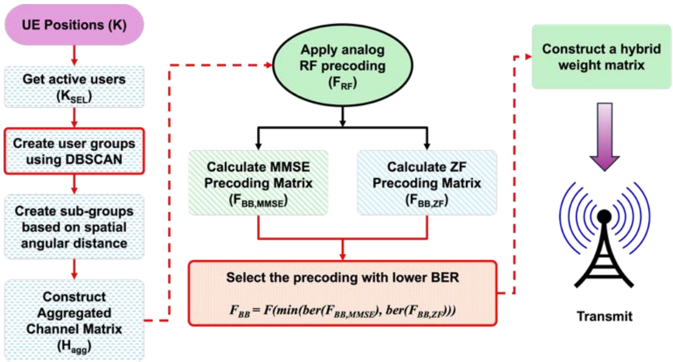

Fig. 1. A flow chart illustrating the process of DBSCAN grouping and adaptive precoding.

Figure 1 presents the methodology of the adaptive beamforming technique, where DBSCAN is first used to assign users to groups, followed by splitting each group into subgroups by performing DBSCAN again on each newly created group. Afterwards, round-robin scheduling is used for each group to select $N_{RF}$ users for data transmission. The selected user channels are used to form the aggregate channel matrix before applying analog precoding. The adaptive precoder then applies either MMSE or ZF precoding, based on which would yield the lowest BER. The final weight matrix is constructed after applying analog and digital precoding before transmitting the signal to each selected user in the matrix.

### *B. User Grouping*

The proposed technique uses DBSCAN to assign users to groups to help minimize signal interference. The DBSCAN algorithm first iterates through all K users assigns them to groups based on the spatial angular difference ($\varepsilon_{group}$) of 15 degrees between the users. Equation 1 defines the minimum number of users, $minPts_{group}$ whose spatial angular difference is within $\varepsilon_{group}$, needed to create a group.

$$minPts_{group} = \max(2, \mathrm{N_{RF}}) \tag{1}$$

Afterwards, DBSCAN is used for each group again to each newly created group into subgroups, with $\varepsilon_{subgroup}$ of 3 degrees and $minPts_{subgroup}$ of 2. Should DBSCAN fail to add a user to a group and treat them as noisy, a group for each individual user is made instead. Unlike centroid-based algorithms such as K-means, DBSCAN strictly groups only high-density users and serves outliers independently, which prevents the angular spread of users in a given group from being inflated and degrading performance. Figure 2 shows the 20 users in the simulation being assigned to different groups, denoted by the dashed ellipses.

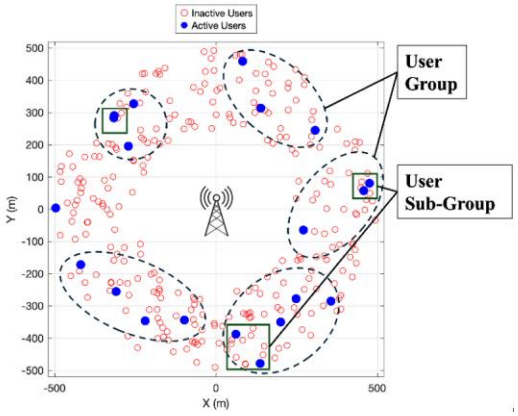

Fig. 2. A scatter plot of how users are being clustered together into groups and subgroups.

Algorithm I provides a high-level look at how groups and subgroups are made with DBSCAN clustering and how users are processed with the round-robin scheduling technique. First, the initial groups are created for K users using DBSCAN. DBSCAN is used again for each newly created group to create the subgroups. Algorithm I iterates through each set of subgroups and employs the round-robin scheduling technique to select a batch of users from the subgroup for data transmission. The aggregated channel matrix is then constructed from each user channel, then analog precoding is applied to create the effective channel matrix. The adaptive precoding technique then simulates the transmission using MMSE and ZF precoding on the digital baseband. The precoding technique that results in the lowest BER is used to create the final weight matrix, before transmitting the signal. If the BER of ZF precoding becomes lower than MMSE, ZF precoding will only be used to prevent redundant simulations.

**Algorithm I:** Adaptive Precoding with User Grouping

```
$G_{main} \leftarrow$ DBSCAN(K, $\epsilon_{group}$, MinPts$_{group}$)
$G_{sub} \leftarrow \{\}$
for each group g in $G_{main}$ do
    subs $\leftarrow$ DBSCAN(g, $\epsilon_{subgroup}$, MinPts$_{subgroup}$)
    $G_{sub}$.add(subs)
end for
for each subgroupSet in $G_{sub}$ do
    while UnservedUsersIn(subgroupSet) do
        $K_{sel} \leftarrow$ RoundRobinSelect(subgroupSet, $N_{RF}$)
        if not IsEmpty($K_{sel}$) then
            $H_{agg} \leftarrow$ ConstructChannelMatrix($K_{sel}$)
            $F_{RF} \leftarrow$ CalculateAnalogPrecoder($K_{sel}$)
            $H_{eff} \leftarrow H_{agg} * F_{RF}$
            Switched $\leftarrow$ false;
            if not switched then
                $F_{BB,ZF} \leftarrow$ CalculateZF(Heff)
                $F_{BB,MMSE} \leftarrow$ CalculateMMSE($H_{eff}$)
                if BER($F_{BB,ZF}$) < BER($F_{BB,MMSE}$) then
                    $F_{BB} \leftarrow F_{BB,ZF}$
                    Switched $\leftarrow$ true;
                else
                    $F_{BB} \leftarrow F_{BB,MMSE}$
                end if
            else
                $F_{BB} \leftarrow$ CalculateZF(Heff)
            end if
            W $\leftarrow$ Normalize($F_{RF} * F_{BB}$)
            TransmitSignal(W, $K_{sel}$)
        end if
    end while
end for
```

### *C. Adaptive Precoding*

After all users have been grouped using DBSCAN, a round-robin scheduling technique is performed on each group to select a subset of users, $K_{sel}$, that are ready for data transmission.

The transmitter and receiver steering vectors, $sv_{rx}(\theta_i)$ and $sv_{tx}(\theta_i)$ respectively, are constructed for each user $i \in K_{sel}$, where $\theta_i$ is the azimuth angle from the transmitter antennas to each user. $sv_{tx}(\theta_i)$ is an $N_T \times 1$ column vector of the transmitter's antenna array signal, while $sv_{rx}(\theta_i)$ is an $N_R \times 1$ column vector of the receiver's antenna array signal [15].

In equation (3), the LoS channel for each user ($H_i$) is constructed from the product of both steering vectors from equations (2a) and (2b). Let $(\cdot)^H$ denote the Hermitian transpose of the given matrix.

$$sv_{tx}(\theta) = \underbrace{\begin{bmatrix} 1 \\ e^{-j\pi sin(\theta)} \\ e^{-2j\pi sin(\theta)} \\ \vdots \\ e^{-(N_T-1)j\pi sin(\theta)} \end{bmatrix}}_{N_T \times 1} \tag{2a}$$

$$sv_{rx}(\theta) = \underbrace{\begin{bmatrix} 1 \\ e^{-j\pi sin(\theta)} \\ e^{-2j\pi sin(\theta)} \\ \vdots \\ e^{-(N_R-1)j\pi sin(\theta)} \end{bmatrix}}_{N_R\times 1} \tag{2b}$$

$$\underbrace{H_i}_{N_R\times N_T} = \underbrace{sv_{rx}(\theta)}_{N_R\times 1} * \underbrace{sv_{tx}^{H}(\theta)}_{1\times N_T} \tag{3}$$

A subset of the individual channel matrix from equation (3) is created to store one receiver antenna from each user. Let $|K_{sel}|$ be the number of selected users. The aggregated channel matrix, $H_{agg}$ is created by vertically stacking the first row vector ($1\times N_T$) from each user's individual channel matrix $H_i$, which allows simultaneous data transmission.

$$H_{agg} = \underbrace{\begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_{|K_{sel}|} \end{bmatrix}}_{|K_{sel}|\times N_T} \tag{4}$$

The adaptive precoding process is used to apply the analog and digital precoder matrices onto the aggregated channel matrix from equation (4). The analog precoding matrix, $F_{RF}$ is the transmitter steering vectors concatenated as columns, where each analog beam is steered toward a selected user.

$$F_{RF} = \underbrace{\left[sv_{tx}(\theta_1), sv_{tx}(\theta_2), \dots, sv_{tx}\left(\theta_{|K_{sel}|}\right)\right]}_{N_T\times|K_{sel}|} \tag{5}$$

The effective channel matrix, $H_{eff}$, is created by multiplying the aggregated channel matrix from equation (4) by the analog precoding matrix.

$$\underbrace{H_{eff}}_{|K_{sel}|\times|K_{sel}|} = \underbrace{H_{agg}}_{|K_{sel}|x N_T} * \underbrace{F_{RF}}_{N_T\times|K_{sel}|} \tag{6}$$

With an analog beam pointing at each user, the digital baseband precoding matrix, $F_{BB}$ is applied to the effective channel matrix to reduce the inter-signal interference between the users and improve performance. As shown in Algorithm I, the selected digital precoding matrix is either ZF ($F_{BB,ZF}$) or MMSE ($F_{BB,MMSE}$) based on which results in the lowest BER. The ZF precoding matrix is constructed by taking the pseudoinverse of the effective channel matrix:

$$F_{BB,ZF} = H_{eff}^{H} * \left(H_{eff} * H_{eff}^{H}\right)^{-1} \tag{7}$$

In a low SNR range where a lot of noise is present in the effective channel, the pseudo-inverse of the effective channel results in the noise being amplified, degrading performance. The adaptive precoder uses MMSE when the noise amplification would result in a lower BER. Let the regularization factor, α, be $|K_{sel}|/SNR_{linear}$.

$$SNR_{linear} = 10^{\left(\frac{SNR_{dB}}{10}\right)} \tag{8}$$

$$F_{BB,MMSE} = H_{eff}^{H} * \left(H_{eff} * H_{eff}^{H} + \alpha I\right)^{-1} \tag{9}$$

Once the final digital precoding matrix is chosen, the final weight matrix (W) is constructed from the product of $F_{RF}$ and $F_{BB}$. The product is then divided by its Frobenius norm to meet the system's fixed total transmission power constraint.

$$\underbrace{W}_{N_T\times|K_{sel}|} = \frac{\underbrace{F_{RF}}_{N_T\times|K_{sel}|} \cdot \underbrace{F_{BB}}_{|K_{sel}|\times|K_{sel}|}}{\left|\left|F_{RF}\cdot F_{BB}\right|\right|_F} \tag{10}$$

## D. *Signal Transmission*

The transmitted signal matrix, x is constructed from the product of the weight matrix and the symbol matrix (s), where the symbol matrix is the $N_s$ symbols transmitted for the selected users.

$$\underbrace{x}_{N_T\times N_S} = \underbrace{W}_{N_T\times|K_{sel}|} \cdot \underbrace{s}_{|K_{sel}|\times N_S} \tag{11}$$

$$Y_i = h_i w_i s_i + \sum_{j=1,\, j\neq i}^{|K_{sel}|} \left(h_i w_j s_j\right) + n \tag{12}$$

Equation (12) models the signal that user receives from equation (11) as $y_i$, which contains inter-user signal interference and additive white gaussian noise, n. Let $w_i$ be the precoding vector from user i from taking the i-th column of the final weight matrix W. Let $s_i$ be the i-th row of s which represents the symbols transmitted to user i. The desired signal is calculated from the product of $h_i$, $w_i$, and $s_i$. The signal interference is the sum of the signals belonging to other users in the group that are receiving data.

## E. *Performance Metrics*

To measure the effectiveness of the beamforming approach, BER, Throughput, and Latency are calculated. The BER is calculated by taking the ratio of incorrect bits to total bits transmitted.

$$BER = \frac{Number\ of\ Errors}{Number\ of\ Bits\ Transmitted} \tag{13}$$

$$Throughput = \log_2(M) \cdot s_r \cdot (1 - BER) \tag{14}$$

The final performance metric calculated is the latency (L). At a high level, the latency is calculated by adding together the processing delay ($L_{proc}$), propagation delay ($L_{prop}$), and transmission delay ($L_{tx}$). The processing delay is the sum of the time required for the channel estimation for the current batch of users ($L_{est}$), the time it takes to calculate the digital baseband precoder ($L_{prec}$), and the time it takes to decode the received symbols ($L_{dec}$). The propagation delay is the time it takes for the signal to travel to the receiver. Finally, the transmission delay is the time it takes for the transmitter to transmit the symbols.

$$L = L_{prop} + L_{tx} + L_{proc} \tag{15}$$

Equation (16) defines the propagation delay, where the distance between the transmitter and user (d) is divided by the signal speed (c). The transmission delay depends on the current Signal-to-Interference-plus-Noise Ratio (SINR) and the number of symbols being transmitted to the user.

$$L_{prop} = \frac{d}{c} \tag{16}$$

Equation (17) defines the effective number of bits per symbol affected by the current SINR and the chosen modulation scheme, where the number of bits is capped based on the current modulation technique and the SINR is used over the SNR [6].

$$b_{eff} = \min(\log_2(M),\ \log_2(1 + SINR)) \tag{17}$$

Equation (18) determines the amount of time needed for the transmitters to transmit all $n_s$ symbols to a single user, where a higher modulation order requires less processing time.

$$L_{tx} = \frac{N_s}{s_r} \cdot \frac{\log_2(M)}{b_{eff}} \tag{18}$$

The processing delay is calculated by adding together the amount of time to estimate the channel delay from the users in the current group ($L_{est}$), the amount of time it takes to calculate $F_{BB}$ ($L_{prec}$), and the amount of time it takes to decode the transmitted symbols.

## IV. RESULTS & ANALYSIS

This research compares the throughput of the adaptive precoding technique, MMSE, and ZF together using some grouping algorithms. The BER and throughput of the adaptive precoding technique are compared with the results found in Cheng's paper [9]. This research then compares the latency of applying MMSE, ZF, and adaptive precoding with each other. Finally, the BER, throughput, and latency are compared in separate simulations of K-MEANS, Tabu, Voronoi, and DBSCAN clustering, which make use of adaptive precoding.

### A. Adaptive Precoding Analysis

Figure 3 presents a graph that shows the throughput when adaptive precoding is applied compared to only using MMSE and ZF precoding with an SNR range between -12 and 0. Figure 4 and Table I compare the latency of using MMSE, ZF, and Adaptive Precoding with QPSK and 16QAM modulation schemes. The latency of the transmitted signal when using ZF is much larger than the latency of Adaptive Precoding, which uses MMSE precoding within the SNR range due to the better performance it offers. As the SNR value increases, the latency of ZF begins to converge with the latency of MMSE precoding. As the modulation order increases, the overall latency increases due to the lower channel capacity as presented equation (17). The high latency is a result of 100,000 symbols being transmitted to each user to get an accurate BER and Throughput estimation.

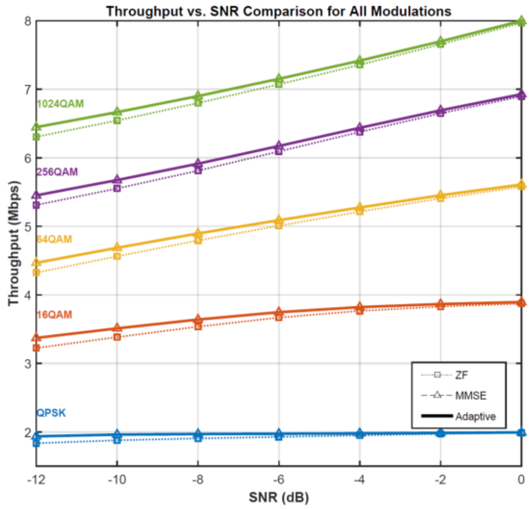

Fig. 3. Throughput comparison versus SNR with ZF, MMSE, and Adaptive Precoding.

TABLE I.

ADAPTIVE PRECODING LATENCY COMPARISON

| Method | QPSK | 16QAM |
|---|---|---|
| Adaptive | 1.1655s | 1.6983s |
| ZF | 1.8449s | 3.05884s |
| MMSE | 1.1655s | 1.6983s |

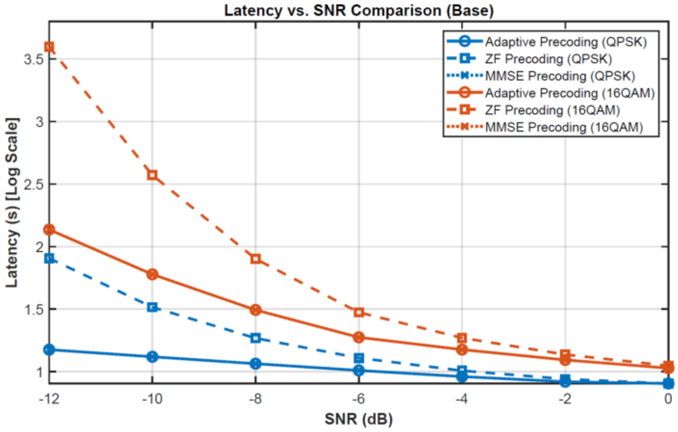

Fig. 4. Latency comparison using adaptive precoding with QPSK and 16QAM modulations. Each group of colored lines represents a different modulation

Figure 5 compares the resulting BER of the adaptive precoding technique with other techniques measured in the paper by Cheng et al, where it is measured on an SNR range between -12 and 0. One important distinction is that their paper uses 2 users for their simulation as opposed to this technique's 20 users. The rationale is that one of the main objectives of this research is to be tied to user grouping. Figure 5 compares the result of a variety of beamforming methods their paper covers with this research's proposed adaptive precoding technique. In a low SNR range, the use of adaptive precoding shows a higher BER and lower throughput compared to Cheng et al's measured techniques. At approximately -3 SNR, the performance of QPSK shows better results compared to Hybrid Precoding using MMSE and Two-Stage ZF precoding. Adaptive precoding using 16QAM performs worse than all of Cheng's measured techniques in a low SNR range.

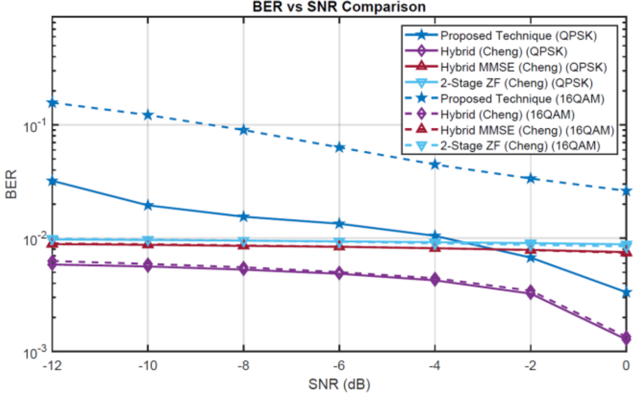

Fig. 5. BER versus SNR comparison between the proposed technique and against a variety of beamforming techniques measured in Cheng's paper.

## B. *Clustering with DBSCAN*

TABLE II.

MEAN METRIC COMPARISON FOR CLUSTERING METHODS

| Method | BER | Latency | Throughput |
|---|---|---|---|
| K-Means | 0.09637 | 3.31797 | $3.61\times10^6$ |
| Tabu | 0.09704 | 2.59703 | $3.61\times10^6$ |
| Voronoi | 0.09393 | 3.04071 | $3.62\times10^6$ |
| DBSCAN | 0.08964 | 2.64559 | $3.64\times10^6$ |

Table II and Figures 6 to 8 show a comparison among the clustering methods: K-Means, Tabu, Voronoi, and DBSCAN, using the same 16-QAM modulation setup. The first three techniques rely only on MMSE precoding, while DBSCAN integrates adaptive precoding, which improves the results noticeably. Figure 6 shows that DBSCAN achieves the lowest average BER of about 0.0896, which is around 7% better than K-Means, 8% better than Tabu, and 4.5% better than Voronoi.

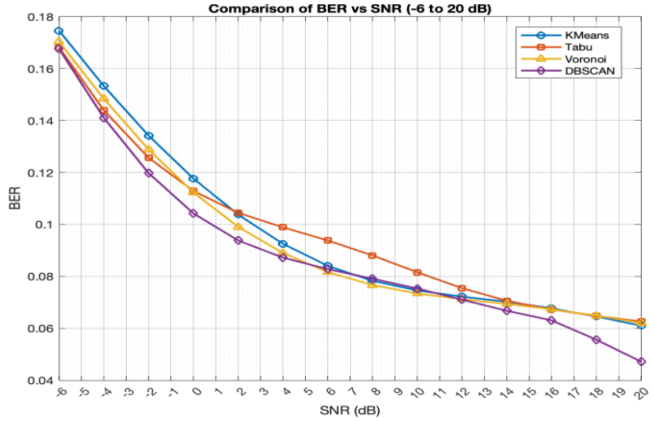

Fig. 6. Average Bit Error Rate (BER) versus SNR for K-Means, Tabu, Voronoi, and DBSCAN clustering methods under 16-QAM modulation.

Figure 6 indicates that the adaptive precoding allows DBSCAN to better adapt to noise variations and fading, providing stronger resistance to interference. The smoother BER curve for DBSCAN, along with increasing SNR, also suggests more stable transmission and less error fluctuation. Because DBSCAN groups users based on density rather than fixed distances, it can recognize overlapping or irregular signal clusters that MMSE-only methods fail to identify efficiently. In contrast, K-Means, Tabu, and Voronoi show limitations at lower SNR ranges due to their static grouping and linear precoding nature. Therefore, the combination of adaptive clustering and adaptive precoding allows DBSCAN to reduce bit errors effectively, giving it a clear advantage in terms of signal reliability and robustness in 5G NR systems.

Latency and throughput results further prove the strength of DBSCAN's adaptive approach. From Figure 7, DBSCAN's average latency is 2.65 ms, which is only 2% slower than Tabu, but 20% faster than K-Means and 13% faster than Voronoi. This means DBSCAN manages to maintain nearly the same responsiveness as Tabu while still providing much better error performance. Figure 8 shows that DBSCAN has the highest throughput at about $3.64\times10^6$ bps, which is 0.8% higher than Voronoi, 0.7% higher than K-Means, and 0.9% higher than Tabu.

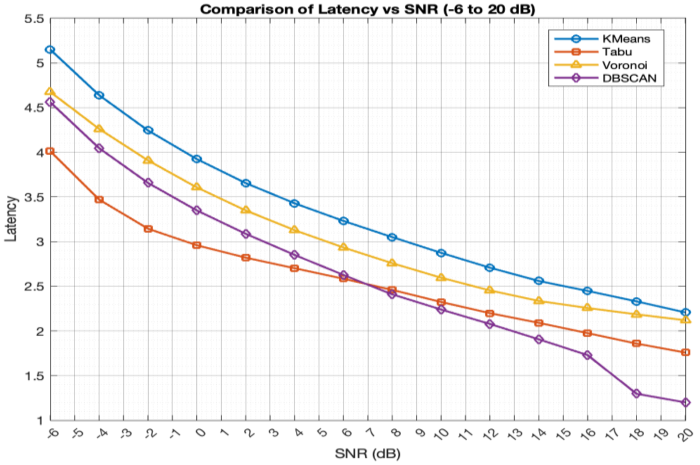

Fig. 7. Average latency performance versus SNR for the compared clustering algorithms using MMSE and adaptive precoding configurations.

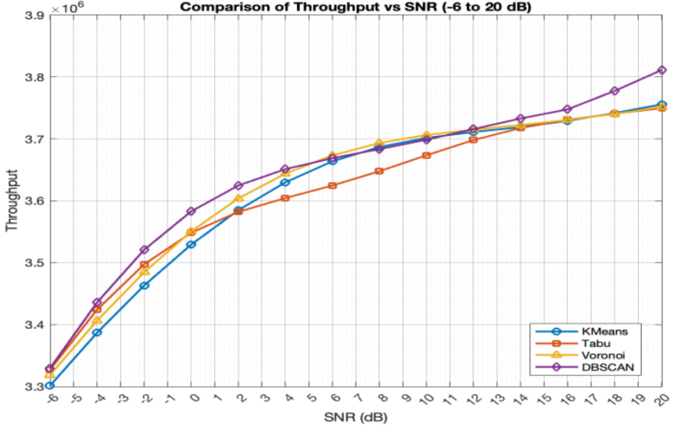

Fig. 8. Average throughput versus SNR showing the QoS improvement achieved by DBSCAN with adaptive precoding over MMSE-only clustering schemes.

Even though these throughput gains seem small, they consistently appear across the whole SNR range, meaning DBSCAN uses the spectrum more efficiently. Its dynamic precoding adjusts beamforming weights in real time according to CSI updates, ensuring more balanced data delivery and a lower retransmission rate. Meanwhile, MMSE-only clustering methods are static and do not adapt during changing channel conditions, which causes resource imbalance and throughput saturation at higher SNR. The density-based nature of DBSCAN also reduces cluster misclassification, improving spectral efficiency and lowering interference. Altogether, DBSCAN shows a balanced gain across all metrics, achieving lower BER, comparable latency, and higher throughput, making it about 5–10% overall better in combined QoS compared to the MMSE-only algorithms.

Table III shows the standard deviation of BER, Throughput, and Latency over 25 runs for clustering methods. It can be seen that Tabu gives the lowest variation in almost all the metrics, showing it has more stable and consistent behavior. For example, the standard deviation of BER and Throughput for Tabu are 1.40% and 1.39%, which are smaller compared to others. DBSCAN, while still performing well overall, has slightly higher variation in BER (1.90%) and Latency (1.82%), meaning it is more sensitive to parameter change. K-Means and Voronoi have moderate fluctuation and stay within a similar range. Overall, all four clustering methods stay under ±2%, which means the system results are stable.

TABLE III. STANDARD DEVIATION OF METRICS OVER 25 RUNS

| **Metric** | **Standard Deviation (±%)** | | | |
|---|---|---|---|---|
| | K-Means | Tabu | Voronoi | DBSCAN |
| BER | 1.77 | 1.40 | 1.81 | 1.90 |
| Throughput | 1.89 | 1.39 | 1.75 | 1.77 |
| Latency | 2.09 | 1.72 | 1.46 | 1.82 |

## V. CONCLUSION

This research proposed a new hybrid beamforming technique that involved dynamically switching between ZF and MMSE precoding, which offers the lowest BER and Latency on top of assigning all users connected to a transmitter into groups from the created clusters using the DBSCAN based on the spatial angular difference between each user. These groups are then split into subgroups by performing the same clustering algorithm to maximize the separation of the transmitted signals to help reduce cross-signal interference and improve the bit error rate. Benchmarking against other works showed that the dynamic precoding is more efficient, and when comparing the results against the Tabu, Voronoi, and K-Means clustering algorithms, results indicated that DBSCAN has accomplished about 8% lower BER and over 20% lower Latency than K-Means-based clustering.

## ACKNOWLEDGMENT

This research is supported by NSF award number #2318634. We want to thank NSF for supporting this research.

## REFERENCES

[1] L. Das, R. Raman Chandan, P. Kaur, A. Singh, A. Rana and B. D. Shivhare, "Advancements in Wireless Network Technologies for Enabling the (IoT): A Comprehensive Review," *2023 6th International Conference on Contemporary Computing and Informatics (IC3I)*, Gautam Buddha Nagar, India, 2023, pp. 807-814, doi: 10.1109/IC3I59117.2023.10397952.

[2] H. Jin *et al.*, "Massive MIMO Evolution Toward 3GPP Release 18," in *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 6, pp. 1635-1654, June 2023, doi: 10.1109/JSAC.2023.3273768.

[3] Z. Chen, X. Zhang, S. Wang, Y. Xu, J. Xiong and X. Wang, "Enabling Practical Large-Scale MIMO in WLANs With Hybrid Beamforming," in *IEEE/ACM Transactions on Networking*, vol. 29, no. 4, pp. 1605-1619, Aug. 2021, doi: 10.1109/TNET.2021.3073160.

[4] Z. Yuan, J. Zhang, V. Degli-Esposti, Y. Zhang and W. Fan, "Efficient Ray-Tracing Simulation for Near-Field Spatial Non-Stationary mmWave Massive MIMO Channel and Its Experimental Validation," in IEEE Transactions on Wireless Communications, vol. 23, no. 8, pp. 8910-8923, Aug. 2024, doi: 10.1109/TWC.2024.3357071.

[5] K. Kaur and P. Zavala, "Analysis of Signal Propagations Between an Access Point and IoT Devices in a Complex Indoor Environment Using Ray Tracing," *2024 International Symposium on Networks, Computers and Communications (ISNCC)*, Washington DC, DC, USA, 2024, pp. 1-7, doi: 10.1109/ISNCC62547.2024.10758988.

[6] K. D. K, M. Prerana, N. Himabindu and C. A. S, "Hybrid Beamforming for Wireless Communication System," *2025 3rd International Conference on Smart Systems for applications in Electrical Sciences (ICSSES)*, Tumakuru, India, 2025, pp. 1-5, doi: 10.1109/ICSSES64899.2025.11009513.

[7] K. K. Gagneja, K. E. Nygard and N. Singh, "Using Tabu-Voronoi clustering heuristics with key management scheme for heterogeneous sensor networks," 2012 Fourth International Conference on Ubiquitous and Future Networks (ICUFN), Phuket, Thailand, 2012, pp. 46-51, doi: 10.1109/ICUFN.2012.6261662

[8] K. K. Gagneja and K. E. Nygard, "A QoS based heuristics for clustering in two-tier sensor networks," 2012 Federated Conference on Computer Science and Information Systems (FedCSIS), Wroclaw, Poland, 2012, pp. 779-784.

[9] W. Ren, J. Deng and X. Cheng, "MMSE Hybrid Beamforming for Multi-User Millimeter Wave MIMO Systems," in IEEE Communications Letters, vol. 27, no. 12, pp. 3389-3393, Dec. 2023, doi: 10.1109/LCOMM.2023.3327451.

[10] J. Zhu, Q. Li, Z. Liu, H. Chen and H. V. Poor, "Enhanced User Grouping and Power Allocation for Hybrid mmWave MIMO-NOMA Systems," in *IEEE Transactions on Wireless Communications*, vol. 21, no. 3, pp. 2034-2050, March 2022, doi: 10.1109/TWC.2021.3109053.

[11] L. Zhu, J. Zhang, Z. Xiao, X. Cao, D. O. Wu and X. -G. Xia, "Millimeter-Wave NOMA With User Grouping, Power Allocation and Hybrid Beamforming," in *IEEE Transactions on Wireless Communications*, vol. 18, no. 11, pp. 5065-5079, Nov. 2019, doi: 10.1109/TWC.2019.2932070

[12] P. -H. Lin and S. -H. Tsai, "Performance Analysis and Algorithm Designs for Transmit Antenna Selection in Linearly Precoded Multiuser MIMO Systems," in *IEEE Transactions on Vehicular Technology*, vol. 61, no. 4, pp. 1698-1708, May 2012, doi: 10.1109/TVT.2012.2190435.

[13] V. Angeline Beulah and S. Markkandan, "Performance analysis of precoding techniques for Massive MU-MIMO systems," *2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, Coimbatore, India, 2015, pp. 1-5, doi: 10.1109/ICIIECS.2015.7192863.

[14] 3rd Generation Partnership Project (3GPP), "NR; Physical channels and modulation (Release 19)," *3GPP TS 38.211 V19.1.0*, Sep. 2025.

[15] H. L. Van Trees, Optimum Array Processing: Part IV of Detection, Estimation and Modulation Theory. New York: John Wiley & Sons, 2002.