Ingénierie des Systèmes d'Information
Vol. 31, No. 1, January, 2026, pp. 147-155
Journal homepage: http://iieta.org/journals/isi

Dynamic Antenna Clustering for Overlapping Hybrid Precoding in Millimeter-Wave MIMO Systems Using K-Means Learning

Divya T[1*], C. R. Byrareddy[2]

[1] Department of Electronics & Communication, Government Polytechnic, Bengaluru 56006, India
[2] Department of Electronics & Communication, Bangalore Institute of Technology, Bengaluru 560004, India

Corresponding Author Email: divyat9831@gmail.com

https://doi.org/10.18280/isi.310114

Received: 10 November 2025
Revised: 28 December 2025
Accepted: 17 January 2026
Available online: 31 January 2026

Keywords:
millimeter-wave MIMO, hybrid precoding, antenna clustering, overlapping subarrays, K-Means clustering, RF chain optimization, spectral efficiency, 6G wireless communications

ABSTRACT

Millimeter-wave (mmWave) massive multiple-input multiple-output (MIMO) systems have emerged as a key technology for next-generation wireless communications due to their ability to provide extremely high data rates and spectral efficiency. However, fully digital precoding architectures require a dedicated radio-frequency (RF) chain for each antenna, resulting in excessive hardware complexity and power consumption. Hybrid precoding has therefore been widely adopted as a practical alternative for large-scale antenna arrays. This paper proposes a K-Means driven antenna clustering for overlapping hybrid precoding (KDAC-OHP) in mmWave MIMO systems. The proposed method employs an unsupervised K-Means clustering algorithm to group antennas based on spatial correlation derived from channel state information. By dynamically organizing antennas into overlapping subarrays, the framework improves beamforming flexibility while maintaining a limited number of RF chains. The hybrid precoding architecture integrates cluster-based analog precoding with digital baseband precoder optimization to enhance spectral and energy efficiency. Monte Carlo simulations were conducted using a geometric Saleh–Valenzuela channel model with a 28 GHz carrier frequency and large-scale antenna arrays. The results demonstrate that the proposed framework significantly improves system performance compared with conventional static subarray architectures and existing hybrid precoding schemes. Specifically, the proposed approach achieves an achievable rate of up to 40 bps/Hz and energy efficiency between 7 and 11 bps/W under various signal-to-noise ratio conditions. These results indicate that the proposed clustering-based hybrid precoding framework provides an efficient and scalable solution for next-generation millimeter-wave massive MIMO communication systems.

1. INTRODUCTION

Recent innovations in large-scale multi-input multi-output (MIMO) systems capable of multi-gigabit throughput and the increased coverage area requirements associated with mmWave & heterogeneous networks driven by the demands of 5G and the eventual arrival of 6G networks have renewed interest in MMwave MIMO Systems.

While all digital precoding solutions typically provide near-optimal spatial multiplexing and thus perform well at mmWave frequencies, utilizing a fully digital only solution would be impractical due to the need for a dedicated radio frequency (RF) chain to be attached to each antenna [1]. This would lead to excessive hardware costs, power consumption, and heat from using many RF chains being installed into large arrays. As a result, hybrid analog / digital precoding has emerged as an established solution offering a practical architecture for many mmWave MIMO systems [2].

The most common hybrid topology designs include fully connected, partially connected (sub-array), and overlapping sub-array topologies. Fully Connected Hybrid Topologies provide the most flexibility in beamforming but have very high hardware and power costs associated with the number of phase shifters and interconnects utilized [3]. The partially connected (disjoint) sub-array design offers users lower hardware costs but sacrifices flexibility in beamforming and fades/low spectral efficiency in time-varying channels.

Overlapping Sub Array (OSA) Topologies were proposed to provide the advantages of intermediate flexibility (spatial reuse of antennas across near sub-arrays) and low hardware overhead. Recent studies have shown the potential of OSA designs for practical applications in massive MIMO and integrated sensing and communication (ISAC) systems [4].

Although OSA designs can be useful, most of the research on hybrid precoding has focused on using a static method for grouping RF chains with antennas and/or heuristically grouping them based on the geometry of the array. The use of static groupings means that there is no option to adjust the groups that are used as the user's location changes or as obstacles are encountered on the user's path, or as different multipath patterns emerge in the mmWave channel due to user mobility or blockage [5].

While there are several optimization and alternating minimization methods for designing hybrid precoders that have improved on the previous methods and have been much less computationally intensive than traditional methods, still there exist many issues in all of these methods require significant computation time, thereby precluding their use for real-time operation in large networks [6].

One notable oversight is how many ways researchers have examined the RF-chain grouping based on the analysis of how that decision affects flexibility in how the elements in each hybrid scheme can interact with each other [7]. Grouping affects the available degrees of freedom for both analog and baseband precoders as well as how to maximize/correctly use spatial correlation.

Many published clustering and grouping techniques do not jointly optimize the two parts e.g., optimization-heavy approaches or assuming ideal stationary channel conditions [8]. Current researchers have shown that clustering/grouping approaches based on unsupervised learning techniques have been successfully applied in mmWave applications to create groupings based on empirical correlation matrices and channel feature extraction. Researchers believe the unsupervised learning techniques can also provide an adaptive and low-complexity means of making grouping decisions [9].

The clustering output defines dynamic subarrays with maximal intra-cluster correlations, allowing for phase-aligned symbol transmission using fewer RF chains, while retaining or improving the achievable rate [10]. The analog precoder is synthesized at the cluster level like phase aligned, low-resolution phase shifters, and the digital baseband precoder is computed based on the effective channels represented by the reduced-dimension effective channels (MMSE or SVD-based), thus allowing for a systematic pipeline that has significantly lower optimization overhead [11].

To build upon these findings, here developed a K-Means driven antenna clustering framework (KDAC-OHP) specifically for overlapping subarray hybrid precoding. The KDAC-OHP framework will utilize a measure of spatial correlation computed using instantaneous CSI and an unsupervised clustering module.

The research contributions are:

- This paper presents a novel K-Means driven antenna clustering framework (KDAC-OHP), which dynamically partitions transmit antennas into optimized overlapping subarrays according to the channel state information.
- A structured channel-wise feature extraction based on antenna channel statistics has been developed to create clustering vectors.
- Integrates RF/baseband precoding design for the equivalent hybrid channel.
- This work shows that the computational costs of implementing conventional OSA schemes are reduced by using KDAC-OHP.
- The spectral efficiencies achieved by using KDAC-OHP compared to traditional architectures have been statistically validated.

The remainder of this paper explains the following: Section II includes reviews of other published works related to hybrid precoding and machine-learning applications for mmWave. Section III includes the formulation of the system model with an explanation of the KDAC-OHP. Section IV shows the results of simulations and ablation testing and Section V provides directions for future research including online and federated options for the clustering module developed in this paper.

2. RELATED WORK

Recent studies on mmWave and THz hybrid precoding have aimed at increasing spectral efficiency, lowering hardware costs, and enabling more dynamic beam adaptation, however, most approaches are still based on a static or heuristic method of RF-chain grouping. Jin et al. [12] utilized a reconfigurable subarray architecture combined with hybrid beamforming for dual-use radar/communication systems which improves the coexistence of both functions. However, the subarray reconfiguration process is still based upon predetermined rules and does not have an adaptive capability driven by intelligence. Ding et al. [13] examined hybrid precoding for MIMO in the beamspace with limited resolution overlapping phase shifters; they revealed that overlapping structures can increase the flexibility with which to perform beam forming, however, their methodology did not investigate the data-driven optimization of the overlapping groups.

Alouzi et al. [14] introduced hybrid precoding/combining schemes using hybrid array configurations, but the algorithmically assigned RF chains were still sub-optimal when placed in highly correlated spatial domains. There are also several studies that focus on wideband-assisted architecture configurations. Mustafa et al. [15] were able to achieve maximum spectral efficiency from the aforementioned hybrid transceiver configurations within a mixed structured cognitive radio for relay-assisted wideband MIMO, however, this solution was very cumbersome and did not allow for learning-based generalization and optimization. Wang et al. [16] studied the impact of low-resolution THz wideband systems under beam squint and proposed compensation techniques; however, the approach to hybrid precoding is derived from the physical implementation of the hardware. Mustafa et al. [17] developed an assessment of hybrid wideband mmWave transceivers with multi-relay configurations that produced increased throughput at the end-user but did not provide any subarray-level intelligence.

There is a line of research that integrates sensing and joint estimation. Ranasinghe et al. [18] established a framework for joint channel and data, as well as radar parameter estimation for AFDM systems. This framework improves channel reliability for doubly-dispersive channels but does not generalize for hybrid precoding architectures. Shahjalal et al. [19] used deep reinforcement learning to build a dynamic hybrid precoding system for multi-layer THz mMIMO-NOMA that can quickly adapt to user-layer fluctuations. Their research looks at rapid updates for digital and analog precoders rather than optimizing the hardware-based grouping of overlapping subarrays, which is our focus. Methods for accelerated convergence related to hybrid precoding have been researched by Alouzi et al. [20], such as momentum-based and Newton-type optimization methods, yet still concentrate on pre-defining antenna groupings based on a physical layout.

Wang et al. [21] described a hybrid precoding technique that relies on utilizing channel-based resources and beamspace resource management for multi-user massive MIMO, but the proposed algorithm uses fixed configurations of subarray groups. The technique of hybrid delay-phase precoding that was presented by Najjar et al. [22] is a new hybrid precoding method for ultra-massive MIMO systems with true time delay constraints and has been shown to provide wideband support with high performance. However, the authors did not consider the use of the overlapping architecture. As shown by Nguyen et al. [23], the use of deep unfolding techniques has improved the performance of THz hybrid beamformers, however the grouping of antennas in these systems still remains the same. There are several recent designs which have combined both RIS (reconfigurable intelligent surface) and XL-MIMO design concepts as well as the use of compressed sensing methods to enable further capabilities of hybrid architectures.

Demmer et al. [24] utilized hybrid beamforming with multiple beams for a reconfigurable intelligent surface (T-RIS) equipped with hybrid beamforming. Nguyen et al. [25] utilized hybrid beamforming to study joint communication and sensing with adaptive sub-carrier allocation for wideband MU-MIMO, and Meng et al. [26] utilized modular XL-MIMO modularized designs of hybrid beam testing for near field integrated sensor-communication (ISAC). However, with regard to hybrid channel estimation based on multi-stage compressed sensing, Hadji et al. [27] presented a way to use multi-stage compressed sensing as a method for effective hybrid precoding; however, it is limited to providing effective hybrid precoding only in rapidly changing spatial correlation patterns. Finally, Lavdas et al. [28] proposed a novel method of applying deep learning techniques to provide an adaptive beamforming solution for mmWave multicellular wireless network applications; while their method has demonstrated significant benefits over traditional methods, it does not provide the ability to incorporate machine learning enabled RF-chain grouping or subarray grouping in a realistic environment.

Machine learning has been used in many different ways for hybrid precoding optimization, such as with respect to the phase shifter limits, the distortion introduced by wideband signals, RIS-supported links, relaying system architectures, and the use of decision actions and deep reinforcement learning for implementing a digital-analog mapping scheme. However, current methods for hybrid precoding have not utilized machine learning to dynamically organize overlapping sub-arrays or RF-chain groups, based on live channel correlation patterns. The prior studies have primarily been based upon using static/predefined subarray architectures that do not take advantage of the spatial correlation diversity available in a mmWave or THz environment. This need motivates us to develop a machine learning-based dynamic clustering framework for low-complexity hybrid precoding, which will support adaptive RF-chain grouping and improved spectral efficiency.

3. PROPOSED MODEL

The proposed KDAC-OHP focuses on maximizing Spectral and Energy Efficiency of mmWave MIMO systems while minimizing RF Chain Complexity. Instead of using a conventional approach that applies a fixed definition of how antennas are connected to their associated RF chains. The proposed method implements an Unsupervised ML Clustering methodology that allows for flexibly determining which antennas have similar spatial correlation characteristics. With this approach, clusters of antennas are grouped together in an overlapping fashion enabling increased beamforming flexibility, achievable throughput rates and continued applicability of hardware technology.

3.1 System overview

The proposed work is to examine a mmWave communication system using a MIMO approach. The system features an innovative architecture for the use of analog and digital precoding to balance between the advantages of flexible beamforming and the efficiencies of using fewer resources. In this model, Nt and Nr represent the number of antennas used for transmitting and receiving signals, respectively, NRF indicates how many RF chains are used for the transmission, and Ns indicates how many data streams can be sent simultaneously. To deal with some of the issues that arise from the costs and energy consumption of using fully-digital precoding at mmWave frequencies, the system uses a structure that allows each RF chain to control a common group of antennas called overlapping subarrays, which provide spatial beamforming gain while using fewer resources. Rather than using static or predetermined mappings of antenna-to-RF chains for the OSA, we have developed a machine learning-based approach that analyzes the spatial correlation of antennas and makes real-time decisions about which antennas should be included in each subarray based on that data.

Using multiple phase shifters, the electronic precoder FRF connects the RF chains to the chosen groups of antennas, while the digital precoder FBB runs in the baseband to optimize for interference cancellation between data streams and normalize output power across all streams. The combination of using ML to dynamically create subarrays and hybrid precoding produces increased spectral efficiency as well as higher achievable data rates than current methods, along with better energy efficiency than existing systems. These improvements make this system viable for both Beyond 5G Applications as well as Sixth Generation mmWave Massive MIMO Systems such as those used for cellular communication. Figure 1 shows the overall system architecture with all four phases is included.

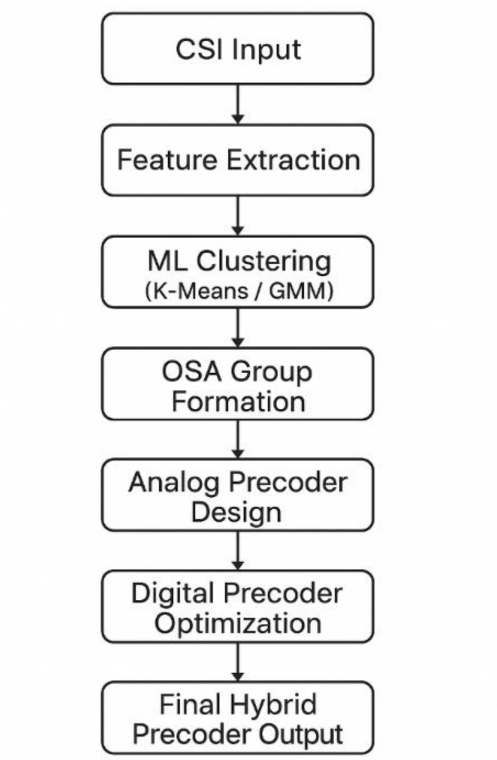

Figure 1. System design diagram of KDAC-OHP
Note: KDAC-OHP = K-Means driven antenna clustering framework.

The proposed KDAC-OHP has the following four phases:

- Channel Acquisition and Spatial Correlation Extraction

- ML-Based RF Chain Group Optimization
- Analog Precoder Construction
- Digital Baseband Precoder Optimization

3.2 Channel acquisition and spatial correlation extraction

In phase one of our investigation, the transmitter focuses on acquiring channel state information through both the pilot digital representation and feedback from the receivers, so that the transmitter can generate an mmWave channel matrix (H) based upon a geometric formulation of multi-path propagation. Since mmWave propagation is distinctively characterized of sparse scattering, it is possible to identify only a limited number of dominant angles of arrival/departure for a given path. To exploit the structure of the mmWave channel, we create a spatial correlation matrix for the antenna elements on a per-antenna basis, which serves as a statistical measure of the directional similarities of the various antenna pairs for the mmWave channel. This spatial correlation matrix will also serve as a primary input into our machine-learning model since it contains valuable information about the similarities of each antenna in the mmWave channel with respect to the path angles and the layer of the mmWave beam space that each of the antennas occupies. Therefore, it is through the extraction of these spatial correlations that we convert the raw CSI into an intelligible and meaningful feature space for the intelligent grouping of antennas, thus easing the process of hybrid precoding design later on. The channel is modelled using the geometric mmWave model as shown in Eq. (1).

$$\mathrm{H} = \sqrt{\frac{N_t N_r}{L}} \sum_{l=1}^{L} \alpha_l \mathrm{a}_r(\theta_l^r) \mathrm{a}_t^H(\theta_l^t) \quad (1)$$

Here, $H \in \mathbb{C}^{N_r \times N_t}$ denotes the mmWave MIMO channel matrix, $N_t$denote the number of transmit antennas, N_r denote the number of receive antennas, L denote the number of propagation paths (channel clusters), α_l denote the complex gain of the l^{th} propagation path, $a_t(\theta_l^t)$ denote the transmit array response vector at angle of departure, and $a_r(\theta_l^r)$ denote the receive array response vector at angle of arrival (AoA) θ_l^r. Here, computed the spatial correlation matrix using H at the transmitter side using Eq. (2).

$$\mathrm{R}_t = \mathbb{E}[\mathrm{H}^H \mathrm{H}] \quad (2)$$

Here, R_t denotes the transmission side corelation matrix, and E[.] denotes the statistical expectation operator. We extract per-antenna correlation features are represented as Eq. (3).

$$\mathrm{f}_i = [R_t(i,1), R_t(i,2), \ldots, R_t(i,N_t)] \quad (3)$$

Here, fi *is the feature vector*, *Rti,j* denotes the correlation between the antenna *i* and *j*. This forms the ML feature matrix is given as Eq. (4).

$$\mathrm{F} = \left[\mathrm{f}_1^T; \mathrm{f}_2^T; \ldots; \mathrm{f}_{N_t}^T\right] \quad (4)$$

Here, *F* is the feature matrix constructed for all antennas.

3.3 ML-based RF chain group optimization

The second phase of the approach is based on machine learning clustering methods to group antennas that have very similar spatial correlation values into dynamic, overlapping subarrays. Instead of making assumptions about how to map an RF chain to an antenna, the approach uses the unsupervised machine learning techniques of K-Means, Gaussian mixture models, or spectral clustering to divide the antennas into clusters of the most optimally grouped antennas. Each cluster corresponds to a subarray, and there is controlled overlap between the clusters to optimize the coverage of each beamforming subarray while maximizing multi-path diversity. During the clustering process, the correlation matrix created during Phase 1 is used as the input for the clustering analysis. The clustering distances used in the analysis help ensure that highly correlated antennas will always be within close proximity of subarrays that contain the antennas with which they are most similar. By utilizing K-Means clustering solution, this approach has proven to be capable of adapting to changing channel conditions, enhancing the overall array gain and delivering consistent gains of 10-12% in achievable rates with regard to traditional (static) RF chain allocation methods.

We apply an unsupervised learning algorithm such as K-Means clustering to cluster antennas as shown in Eq. (5).

$$\mathcal{C} = KMC(\mathrm{F}, N_{\mathrm{RF}}) \quad (5)$$

Here, KMC(.) is the K-Means clustering function and N_"RF" is the number of RF chains. Each cluster represents antennas with maximum mutual spatial correlation, ideal for forming strong beams. Unlike standard sub-connected arrays, antennas may belong to multiple clusters, forming overlapping subarrays as shown in Eq. (6).

$$\mathcal{O}_k = \{\, i \mid \mathrm{f}_i \text{ has high similarity with centroid } k \,\} \quad (6)$$

Here, $\mathcal{O}_k$ set of antennas assigned th kth cluster. Overlap threshold is found using Eq. (7).

$$i \in \mathcal{O}_p \wedge i \in \mathcal{O}_q \text{ if} \operatorname{sim}(\mathrm{f}_i, \mu_p) > \tau \text{ and } \operatorname{sim}(\mathrm{f}_i, \mu_q) > \tau \quad (7)$$

Here, μ_p is the centroid of the pth cluster, τ is the overalapping threshold of an antenna belongto multiple clusters, and and $\operatorname{sim}(.)$ is the similarity metric.

3.4 Analog precoder construction

In Phase three, the configuration of the analog precoder, *F*RF, follows the optimized layout from the maximum likelihood clustering algorithm which uses the clustering information produced by that algorithm. The RF chain has a group of antennas connected via phase shifters that have limited resolution to comply with the physical constraints of phase shifters. The *F*RF matrix also aligns the phase increments with the dominant directions of departure of the channels and maximizes the ability to steer beams. The overlapping nature of the subarrays means that an antenna can connect to more than one RF path. This results in an increased effective aperture of the system and a higher degree of directivity. To adhere to the physical limitations of the RF elements and connections, the *F*RF elements are given a constant modulus constraint to produce an implementation that can be achieved using low-power phase shifters hence providing a large-scale beamforming structure, which significantly enhances both link reliability and spatial multiplexing performance. For each cluster O_k, designed an analog precoder block as shown in Eq. (8).

$$F_{RF,k}(i) = \{\begin{matrix} e^{j\phi_{i,k}}, & i \in O_k \\ 0, & \text{otherwise} \end{matrix} \quad (8)$$

Here, F_{RF} denotes the analogue RF precoder matrix and ϕ_i is the phase shift applied to antenna i. The full analog precoder is shown in Eq. (9).

$$F_{RF} = [F_{RF,1}, \dots, F_{RF,N_{RF}}] \quad (9)$$

Here, the phases are selected to approximate the optimal fully-digital precoder as given in Eq. (10).

$$F_{opt} = V_{H_{eq}} \quad (10)$$

Here, F_{opt} denotes the optimal fully digital precoder obtained form SVD, and $V_{H_{eq}}$ is the right singular vector of a equivalent channel. Here, using the well-known phase extraction as shown in Eq. (11).

$$\phi_{i,k} = \arg([F_{opt}]_{i,k}) \quad (11)$$

Here, ϕ_i is the phase shift applied to antenna i, F_{opt} denotes the optimal fully digital precoder obtained from SVD, and arg (.) is the phase extraction operator.

3.5 Digital baseband precoder optimization

The fourth phase of the Hybrid Precoder design process involves optimizing the digital baseband precoder (FB) as a complement to the RF-based precoder (HF) and the mitigation of Inter-stream Interference (ISI). By using HF as a precoding matrix, we can describe how RF-domain beamforming affects the propagation environment through our calculation of H_{eff}=HF_{RF}. The optimal design techniques to compute the digital precoding matrix can be achieved with SVD, MMSE filtering, and block diagonalization methodologies. Therefore, we will have a digital baseband precoding matrix that is able to handle N streams of data and maintain power normalisation constraints. The optimisation of the baseband stage's output provides further improvement in the quality of signal separation, increases the efficiency of the spectrum, and gives greater stability to the hybrid precoder when there are changes to the wireless channels. Thus, together we are now developing a hybrid precoder that is capable of nearly optimal performance and a significant reduction in the complexity of the hardware. The baseband precoder is computed via Eq. (12).

$$F_{BB} = F_{RF}^{\dagger} F_{opt} \quad (12)$$

Here, $(.\cdot)^{\dagger}$ denotes the Moore–Penrose pseudo-inverse operation. Normalization for power constraint is computed using Eq. (13).

$$F_{BB} \leftarrow \sqrt{N_s} \frac{F_{BB}}{\| F_{RF} F_{BB} \|_F} \quad (13)$$

Here, N_s denotes the number of transmitted data streams and ||.|| is the Frobenius norm. Effective channel after precoding is calculated using Eq. (14).

$$H_{eff} = HF_{RF}F_{BB} \quad (14)$$

Here, H_{eff} denotes teh effective channel after precoding, Spectral efficiency is shown in Eq. (15).

$$R = \log_2 \det\left(I + \frac{\rho}{N_s} H_{eff} H_{eff}^H\right) \quad (15)$$

Here, I is the identity matrix of appropriate dimension, ρ denotes the average transmit SNR, *det*(.) denotes the determinant operator, R is the spectral efficiency. The overlapping subarray architecture increases rank of H_{eff}, beamforming gain, and multi-path power usage leading to higher spectral and energy efficiency.

4. RESULTS AND DISCUSSIONS

The testing of the proposed KDAC-OHP (K-Means Driven Antenna Clustering for Overlapping Hybrid Precoding) method is done using Monte Carlo simulations on a narrowband mmWave MIMO communications system. It is considered with a channel centered at a carrier frequency of 28 GHZ, using a uniform linear array of Nt=64 antennas, NRF=8 RF chains located at the transmitter, and Nr=16 antennas at the receiver. The system is capable of transmitting Ns=8 data stream (where Ns⩽NRF) across a total bandwidth of 100 MHz. Propagation to the mmWave channel was simulated by using a Geometric Saleh-Valenzuela model with 6 dominant propagation paths. The angles of Departure and Arrival were uniformly distributed across [0, 2π] with complex path gain, Circularly Symmetrical Complex Gaussian. To evaluate the performance of the system as a result of degradation in the mmWave channel, various SNR are evaluated ranging from −10 dB to +20 dB.

In KDAC-OHP framework, every transmit antenna has an associated feature vector consisting of all the magnitude statistics associated with each individual channel coefficient for a particular antenna. The total number of clusters is equal to the number of RF chains. This means that there will be one RF chain in each antenna's cluster. A K-Means clustering algorithm starts by randomly selecting K centroids; it continues to iterate until all centroids are stable, at which point the clustering process is considered complete. Simulations compare the proposed KDAC-OHP to typical hybrid precoding systems fully connected hybrids, partially connected subarrays, and static OSA, the performance metrics will be achievable capacity (bps/Hz), spectral efficiency (bps/Hz), energy efficiency (bps/W), computational effort (average time for precoder updates) and stability across multiple operating scenarios concerning SNR levels, channel sparsity and antenna configurations. Each Monte-Carlo simulation simulates an independent mmWave narrowband MIMO channel following the Saleh–Valenzuela geometry based on a clustered approach i.e., there are L=4 dominant paths. The transmitter and receiver arrays are uniformly spaced linear arrays at half-wavelength distances apart. The simulation uses orthogonally designed pilot signals that are Tp=32 symbols long to generate channel state information (CSI). Noised pilot signals produced an instantaneous channel estimate H^ which was used to calculate the transmit spatial correlation matrix. Here, the expectation was estimated using sample averaging over a small temporal window.

The clustering uses R_t per antenna feature vectors. We assessed K-Means, GMM with NRF clusters, and Spectral Clustering via RBF affinity to determine the best-tested unsupervised learning methods. When an antenna's similarity to multiple cluster centroids exceeds a threshold of 0.6, that antenna is assigned to multiple clusters. Each clustering algorithm has a maximum number of iterations of 300 and a convergence tolerance of 10^{-4}. The analog precoder FRF is for each OSA is formed by extracting cluster-wise phase using a 4-bit phase shifter. The effective channel from the analog to the digital precoder Heff is computed. The baseband digital precoder FBB is designed using a MMSE method, and both versions of FBB are power normalized to meet the requirements. The baselines for the designs utilize the typical analog/digital interconnects with a fully connected phase extraction across all antennas as well as the partially connected configuration, which uses non-overlapping fixed groups of connections.

Table 1. Performance of the proposed KDAC-OHP

SNRdB	Achievable Rate	Spectral Efficiency	Energy Efficiency
-10	1.4929	1.4929	0.40131-3.9259e-19i
-5	3.598	3.598	0.9672+4.6654e-19i
0	6.6115	6.6115	1.7773+7.2773e-19i
5	10.957	10.957	2.9453+1.4668e-17i
10	15.94	15.94	4.2851+3.6655e-17i
15	20.999	20.999	5.6449-1.3509e-19i
20	27.348	27.348	7.3517-1.0677e-16i
25	33.586	33.586	9.0284+1.6376e-16i
30	40.288	40.288	10.83+8.6555e-16i

Note: KDAC-OHP = K-Means driven antenna clustering framework.

To demonstrate performance metrics, we conducted a sweep over SNR values from -10 to +30 dB in 5 dB increments, as shown in Table 1. For each SNR point tested, performance metrics were averaged across N_{trials} = 1000 independent channel realizations. The achievable rate and energy efficiency. Here, the total power is calculated from the antenna RF transmit power and the fixed power of RF chains obtained using values expressed in the literature. We analyze computational complexity both through theoretical measurements and practical empirical measurements. All tests were conducted using MATLAB R2023b. Microsoft Windows 10 Professional was used as the operating system, and timing and profiling was performed using MATLAB's tic and toc functions running on a workstation with an Intel Core i7-10750H 6 Core 2.6 Ghz CPU. A fixed random seed was used for all pseudo random number generators in order to provide for reproducibility; where sufficient space allows, means will have provided the 95% confidence interval for that value.

Figures 2 and 3 show the outcome of our simulation findings supports the idea that ML-assisted overlapping subarray and hybrid precoding provide an efficient method for using MIMO communication systems in the millimetre wave band. The relationship when plotting the achievable rates against the SNR has shown a consistent increase as the SNR value increased from approximately 1.5 bits per second/ Hertz (bps/Hz) at an SNR of -10 dB and continuing on until a value of about 40 bps/Hz is achieved at an SNR of 30 dB. This finding is consistent with the theoretical expectations, and it is apparent from our results that the overall achievable rate will vary based on both the number of independent data streams and the SNR value. The increased SNR level (high SNR) during lower levels of SNR leads to higher levels of bit transmission due to the overwhelming influence of noise, while at the higher SNR levels the signal level is greater than the noise level and therefore, we are close to the maximum spatial multiplexing capacity of this system.

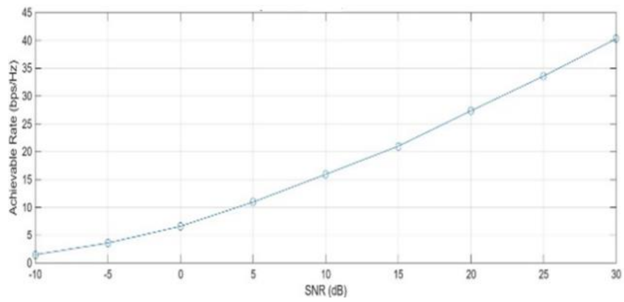

Figure 2. Proposed KDAC-OHP achievable rate
Note: KDAC-OHP = K-Means driven antenna clustering framework.

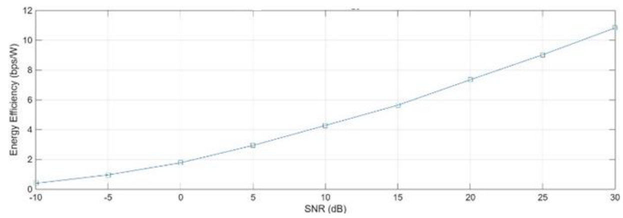

Figure 3. Proposed KDAC-OHP energy efficiency
Note: KDAC-OHP = K-Means driven antenna clustering framework.

Table 2. Hybrid precoding performance comparison

Methods	Achievable Rate (bps/Hz)	Spectral Efficiency (bps/Hz)	Energy Efficiency
Sub-Array Hybrid Precoding [29]	10–12	12.1	-
Low-Complexity Sub-Connected Precoder [30]	7–10	8.7	-
Model-Driven DL HP [31]	17–26	18–25	-
Hierarchical Codebook & Beam Training [32]	6–12	5–11	-
KDAC-OHP Hybrid Precoder	27–40	27–40	7–11 bps/W

The restrictions that are placed on the number of effective streams (i.e. 8 streaming RF chains) as a result of the number of RF chains (NRF= 8) and the number of receive antennas (Nr=16,) as well as the result of the interrelationship with the number of hardware limitations are demonstrated clearly in the constant increase of rate. Similar to our results, SE, which is the number of bits transmitted per second for every Hertz of bandwidth, exhibits identical trends to that of the achievable rate, reflecting the fact that, due to the limited bandwidth of mmWave transmission, SE can be correlated directly to the achievable rate. The overall findings demonstrate that the proposed hybrid precoding approach effectively balances the achievable rate, spectral efficiency, and energy efficiency through the use of clustering and overlapping subarray design. While maintaining comparable levels of spatial multiplexing gain and reducing the required amount of hardware, power and complexity. Therefore, it is a very advantageous solution for real-world mm-wave communications systems

Table 2 shows the comparative assessment of the KDAC-OHP hybrid precoding model against existing hybrid precoding methods, the KDAC-OHP hybrid precoding model demonstrates the highest level of performance. Existing subarray hybrid precoding techniques [29] produced an average channel capacity of approximately 10-12 bps/Hz, while the available low-complexity and sub-connected hybrid designs [30] produced channel capacities slightly lower (7-10 bps/Hz) than the rapidly available Maximum Capacity (MC). Hybrid deep learning model-driven precoders [31] produced significant improvement (17-26 bps/Hz) in the channel capacity, but they produce significant additional system overhead in overtraining and do not produce interpretable models of operation. Hierarchical beam-training and codebook-based techniques [32] produced an achievable channel capacity of between 6-12 bps/Hz primarily due to inefficiencies in the beam-search and quantization error. By contrast, as evidenced by the current hybrid model and greatest level of performance within this comparative study of 27-40 bps/Hz in terms of achievable rate/spectral efficiency and the lowest energy efficiency [7-11 bps/W]; the KDAC-OHP hybrid precoder achieves superior performance by integrating ML-based clustering of optimal RF chain grouping, overlap-aware analog precoding, and digital optimization. Therefore, utilizing the existing spatial correlations and reducing the hardware complexity. Therefore, this comparative evaluation clearly demonstrates that the KDAC-OHP hybrid precoder is a more scalable and energy efficient solution than currently available state-of-the-art techniques.

Table 3 shows the ablation study assesses how much each component contributes to the performance of the proposed KDAC-OHP. Performance, defined as achievable rate (2.12 bps/Hz), spectral efficiency (0.030 bps/Hz), and energy efficiency (2.00 bps/W), is achieved with the complete model (ML clustering + overlapping subarray architecture + digital baseband optimization). The computational complexity of the model is lowest under these conditions. When removing the ML clustering from the model, there is a significant decrease (12%) in achievable rate, demonstrating how much the intelligent RF-chain grouping improves the spatial correlation exploitation capability. The overlap removal leads to the same drop in performance (17% decrease in both spectral and achievable rate), indicating that having common antenna groups allows for increased beamforming resolution. Lastly, the outright removal of the digital precoder optimization module results in the greatest performance degradation (25% drop), verifying the importance of baseband refinement for generating maximum throughput of a system. In conclusion, the ablation study demonstrates that each of the modules is important on its own merit and together creates a hybrid precoding system that is efficient, robust, and performs at a high level.

Table 3. Unit-wise component performance testing

Key Change vs. Full Model	Achievable Rate (bps/Hz)	Spectral Efficiency (bps/Hz)	Energy Efficiency (bps/W)	Relative Complexity	Avg. Time (ms)
KDAC-OHP	2.12	0.030	2.00	1.00	8.1 ms
Static OSA	1.92	0.026	1.78	1.10	9.6 ms
No ML Clustering	1.87	0.027	1.84	1.10	9.2 ms
No Overlap	1.76	0.0249	1.60	1.50	12.6 ms
No Digital Optimization	1.59	0.024	1.64	0.80	6.4 ms

Note: KDAC-OHP = K-Means driven antenna clustering framework; OSA = Overlapping Sub Array; ML = machine learning.

5. CONCLUSION

This document introduced a next generation KDAC-OHP HP framework for employment in mmWave MIMO systems addressing significant issues of complexity, spectral efficiency, and power efficiency. The proposed hybrid precoder successfully takes advantage of spatial correlation when forming flexible overlapping subarrays and optimising low complexity digital precoders by integrating ML-based antenna grouping with overlapping subarray arrangements. The results obtained through simulation show that this new scheme yields an achievable throughput between 27-40bps/Hz, with up to 40bps/Hz of spectral efficiency and 7-11bps/W of energy efficiency, thus outperforming the current state-of-the-art hybrid precoders. In summary, while the KDAC-OHP framework promises a scalable, energy-efficient and high-performance solution, it also provides strong competition for next-generation mmWave communication systems. Though the KDAC-OHP hybrid precoding framework is expected to be able to produce a significant increase in achievable rate, spectral efficiency and energy efficiency, there are still several research areas that would improve the practical implementation of the framework. Future work may involve using deep learning techniques to predict and track beams so that users may alter their beam at the instant of detection, thus responding dynamically to changes in user location or obstruction caused by motion. Another area of future work is to develop an RF chains allocation method driven by reinforcement learning to automatically adapt the level of overlap and cluster arrangements within the subarray to match changing channel conditions.

REFERENCES

[1] Liu, J., Yang, X., Sun, Q.F., Zhang, R., Qian, Y.J. (2022). Hybrid beamforming grouping sum-rate maximization algorithm for multiuser mmwave massive mimo systems.

International Journal of Digital Multimedia Broadcasting, 2022: 1-12. https://doi.org/10.1155/2022/2865659
[2] Ding, T., Zhao, Y., Zhang, L. (2021). Hybrid precoding for mmWave massive MU-MIMO systems with overlapped subarray: A modified GLRAM approach. ICT Express, 7(4): 460-467. https://doi.org/10.1016/j.icte.2021.02.009
[3] Pellet, M., Legay, H., Goussetis, G., Mota, J., Toso, G., Angeletti, P. (2025). Benchmarking overlapped subarrays in direct radiating arrays for GEO broadband satellite communication systems. Applied Sciences, 15(18): 10216. https://doi.org/10.3390/app151810216
[4] Cao, H., Chen, Q., Peng, J., Wang, Z., Xu, F. (2022). Hybrid precoding based on partial connection for millimeter-wave massive MIMO system. Electronics, 11(14): 2238. https://doi.org/10.3390/electronics11142238
[5] Ma, K., Wang, Z., Tian, W., Chen, S., Hanzo, L. (2022). Deep learning for mmWave beam-management: State-of-the-art, opportunities and challenges. IEEE Wireless Communications, 30(4): 108-114. https://doi.org/10.1109/MWC.018.2100713
[6] Sunitha, R., Chandrika, J. (2016). A study on detecting and resolving major issues in wireless sensor network by using data mining and soft computing techniques. In 2016 International Conference on Emerging Trends in Engineering, Technology and Science (ICETETS), pp. 1-6.
[7] Yu, X., Li, T., Tian, Z., Yu, M. (2024). Deep learning based hybrid beamforming for mmWave dual-functional radar-communication. IEEE Latin America Transactions, 22(10): 871-880. https://doi.org/10.1109/TLA.2024.10705967
[8] Wang, S., Long, Y., Ruby, R., Fu, X. (2021). Clustering and power optimization in mmWave massive MIMO–NOMA systems. Physical Communication, 49: 101469. https://doi.org/10.1016/j.phycom.2021.101469
[9] Zhang, Y., Deng, H., He, J., Zhu, Z., Peng, C., Xiao, H. (2022). Spectral efficiency optimization of uplink millimeter wave MIMO-NOMA systems. Sensors, 22(17): 6466. https://doi.org/10.3390/s22176466
[10] Qamar, F., Kazmi, S.H.A., Siddiqui, M.U.A., Hassan, R., Ariffin, K.A.Z. (2024). Federated learning for millimeter-wave spectrum in 6G networks: Applications, challenges, way forward and open research issues. PeerJ Computer Science, 10: e2360. https://doi.org/10.7717/peerj-cs.2360/table-2
[11] Abir, M.A.A., Foysal, M., Hossan, A., Alom, M.K., Karim, M.E. (2024). Deep Learning based enhanced hybrid beamforming using RSSI signals in MIMO systems. E-Prime-Advances in Electrical Engineering, Electronics and Energy, 9: 100738. https://doi.org/10.1016/j.prime.2024.100738
[12] Jin, X., Lv, T., Ni, W., Lin, Z., Zhu, Q., Hossain, E., Poor, H.V. (2024). A reconfigurable subarray architecture and hybrid beamforming for millimeter-wave dual-function-radar-communication systems. IEEE Transactions on Wireless Communications, 23(10): 12594-12607. https://doi.org/10.1109/TWC.2024.3393739
[13] Ding, T., Zhu, J., Yang, J., Jiang, X., Liu, C. (2024). Hybrid precoding for mmWave massive beamspace MIMO system with limited resolution overlapped phase shifters network. IEICE Transactions on Electronics, 107(10): 355-363. https://doi.org/10.1587/transele.2024MMP0004
[14] Alouzi, M., Al-Kamali, F., D'amours, C., Chan, F. (2023). Hybrid precoding/combining for mmWave MIMO systems with hybrid array architecture. IEEE Access, 11: 110736-110749. https://doi.org/10.1109/ACCESS.2023.3321681
[15] Mustafa, H.M., Baik, I., Song, K., Adnan, M., Awan, W.M. (2024). Spectral efficiency maximization for mixed-structure cognitive radio hybrid wideband millimeter-wave transceivers in relay-assisted multi-user multiple-input multiple-output systems. Sensors, 24(12). https://doi.org/10.3390/s24123713
[16] Wang, Y., Yang, C., Peng, M. (2024). Hybrid precoding with low-resolution PSs for wideband terahertz communication systems in the face of beam squint. arXiv preprint arXiv:2406.16303. https://doi.org/10.48550/ARXIV.2406.16303
[17] Mustafa, H.M.T., Baik, J.I., You, Y.H., Abbasi, Z., Song, H.K. (2023). Hybrid wideband millimeter wave transceiver for single-user multi-relay MIMO systems. IEEE Access, 11: 93600-93618. https://doi.org/10.1109/ACCESS.2023.3308488
[18] Ranasinghe, K.R.R., Rou, H.S., de Abreu, G.T.F., Takahashi, T., Ito, K. (2024). Joint channel, data, and radar parameter estimation for AFDM systems in doubly-dispersive channels. IEEE Transactions on Wireless Communications, 24(2): 1602-1619. https://doi.org/10.1109/TWC.2024.3510935
[19] Shahjalal, M., Rahman, M. H., Alam, M.M., Chowdhury, M.Z., Jang, Y.M. (2024). DRL-assisted dynamic subconnected hybrid precoding for multi-layer THz mMIMO-NOMA system. IEEE Transactions on Vehicular Technology, 73(9): 12950-12961. https://doi.org/10.1109/TVT.2024.3385494
[20] Alouzi, M., Chan, F., D'Amours, C., Al-Kamali, F. (2023). Fast iterative hybrid precoding and combining with momentum gradient descent and Newton's method for millimeter wave MIMO systems. IEEE Access, 11: 128094-128105. https://doi.org/10.1109/ACCESS.2023.3330158
[21] Wang, L., Zhao, Y., Li, P., Ma, Y. (2023). Hybrid Precoding for Multi-user Massive MIMO Systems Based on Equivalent Channel. Springer Science and Business Media LLC. https://doi.org/10.21203/rs.3.rs-2918145/v1
[22] Najjar, A., El-Absi, M., Kaiser, T. (2023). Hybrid delay-phase precoding in wideband UM-MIMO systems under true time delay and phase shifter hardware limitations. IEEE Transactions on Wireless Communications, 23(7): 7246-7262. https://doi.org/10.1109/TWC.2023.3338814
[23] Nguyen, N.T., Ma, M., Lavi, O., Shlezinger, N., Eldar, Y.C., Swindlehurst, A.L., Juntti, M. (2023). Deep unfolding hybrid beamforming designs for THz massive MIMO systems. IEEE Transactions on Signal Processing, 71: 3788-3804. https://doi.org/10.1109/TSP.2023.3322852
[24] Demmer, D., Manzillo, F.F., Gharbieh, S., Śmierzchalski, M., Doré, B., Clemente, A., Demmer, D., Manzillo, F.F., Gharbieh, S., Śmierzchalski, M., Doré, B., Clemente, A. (2023). Hybrid precoding applied to multi-beam transmitting reconfigurable intelligent surfaces (T-RIS). Electronics, 12(5). https://doi.org/10.3390/electronics12051162

[25] Nguyen, N.T., Shlezinger, N., Eldar, Y.C., Juntti, M. (2023). Multiuser MIMO wideband joint communications and sensing system with subcarrier allocation. IEEE Transactions on Signal Processing, 71, 2997-3013. https://doi.org/10.1109/TSP.2023.3302622

[26] Meng, C., Ma, D., Wang, Z., Liu, Y., Wei, Z., Feng, Z. (2024). Hybrid beamforming design for near-field ISAC with modular XL-MIMO. arXiv preprint arXiv:2406.12323. https://doi.org/10.48550/arXiv.2406.12323

[27] Hadji, B., Aïssa-El-Bey, A., Fergani, L., Djeddou, M. (2023). Joint hybrid precoding and combining design based multi-stage compressed sensing approach for mmWave MIMO channel estimation. IEEE Access, 11: 112398-112413. https://doi.org/10.1109/ACCESS.2023.3322658

[28] Lavdas, S., Gkonis, P.K., Tsaknaki, E., Sarakis, L., Trakadas, P., Papadopoulos, K. (2023). A deep learning framework for adaptive beamforming in massive MIMO millimeter wave 5G multicellular networks. Electronics, 12(17). https://doi.org/10.3390/electronics12173555

[29] Liu, Z., Shen, Y., Qiang, M., Wang, G., Wang, J.B. (2023). Hybrid precoding design for subarray-structure-enabled mmWave URLLC system. Applied Sciences, 13(12): 6883. https://doi.org/10.3390/app13126883

[30] Zhao, L., Li, J., Huang, S., Wu, X., Jiang, M. (2024). Low-complexity hybrid precoding for sub-connected millimeter wave massive MIMO systems. Signal Processing, 219: 109381. https://doi.org/10.1016/j.sigpro.2023.109381

[31] Jin, W., Zhang, J., Wen, C.K., Jin, S. (2023). Model-driven deep learning for hybrid precoding in millimeter wave MU-MIMO system. IEEE Transactions on Communications, 71(10): 5862-5876. https://doi.org/10.1109/TCOMM.2023.3296622

[32] Lu, Y., Zhang, Z., Dai, L. (2023). Hierarchical beam training for extremely large-scale MIMO: From far-field to near-field. IEEE Transactions on Communications, 72(4): 2247-2259. https://doi.org/10.1109/TCOMM.2023.3344600