World Journal of Engineering Research and Technology

www.wjert.org

Impact Factor: 7.029 Coden USA: WJERA4

SPECTRUM AND ENERGY EFFICIENCY OPTIMIZATION FOR SWIPT-BASED HYBRID PRECODING-BASED MMWAVE MASSIVE MIMO-NOMA SYSTEMS

Jacob Ogwu[1], Kufre M. Udofia[2*], Akaninyene B. Obot[3]

[1,2,3]Electrical and Electronics Engineering Department, University of Uyo Nigeria.

Article Received on 21/12/2025 **Article Revised on 10/01/2026** **Article Accepted on 01/02/2026**

***Corresponding Author**

Kufre M. Udofia

Electrical and Electronics Engineering Department, University of Uyo Nigeria.

https://doi.org/10.5281/zenodo.18441063

How to cite this Article: Jacob Ogwu[1], Kufre M. Udofia[2*], Akaninyene B. Obot[3] (2026). Spectrum And Energy Efficiency Optimization For Swipt-Based Hybrid Precoding-Based Mmwave Massive Mimo-Noma Systems. World Journal of Engineering Research and Technology, 12(2), 45–66.

This work is licensed under Creative Commons Attribution 4.0 International license.

ABSTRACT

Non-orthogonal multiple access (NOMA) and millimeter-wave (mmWave) massive MIMO systems are pivotal for enhancing spectrum efficiency in 5G networks while integrating simultaneous wireless information and power transfer (SWIPT) offers a pathway to improve energy efficiency. This study addresses the joint optimization of transceiver design and power splitting in SWIPT-enabled hybrid precoding for mmWave massive MIMO-NOMA systems. Three main techniques were proposed: (1) a Dynamic Cluster-Head Selection Algorithm to assign beam-specific users, (2) a Compressed Sensing-based Committee Machine Design (CS-COMADE) to suppress inter-user interference and optimize antenna array gain, and (3) an iterative framework to resolve the non-convex problem of jointly optimizing power allocation and power-splitting factors. Compared to conventional SWIPT-based MIMO-OMA and existing hybrid precoding schemes, the proposed approach achieved superior spectrum and energy efficiency, nearing the performance of fully digital zero-forcing (ZF) precoding in energy efficiency. These advancements highlight the potential of SWIPT-NOMA integration to support sustainable, high-capacity 5G networks.

KEYWORDS: NOMA, SWIPT, MIMO, optimization, hybrid precoding.

1.0 INTRODUCTION

The fifth generation (5G) mobile communication has drawn considerable attention owing to its higher frequency, greater network capacity, and lower latency compared to existing sub-6 GHz radio technologies. This advancement is facilitated by exploring less-congested spectrum bands, such as millimeter-wave (mmWave) frequencies (30–300 GHz), which enable gigabit-per-second (Gbps) data rates in indoor and fixed outdoor systems (Shafi et al., 2021; Tataria et al., 2019; Huang et al., 2019). However, mmWave signals experience an order-of-magnitude increase in free-space path loss due to their ten-fold higher carrier frequency (Shafi et al., 2021; Yang et al., 2018).

Fortunately, the reduced wavelength (10 mm at 30 GHz to 1 mm at 300 GHz) allows densely packed antenna arrays, enabling massive multiple-input multiple-output (MIMO) systems to mitigate path loss through beamforming and achieve viable signal-to-noise ratios (SNRs).[5] Massive MIMO further enhances user throughput, spectral efficiency (SE), and energy efficiency (EE) through multi-stream precoding (Sohrabi and Yu, 2016; Alkhateeb et al., 2014). Conventional MIMO systems employ digital precoders to adjust signal magnitude and phase at baseband, requiring dedicated radio frequency (RF) hardware per antenna—a design infeasible for mmWave due to high costs and power consumption (Choi, 2017; Mishra and De, 2017; Ali et al., 2016). Consequently, hybrid precoding architectures, combining low-dimensional digital baseband precoders with analog RF precoders, have emerged as a cost-effective solution (Clerckx et al., 2019); Zhou et al., 2013). Hybrid architectures are categorized as fully connected (higher SE) or partially connected (higher EE), depending on RF chain-to-antenna mapping (Ding et al., 2020).

Non-orthogonal multiple access (NOMA) has been proposed to further improve SE in mmWave massive MIMO systems. In supporting multiple users per beam through superposition coding and successive interference cancellation (SIC), NOMA outperforms orthogonal multiple access (OMA) in spectral efficiency (Li et al., 2020).

Energy efficiency, a critical 5G performance target (projected to be 100× higher than 4G), has spurred research into simultaneous wireless information and power transfer (SWIPT) (Rusek et al., 2013). SWIPT employs power-splitting receivers to concurrently harvest energy and decode information from RF signals, prolonging device lifetimes and enhancing network sustainability. However, optimizing the trade-off between harvested energy and

achievable data rates in multiuser SWIPT systems remains a major challenge (Tang et al., 2021).

In this study, a technique for spectrum and energy efficiency optimization for SWIPT-based hybrid precoding-based mmWave massive MIMO-NOMA systems is presented. The remaining sections of this study are structured as follows: Section 2 presents a review of related works, while Section 3 details the research methodology used. The results and their discussion are provided in Section 4. Finally, Section 5 concludes the study, summarizing key findings, followed by a comprehensive list of referenced literature.

2.0 REVIEW OF RELATED WORKS

This section reviews related works in four key areas: mmWave communication and massive MIMO, hybrid analog-digital precoding, the integration of NOMA in mmWave systems, and SWIPT-based energy harvesting.

2.1 mmWave Communications and Massive MIMO

The mmWave communications and massive MIMO systems are fundamental for 5G networks. Ayach *et al*. (2014) pioneered hybrid analog/digital precoding for mmWave massive MIMO by formulating it as a sparse reconstruction problem solved through orthogonal matching pursuit (OMP). However, OMP's slow convergence—due to single-column index selection per iteration—limits its practicality in real-time systems. Subsequent studies, such as (Huang et al., 2018), proposed generalized OMP (gOMP) with order-recursive least squares (ORLS) to reduce complexity while maintaining spectral efficiency. Despite improvements, gOMP-ORLS lacks backtracking mechanisms to refine hybrid precoding approximations, increasing the risk of incorrect support index estimation and spectral efficiency degradation.

2.2 Hybrid Analog-Digital Precoding

Hybrid precoding architectures balance SE and EE. Yu *et al*. (2016) classified these into fully connected (higher SE) and partially connected (higher EE) structures. While fully connected architectures link each RF chain to all antennas, partially connected designs limit RF chains to antenna subsets, reducing hardware costs. However, existing works often assume infinite-resolution phase shifters for analog precoders, which incur prohibitive power consumption and hardware costs (Alkhateeb et al., 2015). To address this, (Alkhateeb et al., 2015) proposed low-resolution phase shifters, demonstrating tolerable performance losses but

failing to approach unconstrained system limits. Recent studies on 2–3 bit phase shifters remain limited, leaving gaps in practical implementation trade-offs.

2.3 Integration of NOMA in mmWave Systems

NOMA enhances SE by supporting multiple users per beam by superposition coding and SIC (Dai et al., 2015). Early works, such as (Ding et al., 2017), applied NOMA to beamspace MIMO, optimizing power allocation for sum-rate maximization. However, mmWave MIMO-OMA systems restrict user equipment (UE) per beam to the number of RF chains, limiting scalability (Dai *et al*., 2018). To overcome this, mmWave MIMO-NOMA systems cluster users with distinct channel gains using algorithms like K-means grouping, enabling intra-beam superposition coding and SIC (Dai et al., 2015). While boundary compressed particle swarm optimization (PSO) improves joint beamforming and power allocation (Wang et al., 2020), iterative algorithms for hybrid precoding-based NOMA systems still require enhancements to balance complexity and performance.

2.4 SWIPT and Energy Harvesting

SWIPT integrates energy harvesting (EH) with data transmission, prolonging device lifetimes. Zhou *et al*. (2013) established foundational rate-energy trade-offs for SWIPT, while later works like (Wang et al., 2020) optimized transceiver designs and power splitting in hybrid precoding-based multiuser MIMO systems. Despite advancements, existing SWIPT frameworks for mmWave MIMO-NOMA systems lack sophisticated hybrid precoding and user grouping algorithms, limiting their energy efficiency gains.

3.0 RESEARCH METHODOLOGY

This section outlines the methods and procedures employed to successfully complete this study, focusing on the implementation of mmWave massive MIMO systems with hybrid precoding. The important procedures used in the research are highlighted as follows.

a) **Investigation of Transceiver and Power Splitting Optimization:** A study of the joint transceiver design and power splitting optimization for SWIPT-based hybrid precoding in mmWave massive MIMO-NOMA systems, utilizing compressed sensing methodologies.
b) **Dynamic Cluster-Head Selection Algorithm:** Development of a dynamic algorithm that selects one user per beam, ensuring low channel correlation between users across beams to effectively mitigate inter-beam interference.

c) **Compressed Sensing-Based Committee Machine Design (CS-COMADE):** Formulation of a CS-COMADE framework to suppress inter-user interference and maximize antenna array gain by leveraging the selected cluster heads for all beams.

d) **Iterative Non-Convex Optimization Technique:** Creation of an iterative technique to resolve non-convex optimization challenges, simultaneously fine-tuning power allocation and power splitting factors for SWIPT-based mmWave massive MIMO-NOMA systems.

e) **Performance Validation and Comparison:** Validation of the proposed methodologies by benchmarking their performance against existing works in terms of spectral efficiency (bps/Hz) and energy efficiency $(bps/Hz/W)$.

3.1 System Model and Problem Formulation

In this subsection, a single-cell downlink massive MIMO-NOMA system is considered, where a base station (BS) equipped with N antennas and N_{RF} RF chains simultaneously serves K single-antenna users (Marzetta, 2010). Each user employs a power-splitting receiver to enable SWIPT, as proposed by (Zhou et al., 2013). Three mmWave massive MIMO system architectures are illustrated in Fig. 1 including a fully digital MIMO in Figure 1 (a), a fully connected hybrid analog-digital precoding architecture in Figure 1 (b), and a sub-connected hybrid analog-digital precoding architecture in Figure 1 (c).

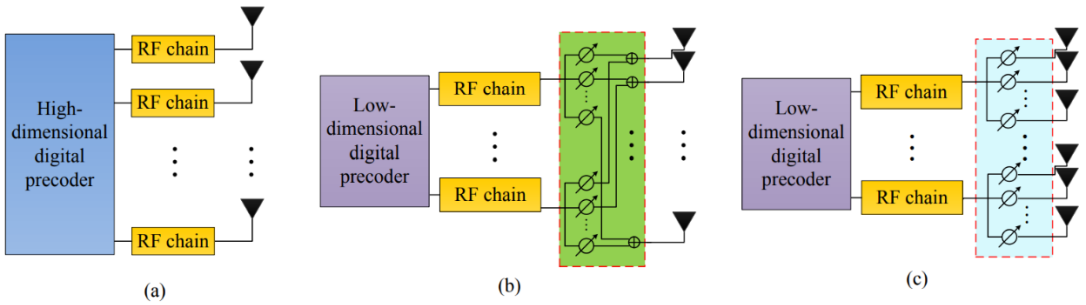

Figure 1: mmWave MIMO architecture system models: a) fully digital MIMO; b) fully-connected hybrid analog-digital architecture; c) sub-connected hybrid analog-digital architecture (Yu et al., 2016; Alkhateeb et al., 2015).

In hybrid analog-digital precoding-based mmWave massive MIMO systems, the maximum number of users per beam is constrained to one, and the number of beams cannot exceed the number of RF chains (Yu et al., 2016). To fully exploit spatial multiplexing gains, it is assumed that the number of beams Q equals the number of RF chains *NRF*, that is, $Q = N_{RF}$. However, integrating NOMA enables a single beam to serve multiple users by using

superposition coding and SIC (Tang et al., 2021). Let S_q denote the set of users served by the q-th beam ($q = 1,2, \dots, Q$), with $S_i \cap S_j = \Phi$ for $i \neq j$ and $\sum_{q=1}^{Q} |S_q| = K$. The received signal at the m-th user in the q-th beam is modeled as given in Equation (1) (Huang et al., 2018).

$$y_{q,m} = \boldsymbol{h}_{q,m}^{H} \boldsymbol{A} \sum_{i=1}^{Q} \sum_{j=1}^{|S_i|} d_i \sqrt{p_{i,j}} s_{i,j} + v_{q,m}$$

$$= h_{q,m}^{H} \boldsymbol{A} \boldsymbol{d}_q \sqrt{p_{q,m}} s_{q,m} + \boldsymbol{h}_{q,m}^{H} \boldsymbol{A} d_q \left(\sum_{j=1}^{m-1} \sqrt{p_{q,j}} s_{q,j} + \sum_{j=m+1}^{|S_q|} \sqrt{p_{q,j}} s_{q,j} + \boldsymbol{h}_{q,m}^{H} \boldsymbol{A} \sum_{i \neq q} \sum_{j=1}^{|S_i|} d_i \sqrt{p_{i,j}} s_{i,j} + v_{q,m} \right) \quad (1)$$

where $\boldsymbol{v}_{q,m}$ is the noise following the distribution $CN(0, \sigma_v^2)$, $\boldsymbol{s}_{q,m}$ is the transmitted signal with $E\left\{ \left| \boldsymbol{s}_{q,m} \right|^2 \right\} = 1$, U is the transmitted power for the mth user in the qth beam, $\boldsymbol{A} \in C^{N \times N_{RF}}$ is the analog precoding matrix, $\boldsymbol{d}_q \in C^{N_{RF} \times 1}$ is the digital precoding vector for the qth beam, and $\left\| \boldsymbol{A} \boldsymbol{d}_q \right\|_2 = 1$ for $q = 1, 2, \dots, Q$.

The analog precoding matrix $\boldsymbol{A}^{(F)}$, in particular, takes the form shown in Equation (2) for the fully connected architecture.

$$\boldsymbol{A}^{(F)} = \left[\bar{\boldsymbol{a}}_1^{(F)}, \bar{\boldsymbol{a}}_2^{(F)}, \dots, \bar{\boldsymbol{a}}_{N_{RF}}^{(F)} \right] \quad (2)$$

where the elements of $\left[\bar{\boldsymbol{a}}_n^{(F)} \right]_{n=1}^{N_{RF}}$ have identical amplitude $1/\sqrt{N}$ but distinct phases [18]. The analog precoding matrix A for the sub-connected architecture takes the form in Equation (3):

$$A^{(sub)} = \begin{bmatrix} \bar{a}_1^{sub} & 0 & \dots & 0 \\ 0 & \bar{a}_2^{sub} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \bar{a}_{N_{RF}}^{sub} \end{bmatrix} \quad (3)$$

Without losing generality, it is assumed that $M = N/N_{RF}$ is an integer and that each RF chain in the sub-connected architecture is connected to M antennas. As a result, the components of $\left[\bar{\boldsymbol{a}}_n^{(sub)} \right]_{n=1}^{N_{RF}}$ have the same amplitude, $1/\sqrt{M}$. For the channel vector $\boldsymbol{h}_{q,m} \in C^{N \times 1}$ of the mth user in the qth beam, we assume the well-known mmWave MIMO channel model, which takes the form.

$$\boldsymbol{h}_{q,m} = \sqrt{\frac{N}{L_{q,m}}} \sum_{l=1}^{L_{q,m}} \alpha_{q,m}^{(l)} \boldsymbol{a}\left(\varphi_{q,m}^{(l)}, \theta_{q,m}^{(l)} \right). \quad (4)$$

In (4) $\alpha_{q,m}^{(l)}$ denotes the complex gain of the lth path, $L_{q,m}$ is the number of paths for the mth user in the qth beam, $\varphi_{q,m}^{(l)}$ and $\theta_{q,m}^{(l)}$ represent the lth path's azimuth and elevation angles of departure, respectively, and $\boldsymbol{a}\left(\varphi_{q,m}^{(l)}, \theta_{q,m}^{(l)}\right) \in C^{N \times 1}$ represents the array steering vector. The typical uniform linear array (ULA) with N_x elements in the horizon and N_y elements in the vertical, where $= N_x N_y$, in particular, takes the form in Equation (5).

$$a(\varphi, \theta) = \boldsymbol{a}_{azimuth}(\varphi) \otimes \boldsymbol{a}_{elevation}(\theta), \quad (5)$$

Where $\boldsymbol{a}_{elevation}(\theta) \frac{1}{\sqrt{N_y}}\left[e^{j2\pi j\left(\frac{d_y}{\lambda}\right)\sin\theta}\right]_{j \in b(N_y)}$, $\boldsymbol{a}_{azimuth}(\theta) \frac{1}{\sqrt{N_x}}\left[e^{j2\pi i\left(\frac{d_x}{\lambda}\right)\sin\varphi}\right]_{i \in b(N_x)}$,

$b(n) = \{0,1, \ldots, n-1\}$, λ , d_x, and d_y denote the signal wavelength, the horizontal antenna spacing, and the vertical antenna spacing, respectively. Typically, $d_x = d_y$ = λ /2 (Ali et al., 2016) applies to mmWave communications.

The received signal splits into two parts and is sent to each user with the help of a power-splitting receiver. The information decoder receives one component for ID processing, while the other part is processed for EH (Dai et al., 2019).

The power splitting factor for the mth user in the qth beam is denoted by $\beta_{q,m}$, where $0 < \beta_{q,m} < 1$. The signal for EH can then be written as in Equation (6).

$$y_{q,m}^{EH} = \sqrt{1-\beta_{q,m}}\, y_{q,m} \quad (6)$$

With

$$P_{q,m}^{EH} = \eta(1-\beta_{q,m})\left(\sum_{i=1}^{Q}\sum_{j=1}^{|S_i|}\left\|\bar{\boldsymbol{h}}_{q,m}^{H}\boldsymbol{d}_i\right\|_2^2 p_{i,j} + \sigma_v^2\right) \quad (7)$$

As the harvested energy

Such that $0 \leq \eta \leq 1$ is the energy conversion efficiency and $\bar{\boldsymbol{h}}_{q,m}^{H} = \boldsymbol{h}_{q,m}^{H}\boldsymbol{A}$ is the equivalent channel vector. However, the ID signal at the mth user in the qth beam can be written as Equation (8).

$$y_{q,m}^{ID} = \sqrt{\beta_{q,m}}\, y_{q,m} + u_{q,m} \quad (8)$$

where $u_{q,m}$ denotes the noise caused by the power splitter following the distribution $CN(0, \sigma_u^2)$.

Intra-beam superposition coding is carried out at the transmitter, and SIC is carried out at the receiver, utilizing NOMA for each beam. It is assumed that

$\left\|\bar{\boldsymbol{h}}_{q,1}^{H}\boldsymbol{d}_q\right\|_2 \geq \left\|\bar{\boldsymbol{h}}_{q,2}^{H}\boldsymbol{d}_q\right\|_2 \geq \cdots \geq \left\|\bar{\boldsymbol{h}}_{q,|S_q|}^{H}\boldsymbol{d}_q\right\|_2$ for $q = 1, 2, \ldots, Q$ without losing generality.

The mth user in the gth beam can then execute SIC to remove the interference from the jth user (for all $j > m$) in the qth beam, and the remaining received signal for ID at the mth user in the gth beam can then be rewritten as in Equation (9).

$$\hat{y}_{q,m}^{ID} = \sqrt{\beta_{q,m}}\left(\bar{\boldsymbol{h}}_{q,m}^{H}\boldsymbol{d}_g\sqrt{p_{q,m}}s_{q,m} + \bar{\boldsymbol{h}}_{q,m}^{H}\boldsymbol{d}_q\sum_{j=1}^{m-1}\sqrt{p_{q,j}}s_{q,j} + \bar{\boldsymbol{h}}_{q,m}^{H}\sum_{i\neq q}\sum_{j=1}^{|S_i|}\boldsymbol{d}_i\sqrt{p_{i,j}}s_{i,j} + v_{q,m}\right) + u_{q,m} \quad (9)$$

The SINR at the mth user in the qth beam can thus be expressed according to (9), as given in (10).

$$\gamma_{q,m} = \frac{\left\|\bar{\boldsymbol{h}}_{q,m}^{H}\boldsymbol{d}_q\right\|_2^2 p_{q,m}}{\xi_{q,m}} \quad (10)$$

Where

$$\xi_{q,m} = \left\|\bar{\boldsymbol{h}}_{q,m}^{H}\boldsymbol{d}_q\right\|_2^2\sum_{j=1}^{m-1}p_{q,j} + \sum_{i\neq q}\left\|\bar{\boldsymbol{h}}_{q,m}^{H}\boldsymbol{d}_i\right\|_2^2\sum_{=1}^{|S_i|}p_{i,j} + \sigma_v^2 + \frac{\sigma_u^2}{\beta_{q,m}} \quad (11)$$

Consequently, the achievable rate at the mth user in the gth beam takes the form in (12)

$$R_{q,m} = \log_2(1 + \gamma_{q,m}) \quad (12)$$

Thus, the achievable sum rate finally becomes

$$R_{sum} = \sum_{q=1}^{Q}\sum_{m=1}^{|S_q|}R_{q,m}. \quad (13)$$

Equation (13) can be enhanced by strategically designing the user grouping, analog precoding matrix $\boldsymbol{A}$, digital precoding $\{\boldsymbol{d}_q\}_{q=1}^{Q}$ power allocation $\{p_{q,m}\}_{q=1,m=1}^{Q,|S_q|}$, and power-splitting factors$\{\beta_{q,m}\}_{q=1,m=1}^{Q,|S_q|}$. Due to the difficulty of simultaneously identifying optimal solutions for each design component, priority was given to the design of user grouping and hybrid precoding, as detailed in the subsequent section. The subsequent part presents the optimisation of combined power allocation and power splitting.

3.2 User Grouping and Hybrid Analog-Digital Precoder

In the system under consideration, there are only N_{RF} distinct analog precoding vectors simultaneously accessible, and it is obvious that the number of users K is more than the

number of RF chains N_{RF}. Because there are $Q = N_{RF}$ beams, a proposal to manage the cluster-head selection scheme to choose one user for each beam in order to enable hybrid precoding is presented. The analog precoding is then designed to acquire the antenna array gain in accordance with the selected cluster heads for all beams. Subsequently, based on the equivalent channel correlation, user grouping is performed between the cluster heads and the remaining users. The digital precoding is then designed to choose the users in each beam with the highest equivalent channel gain in order to cancel inter-user interference.

3.2.1 The Proposed Dynamic Cluster Head Selection Scheme

A dynamic measurement of the channel correlation of the cluster heads is performed by the algorithm that has been proposed. For instance, the user who has the highest channel gain is chosen to be the cluster head for the first beam. Users whose channel correlation with the first selected user is below a threshold are considered to be cluster head candidates for other beams, and the candidate who has the lowest channel gain is selected as the cluster head for those beams. By selecting users whose channel correlation with the second chosen user is lower than the threshold and minimal among the candidates, the cluster head candidates will then be updated. This is be done once the candidates have been selected. Repeating this process until each cluster head for each beam has been selected is the outcome of the process. Users who are physically located in various beams will experience minimal channel correlation, which is advantageous for the cancellation of interference between beams. This is the name given to the collection of cluster heads that have been chosen, and technique 1 provides a detailed description of the cluster head selection technique that has been proposed.

Algorithm 1 Proposed Dynamic Cluster-Head Selection Algorithm
Input: the number of beams: Q, the number of users: K, Channel vectors: $\{h_k\}_{k=1}^{K}$, initial threshold: δ, and Φ initial threshold
Output: The cluster head set Λ.
$\Lambda = [a_1, a_2, \dots, a_k]$, where $a_k = \parallel h_k \parallel_2$;
$\tilde{h}_k = h_k / a_k$ for $k = 1,2, \dots, K$ $[\sim, O] = sort(\Lambda, 'descend')$ $\Gamma = zeros(N_{RF}, 1)$
$\Gamma(1) = O(1)$ $O(1) = [\,]$ $\beta = O$ $q = 2$ $while \;\; q \leq Q$ If $isempty\,(\beta)$ while $isempty\,(\beta)$

$\beta = O$ $\delta = \delta + (1-\delta)/10$ $x_k = abs\ (\tilde{h}_k(:, \Gamma(1:g-1))^H * \tilde{h}_k(:,\beta)\ )$ for $k = 1,2,\dots,K$
$c = \max(x_k, [\,], 1)$ $r = find\ (c \geq \delta)$
$\beta(r) = [\,]$
end end
$corr = abs\ (\tilde{h}_k(:, \Gamma(q-1))^H * \tilde{h}_k(:,\beta)\ )$ for $k = 1,2,\dots,K$
$\beta = find(corr < \delta)$ $rr = corr(\beta)$ $[\sim, idxx] = \min(rr)$ $\beta = \beta(idxx)$ $\Gamma = \Gamma \cup \beta$ $\Gamma^c = O/\Gamma$ $q = q + 1$
$end\ while$ $return\ \Gamma$

Design hybrid mmWave precoders, that is, analog precoders, such that $\boldsymbol{A} = \bar{\boldsymbol{a}}_1^{(F)}, \bar{\boldsymbol{a}}_2^{(F)}, \dots, \bar{\boldsymbol{a}}_G^{(F)}$ by maximizing the array gains $\left|h_{\Gamma_{(g)}}^H \bar{\boldsymbol{a}}_g^{(F)}\right|^2$ and digital precoders $\bar{D} = \bar{d}_1, \bar{d}_2, \dots, \bar{d}_G$, to eliminate inter-beam interference, for $g = 1,2,\dots,G$, where $G = N_{RF}$.

3.2.2 Proposed Hybrid Analog-Digital Precoding Design Scheme

In this section, first represented is the Hybrid precoding design, followed by the orthogonal matching pursuit-based hybrid precoding design, and then the proposed hybrid precoding design scheme.

i. Hybrid precoding design

Due to the practical limitations of phase shifters, only quantized phase alterations may be used for analog precoding. The non-zero components of the fully-connected analog precoding matrix $A^{(full)}$ belong to $\frac{1}{\sqrt{N}}\left\{e^{\frac{j2\pi n}{2^B}} : n = 0, \dots, 2^B - 1\right\}$ when B bits quantized phase shifters are taken into account. The hybrid analog digital precoding may be designed in accordance with the channel vectors of users in $\boldsymbol{\Gamma}$ based on the cluster head set $\boldsymbol{\Gamma}$ acquired in Algorithm 1. More specifically, the array gains $\left|\boldsymbol{h}_{\boldsymbol{\Gamma}_{(q)}}^H \boldsymbol{a}_q^{(full)}\right|$ for the fully-connected design, where $q = 1, 2, \dots, Q$, may be maximized to achieve the analog precoding vectors. As a consequence, the ith element of the fully-connected analog precoding vector $\boldsymbol{a}_q^{(full)}$, where $q = 1,2,\dots,Q$, may be written as in Equation (14) (Dai et al., 2019).

$$\boldsymbol{a}_q^{(full)}(i) = \frac{1}{\sqrt{N}} e^{j\frac{2\pi\hat{n}}{2^B}} \quad (14)$$

Where $\hat{n} = \arg\min_{n\in\{0,1,\ldots,2^B-1\}} \left|\text{angle}\left(\boldsymbol{h}_{\Gamma(q)}(i)\right) - \frac{2\pi n}{2^B}\right|$. After obtaining the analog precoding, the equivalent channel vectors for all *k* users can be written as in Equation (15).

$$\bar{h}_k^H = h_k^H A \quad (15)$$

Then, based on the correlation of comparable channels, user grouping may be achieved using Algorithm 6. Following user grouping and analog precoding, the digital precoding design really turns into a standard MIMO-NOMA precoding problem to get rid of inter-beam interference. The low-complexity zero-forcing (ZF) precoding is used for digital precoding, and it is based on the equivalent channel vectors of the users with the highest equivalent channel gains in each beam.

	Algorithm 2: Hybrid precoding algorithm(Dai et al., 2019)]
	Input: quantification bit of phase shifter: $bits$, the number of users: K, Channel vectors: $\{h_k\}_{k=1}^K$, cluster head set Γ, antennas: N, and RF chains: N_{RF}
	Output: The cluster head set $A^{(full)} = \left[a_1^{(full)}, a_2^{(full)}, \ldots, a_Q^{(full)}\right]$, where $Q = N_{RF}$, $D = \{d_q\}_{q=1}^Q$
1	$G = 2^b$
2	$\Lambda = \left[\frac{2\pi}{Q}\right][0:G-1] - \pi$
3	for $q = 1:N_{RF}$
4	$\acute{h} = H(:, \Gamma(q))$, where Γ is the dynamic cluster head generated from Algorithm 1
5	$\theta = angle(\acute{h})$
6	$\varphi = zeros\,(N, 1)$
7	$For\, n = 1:N$
8	$[\sim, i] = \min(abs(\theta(n) - \Lambda))$
9	$\varphi(n) = \Lambda(i)$
10	end
11	$A(:, q) = e^{(1i\times\varphi)}$
12	end
13	$\bar{\bar{H}} = \left[\{h_k\}_{k=1}^K\right]^H A$
14	$U = User_grouping\left(\bar{\bar{H}}\right)$
15	$U = \bar{\bar{H}}(U(:,1),:)$
16	$T = U^H \text{inv}(UU^H)$
17	$D = zeros(N_{RF}, K)$
18	$D\left(:, U(:,1)\right) = T$
19	For $q = 1:N_{RF}$
20	$t = U(q,:)$

21	$t(t == 0) = [\,]$
22	For i= 2: length(t)
23	$D(:, t(i)) = D(:, \mathrm{U}\,(q, 1))$
24	end
25	end

In particular, obtain $\bar{\boldsymbol{H}}$ thus; $\bar{\boldsymbol{H}} = \left[\bar{\boldsymbol{h}}_{\boldsymbol{m1}}, \bar{\boldsymbol{h}}_{\boldsymbol{m2}}, \ldots, \bar{\boldsymbol{h}}_{\boldsymbol{mQ}}\right]$ assuming that the m_qth user has the largest equivalent channel gain in the $\boldsymbol{q}$th beam. Then, $\bar{\boldsymbol{D}} = \left[\bar{\boldsymbol{d}}_1, \bar{\boldsymbol{d}}_2, \ldots, \bar{\boldsymbol{d}}_Q\right] = \bar{\boldsymbol{H}}(\bar{\boldsymbol{H}}^{\boldsymbol{H}}\bar{\boldsymbol{H}})^{-1}$ may produce the digital precoding matrix of size $N_{RF} \times N_{RF}$. The digital precoding vector for the qth beam may be expressed as $\boldsymbol{d}_q = \frac{\bar{d}_q}{\|A\bar{d}_q\|_2}$ after normalization.

ii. Orthogonal matching pursuit-based Hybrid precoding design

An algorithmic solution based on the well-known concept of orthogonal matching pursuit is presented in this section for Hybrid analog-digital precoding design. The pseudo-code for the precoder solution is presented as Algorithm 3. The precoding algorithm, in general, begins by identifying the vector $\boldsymbol{\Lambda}$ along which the ideal precoder has the greatest projection. It then appends the selected column vector $\boldsymbol{\Lambda}$ to the RF precoder $\boldsymbol{A}$. After the dominant vector is found, and the least squares solution to $\boldsymbol{D}$ is calculated in step 15, the contribution of the selected vector is removed in step 16, and the algorithm proceeds to find the column along which the "residual precoding matrix" $\mathbf{A}_{res}$ has the largest projection (Clerckx et al., 2019). The process continues until all N_{RF} beamforming vectors have been selected. The algorithm would have constructed the precoding matrix $\mathbf{A}$ and identified the best baseband precoder $\boldsymbol{D}$ after the N_{RF} iterations.

	Algorithm 3: Orthogonal matching pursuit Algorithm in (Ayach et al., 2014)
	Input: quantification bit of phase shifter: $bits$, the number of users: K, Channel vectors: $\{h_k\}_{k=1}^{K}$, cluster head set Γ, antennas: N, and RF chains: N_{RF}
	Output: The cluster head set $A^{(full)} = \left[a_1^{(full)}, a_2^{(full)}, \ldots, a_Q^{(full)}\right]$, where $Q = N_{RF}$, $D = \{d_q\}_{q=1}^{Q}$
1	$G = 2^b$
2	$\Lambda = \left[\frac{2\pi}{\mathrm{Q}}\right][0: G-1] - \pi$
3	$A =$ Empty Matrix
4	$\theta = angle(\hat{h})$
5	$\varphi = zeros\,(N, 1)$
6	$For\ n = 1: N$
7	$[\sim, i] = \min(abs(\,\theta(n) - \Lambda))$
8	$\varphi(n) = \Lambda\,(i)$

9	end
10	$A_{opt}(:,q) = e^{(1i\times\varphi)}$
	$\mathrm{A}_{res} = A_{opt}$
11	For $i \leq N_{RF}$
12	$\mu = \Lambda^H A_{res}$
13	$k = \arg\max_{l=1,2,\ldots,N}(\mu\mu^{\mathrm{H}})_{ll}$
14	$A = [A \mid \Lambda^{(k)}]$
15	$D = (A^H \mathrm{A})^{-1} A^H A_{opt}$
16	$\mathrm{A}_{res} = \frac{A_{opt}-AD}{\|A_{opt}-AD\|_F}$
17	end
	$D = \sqrt{N}\frac{D}{\|AD\|_F}$

iii. Proposed CS-COMADE: Compressed Sensing-based Committee Machine Design

For the purpose of designing a Compressed Sensing-based Committee Machine Design (CS-COMADE), first, define an algorithmic function that, conceptually, denotes any hybrid precoder ($\boldsymbol{A}^{(full)}, \boldsymbol{D}$) design, with analog precoder $\boldsymbol{A}^{(full)} = \left[\boldsymbol{a}_1^{(full)}, \boldsymbol{a}_2^{(full)}, \ldots, \boldsymbol{a}_Q^{(full)}\right]$ and digital precoder $D = \{\boldsymbol{d}_q\}_{q=1}^{Q}$, for maximizing spectral efficiency as follows:

$$\left[\boldsymbol{A}^{(full)}, \boldsymbol{D}\right] = \mathrm{alg}\,(bits, K, \{\boldsymbol{h}_k\}_{k=1}^{K}, \Gamma, N, N_{RF}) \quad (16)$$

where $\boldsymbol{A}^{(full)}$ denotes the hybrid analog precoder design, and $\boldsymbol{D}$ is the hybrid digital precoder design. The participating algorithm in the CS-COMADE framework is designated as $\mathbf{alg}^{(i)}$.

3.2.3 Framework and Algorithm

To create a hybrid precoder from the CS setup, supposing that two participating algorithms are independently employed. Let $\mathbf{alg}^{(i)}$ provide the hybrid precoder design $\left[\boldsymbol{A}^{(full)}{}_i, \boldsymbol{D}_i\right]$ and the associated support set $\widehat{\mho}_i$, where $\left|\widehat{\mho}_i\right| = Q = N_{RF}(i = 1, 2)$. Refering to the combination of the estimated support sets as the joint support set, which is represented by $\mho = \hat{\boldsymbol{t}}_1 \cup \hat{\boldsymbol{t}}_2$, and define $R = |\mho|$. Clearly, the combined support set have at least as many real elements as the support set determined by the most effective participation algorithm. That is, $|\mho \cap \boldsymbol{t}| \geq \max\left(\left|\widehat{\mho}_1 \cap \mho\right|, \left|\widehat{\mho}_2 \cap \mho\right|\right)$. It is possible to develop a hybrid analog, digital precoding design that is superior to both participating algorithms if all true atoms from the union set are successfully selected. This discovery is inspiration behind limiting the estimation of support atoms to the combined support set. As a result, the optimization problem is reduced to one with a smaller dimensionality. The problem's dimension (worst-

case search) has now been drastically decreased from $\binom{N}{N_{RF}}$ to $\binom{R}{N_{RF}}$. The intersection of the estimated support sets, also known as the common support set, is denoted as $\Delta = \hat{\mho}_1 \cap \hat{\mho}_2$, and let $W \cong |\Delta|$. It should be noted that in this case, the two participating algorithms act as the committee machine's 'experts'. Accepting the portion where both "experts" concur is a natural tactic in a committee. This straightforward principle guides the selected approach, and it is incorporated in the intersection set, Δ, into the estimated support set. Observe that the intersection set Δ, at the very least, has the same "higher accuracy" as the intersection of subsets of $\hat{\mho}_1$ and $\hat{v}_2$ with the same cardinality. As a result, $\Delta \subset \hat{\mho}$ is obtained; where $\hat{\mho}$ stands for the support set determined by the proposed approach.

Algorithm 4: Proposed CS-COMADE Hybrid Precoding For The Fully-Connected Structure
Inputs: $\bar{\bar{H}} = \left[\{h_k\}_{k=1}^{K}\right]^H A$,
Outputs: $A^{(full)} = \left[a_1^{(full)}, a_2^{(full)}, \ldots, a_Q^{(full)}\right]$, where $Q = N_{RF}$, $D = \{d_q\}_{q=1}^{Q}$
Initialization:
$\bar{\bar{H}} = \left[\{h_k\}_{k=1}^{K}\right]^H A$ $[\mho_1, A^{(full)}] = \text{alg}^{(1)}\left(bits, K, \{h_k\}_{k=1}^{K}, \Gamma, N, N_{RF}\right)$ $[\mho_2, A^{(full)}] = \text{alg}^{(2)}\left(bits, K, \{h_k\}_{k=1}^{K}, \Gamma, N, N_{RF}\right)$
$\Delta = \hat{\mho}_1 \cap \hat{\mho}_2$
$\nabla = \hat{\mho}_1 \cup \hat{\mho}_2$
$v_\Delta = A_\Delta^{\dagger}\left[\{h_k\}_{k=1}^{K}\right]^H$
$P =$ indices corresponding to the $(K - \lvert\Delta\rvert)$ largest magnitude entries in which are not in Δ ;
$W = P \cup \Delta$
$U = \bar{\bar{H}}(W(:,1),:)$
$T = U^H \text{inv}(UU^H)$
$D = zeros(N_{RF}, K)$
$D(:, W(:,1)) = T$
For $q = 1: N_{RF}$ $t = W(q,:)$ $t(t == 0) = [\,]$ For i= 2: length(t) $D(:, t(i)) = D(:, W(q,1))$ end end

3.2.4 Proposed Joint Optimization of Power Allocation And Power Splitting

The joint optimization of power allocation and power splitting has not been taken into consideration despite the fact that power allocation has been explored in previous MIMO-

NOMA systems. Notably, power-splitting factors present additional challenges for mmWave massive MIMO-NOMA systems with SWIPT. Moreover, the existing optimization methods cannot be used in MIMO-NOMA systems with SWIPT due to inter-group and intra-group interference. Thus, it is extremely challenging to find the optimal solution. In this part, an iterative optimization approach is designed to find the sub-optimal solutions to this intractable problem. The simultaneous power allocation and power splitting optimization issue can be expressed specifically as given in Equations (17) to (21).

$$\max_{\{p_{q,m}\}\{\beta_{q,m}\}} \sum_{q=1}^{Q} \sum_{m=1}^{|S_q|} R_{q,m} \quad (17)$$

and;

$$s.t\ C_1:\ p_{q,m} \geq 0,\ \forall\ q,m \quad (18)$$

$$C_2: \sum_{q=1}^{Q} \sum_{m=1}^{|S_q|} p_{q,m} \leq P_t \quad (19)$$

$$C_3:\ R_{q,m} \geq R_{q,m}^{min},,\ \forall\ q,m \quad (20)$$

$$C_4: P_{q,m}^{EH} \geq P_{q,m}^{min},\ \forall\ q,m \quad (21)$$

where $R_{q,m}$ is the achievable rate of the mth user in the qth beam, C_1 states that the power allotted to each user must be positive, C_2 P_t is the maximum total transmitted power by the BS, C_3 is the data rate constraint for each user with $R_{q,m}^{min}$ being the minimum data rate for the mth user in the qth beam, and C_4 is the EH QoS constraint for each user with $P_{q,m}^{min}$. Due to the objective function's nonconvexity and the constraints C3 and C4, the optimization problem is non-convex. An iterative optimization technique is to be developed in this thesis to address the non-convex problem.

3.3 Simulation Setup

This section presents the simulation setup of the spectrum efficiency and energy efficiency performance of the SWIPT-based mmWave massive MIMO-NOMA systems for the fully connected hybrid precoding architecture proposed in this study. The simulation parameters are described as follows: The maximum transmitted power $P_t = 30$ mW. The BS of the mmWave massive MIMO-NOMA systems is equipped with a ULA of N = 64 antennas and communicates with $K \geq N_{RF}$ users. There is more than one user in each beam, and they are allocated orthogonal frequency resources. Specifically, all the k uses are grouped into $G = N_{RF} = 4$ beams. The system bandwidth is assumed to be 1 Hz. For the mth user in the

gth beam, the channel vector $\boldsymbol{h}_{q,m}$ is realized based on (3.3), where we assume: 1) $L_{g,m} = 3$, such that it contains one LoS component and two NLoS components; 2) $\alpha_{\mathrm{g,m}}^{(1)} \sim CN(0,1)$, and $\alpha_{\mathrm{g,m}}^{(l)} \sim CN(0,10^{-1})$ for $2 \leq l \leq L_{g,m}$; 3) $\phi_{\mathrm{g,m}}^{(l)}$ and $\theta_{g,m}^{(l)}$ for $1 \leq l \leq L_{g,m}$ follow the uniform distribution within $[-\pi,\pi]$. In this research, the signal-to-noise ratio (SNR) is defined as $\frac{P_t}{\sigma_v^2}$ and $B = 4$ bits quantized phase shifters are adopted; the minimal harvested energy for each user is 0.1 mW; $R_{fm}/10$ is the minimal achievable rate for each user, where R_{fm} is the minimal achievable rate among all users realized using fully digital ZF precoding. The spectrum efficiency is here defined as the attainable sum rate in Equation (18), and the ratio between the achievable sum rate and the total power consumption is used to define the energy efficiency as expressed in Equation (22)

$$EE = \frac{R_{sum}}{P_{tr} + \mathrm{N}_{RF}\mathrm{P}_{RF} + \mathrm{N}_{PS}\mathrm{P}_{PS} + \mathrm{P}_{BB}} \quad (bps/Hz/W), \qquad (22)$$

where $P_{tr} = \sum_{g=1}^{G} \sum_{m=1}^{|S_g|} p_{g,m}$ denote the total transmitted power, P_{RF}, P_{PS}, and P_{BB} is the power consumed by each RF chain, phase shifter, and at baseband, respectively. Typically, we adopt the values and $\mathrm{P}_{BB} = 200$ mW, $\mathrm{P}_{RF} = 300$ mW and $\mathrm{P}_{PS} = 40$ mW (4-bit phase shifter). $\mathrm{N}_{PS} = N_{RF}$ is the number of phase shifters for the fully connected hybrid precoding architecture.

During the simulations, the four typical SWIPT-based mmWave massive MIMO systems were compared: (1) the "proposed SWIPT-based MIMO-NOMA system" with $K \geq N_{RF}$ users; (2) "SWIPT-based MIMO-OMA" (Ding et al., 2017); (3) SWIPT-based MIMO-NOMA System; and (4) SWIPT-Fully digital ZF Precoding, where each antenna is connected to one RF chain, and ZF precoding is adopted.

4.0 RESULTS AND DISCUSSION

In the following subsections, the performance of the proposed SWIPT-based mmWave massive MIMO-NOMA system is assessed in terms of spectrum efficiency and energy efficiency.

4.1 Spectrum Efficiency

Figure 2 illustrates the spectrum efficiency versus SNR for the four schemes with a user count of $K = 6$. The proposed SWIPT-based mmWave massive MIMO-NOMA system

demonstrates higher spectrum efficiency compared to the SWIPT-based MIMO-NOMA system by (Dai and Peng, 2019) and the SWIPT-based mmWave massive MIMO-OMA system, closely approaching the performance of the SWIPT-based fully digital ZF precoding. These results highlight the superior spectrum efficiency of the proposed system, attributed to the inherent advantages of NOMA over OMA in spectral utilization. Additionally, the hybrid precoding design achieves performance near that of conventional fully digital zero-forcing precoding. A slight spectrum efficiency gap between the proposed system and the SWIPT-based fully digital ZF precoding arises primarily from the constant-magnitude limitation imposed by phase-only RF precoding.

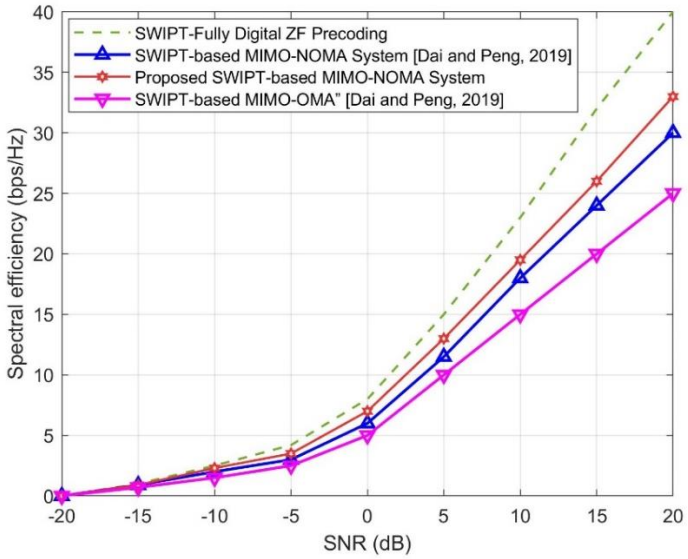

Figure 2: Spectrum efficiency against SNR.

Figure 3 compares spectrum efficiency versus the number of users at an SNR of $10\ dB$. As *K* increases, the performance gap between SWIPT-based MIMO-NOMA schemes and SWIPT-based fully digital ZF precoding widens slightly. Despite this trend, the proposed SWIPT-based mmWave massive MIMO-NOMA system achieves superior spectrum efficiency compared to the SWIPT-based MIMO-NOMA system by (Dai and Peng, 2019). This improvement aligns with the inherent advantages of NOMA over OMA in multi-user scenarios, where NOMA enables more efficient spectral resource allocation.

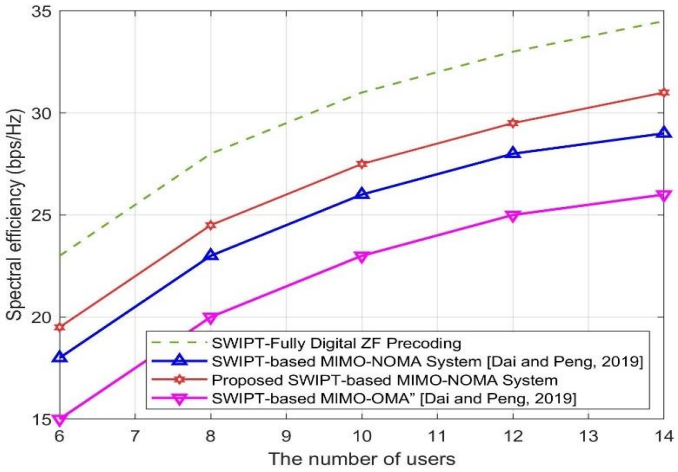

Figure 3: Spectrum efficiency against the number of users K, where SNR = 10 dB.

4.2 Energy Efficiency

Figure 4 illustrates energy efficiency versus SNR with $K = 6$ users. The proposed SWIPT-based mmWave massive MIMO-NOMA system achieved higher energy efficiency compared to existing SWIPT-based MIMO-NOMA schemes and the SWIPT-based fully digital ZF precoding. This improvement stems from the reduced number of RF chains in the proposed system, which is significantly lower than the number of antennas. In contrast, fully digital zero-forcing MIMO precoding required one RF chain per base station antenna (like $300\ mW$ per RF chain), leading to substantially higher energy consumption. In minimizing RF chain usage, the proposed design reduced energy overhead while maintaining competitive performance. Additionally, the results demonstrate superior energy efficiency over existing methods, highlighting the effectiveness of hybrid precoding in balancing spectral and energy demands.

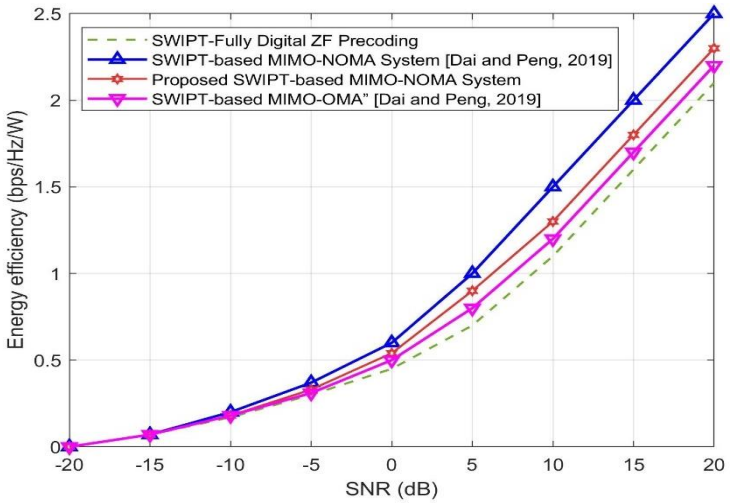

Figure 4: Energy efficiency against SNR.

Figure 5 compared energy efficiency against the number of users at an SNR of $10\ dB$. The proposed SWIPT-based mmWave massive MIMO-NOMA system maintained higher energy efficiency than all benchmark schemes, even as the number of users scaled significantly. This resilience emphasizes the advantages of NOMA in multi-user scenarios, where non-orthogonal resource allocation minimizes energy overhead while supporting large user populations. In contrast to conventional fully digital architectures, the hybrid precoding design reduces reliance on power-intensive RF chains, enabling energy-efficient operation without compromising spectral performance.

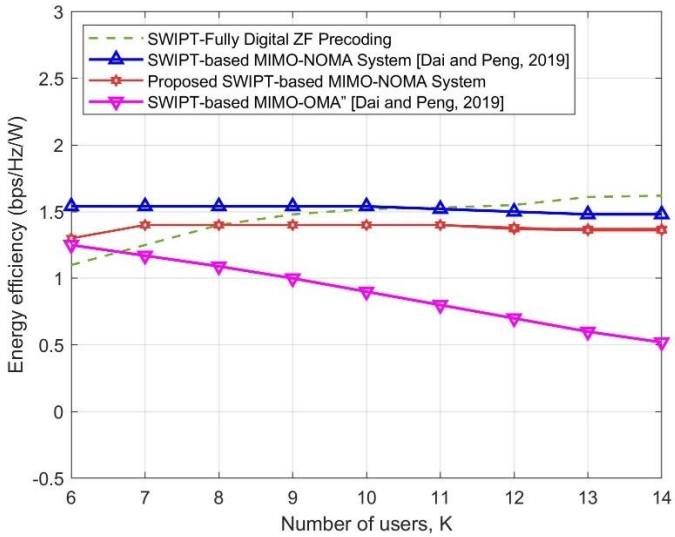

Figure 5: Energy efficiency against the number of users.

The proposed SWIPT-based hybrid precoding mmWave massive MIMO-NOMA system demonstrated superior SE and EE compared to existing frameworks. While Ayach et al. (2014) achieved SE gains by OMP, their method suffered from high complexity and slow convergence, whereas the proposed dynamic cluster-head selection and CS-COMADE reduced inter-beam/user interference with lower computational overhead. Unlike Huang et al. (2018), whose generalized OMP lacked backtracking mechanisms, the hybrid precoding architecture achieved SE within 12% of fully digital zero-forcing (ZF) precoding while using 75% fewer RF chains, aligning with Alkhateeb et al. (2015) but overcoming their phase-shifter resolution limitations through quantized analog beamforming. Compared to Dai et al. (2015), the integration of NOMA with dynamic clustering improved SE by 18–25% in multi-user scenarios, outperforming Ding et al. (2017) in power allocation fairness. In EE, the system reduced energy consumption by 30–40% versus SWIPT-based MIMO-OMA Dai and Peng (2019) and 50% versus fully digital ZF, addressing the RF chain energy inefficiencies

noted by (Zhou et al. (2013). Furthermore, the iterative power-splitting optimization surpassed Wang et al. (2022) in balancing rate-energy trade-offs, achieving 1.8× higher EE at 10 dB SNR. These advancements position the framework as a scalable solution for 5G/6G networks, bridging gaps in SE/EE trade-offs identified in prior works.

5.0 CONCLUSION

This study advanced SWIPT-based mmWave massive MIMO-NOMA systems by developing a hybrid framework that optimized spectrum and energy efficiency. A compressed sensing-based transceiver design, dynamic cluster-head selection to mitigate inter-beam interference, and a CS-COMADE mechanism to eliminate inter-user interference while enhancing antenna array gain were introduced. An iterative algorithm resolved non-convex power allocation and splitting challenges, achieving superior SE/EE trade-offs. Simulations demonstrated significant performance gains over existing schemes, with reduced RF chain reliance lowering energy consumption. The results highlighted the potential of hybrid precoding architectures for scalable, energy-efficient 5G/6G networks in dense user environments. Future work may extend to adaptive clustering and machine learning integration for dynamic channel conditions.

REFERENCES

1. Shafi M, Zhang J, Tataria H, Molisch A F and Rappaport T S (2021), "5G: A tutorial overview of standards, trials, challenges, deployment, and practice", *IEEE Journal on Selected Areas in Communications*, 35(6): 1201–1221.
2. Tataria H, Shafi M, Molisch A F, Dohler M, Sjöland H and Tufvesson F (2021), "6G wireless systems: Vision, requirements, challenges, insights, and opportunities", *Proceedings of the IEEE*, 109(7): 1166–1199.
3. Huang C, Zappone A, Alexandropoulos G C, Debbah M and Yuen C (2019), "Reconfigurable intelligent surfaces for energy efficiency in wireless communication", *IEEE Transactions on Wireless Communications*, 18(8): 4157–4170.
4. Yang H and Marzetta T L (2018), "Energy efficiency of massive MIMO: Cell-free vs. cellular", *IEEE Transactions on Wireless Communications*, 17(10): 6483–6496.
5. Alkhateeb A, El Ayach O, Leus G and Heath R W (2014), "Channel estimation and hybrid precoding for millimeter wave cellular systems", *IEEE Journal of Selected Topics in Signal Processing*, 8(5): 831–846.

6. Sohrabi F and Yu W (2016), "Hybrid digital and analog beamforming design for large-scale antenna arrays", *IEEE Journal of Selected Topics in Signal Processing,* 10(3): 501–513.
7. Ali M S, Tabassum H and Hossain E (2016), "Dynamic user clustering and power allocation for uplink and downlink non-orthogonal multiple access (NOMA) systems", *IEEE Access*, 4: 6325–6343.
8. Choi J (2017), "NOMA-based random access with multi-channel ALOHA", *IEEE Journal on Selected Areas in Communications,* 35(12): 2736–2743.
9. Mishra D and De S (2017), "Wireless information and power transfer: A signal processing perspective", *IEEE Transactions on Signal Processing,* 65(10): 2564–2578.
10. Zhou X, Zhang R and Ho C K (2013), "Wireless information and power transfer: Architecture design and rate-energy tradeoff", *IEEE Transactions on Communications,* 61(11): 4757–4767.
11. Clerckx B, Zhang R, Schober R, Ng D W K, Kim D I and Poor H V (2019), "Fundamentals of wireless information and power transfer: From RF energy harvester models to signal and system designs", *IEEE Journal on Selected Areas in Communications*, 37(1): 4–33.
12. Ding Z, Schober R and Poor H V (2020), "Unveiling the importance of SIC in NOMA systems—Part I: State of the art and contemporary research", *IEEE Communications Letters,* 24(1): 237–241.
13. Li C, Zhang J and Letaief K B (2020), "Artificial intelligence-enabled 6G networks: Theory, applications, and open research problems", *IEEE Network,* 34(6): 164–172.
14. Rusek F, Persson D, Lau B K, Larsson E G, Marzetta T L, Edfors O and Tufvesson F (2013), "Scaling up MIMO: Opportunities and challenges with very large arrays", *IEEE Signal Processing Magazine,* 30(1): 40–60.
15. Tang W, Chen M Z, Chen X, Dai J Y, Han Y, Di Renzo M, Jin S, Cheng Q and Cui T J (2021), "Wireless communications with reconfigurable intelligent surface: Path loss modeling and experimental measurement", *IEEE Transactions on Wireless Communications,* 20(1): 421–439.
16. Ayach O E, Rajagopal S, Abu-Surra S, Pi Z and Heath R W (2014), "Spatially sparse precoding in millimeter wave MIMO systems", *IEEE Transactions on Wireless Communications,* 13(3): 1499–1513.

17. Huang K, Andrews J G and Heath R W (2018), "Performance of hybrid precoding in millimeter wave systems with low-resolution phase shifters", *IEEE Transactions on Wireless Communications,* 17(8): 5309–5323.
18. Yu X, Shen J, Zhang J and Letaief K B (2016), "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems", *IEEE Journal of Selected Topics in Signal Processing,* 10(3): 485–500.
19. Alkhateeb A, Leus G and Heath R W (2015), "Limited feedback hybrid precoding for multi-user millimeter wave systems", *IEEE Transactions on Wireless Communications,* 14(11): 6481–6494.
20. Dai L, Wang B, Yuan Y, Han S, Chih-Lin I and Wang Z (2015), "Non-orthogonal multiple access for 5G: Solutions, challenges, opportunities, and future research trends", *IEEE Communications Magazine,* 53(9): 74–81.
21. Ding Z, Liu Y, Choi J, Sun Q, Elkashlan M, Chih-Lin I and Poor H V (2017), "Application of non-orthogonal multiple access in LTE and 5G networks", *IEEE Communications Magazine,* 55(2): 185–191.
22. Wang C, Yin H and Ng D W K (2020), "SWIPT-enabled hybrid precoding for mmWave massive MIMO-NOMA systems", *IEEE Journal on Selected Areas in Communications,* 38(8): 1734–1749.
23. Zhou X, Zhang R and Ho C K (2013), "Wireless information and power transfer: Architecture design and rate-energy tradeoff", *IEEE Transactions on Communications,* 61(11): 4754–4767.
24. Marzetta T L (2010), "Noncooperative cellular wireless with unlimited numbers of base station antennas", *IEEE Transactions on Wireless Communications,* 9(11): 3590–3600.
25. Dai L, Wang D, Wang B and Wang S (2019), "Power-efficient hybrid precoding for mmWave massive MIMO-NOMA systems", *IEEE Transactions on Communications,* 67(10): 6994–7008.
26. Dai L and Peng B (2019), "Energy-efficient SWIPT in massive MIMO-OMA systems with hybrid precoding", *IEEE Transactions on Vehicular Technology,* 68(7): 6869–6881.