IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 28 January 2023, accepted 8 March 2023, date of publication 31 March 2023, date of current version 6 April 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3263548

RESEARCH ARTICLE

Modified User Grouping and Hybrid Precoding for Information Decoding and Energy Harvesting in Hardware Impaired Point-To-Point MM-Wave Massive MIMO NOMA

WAHEED M. AUDU, (Member, IEEE), AND OLUTAYO O. OYERINDE, (Senior Member, IEEE)
School of Electrical and Information Engineering, University of the Witwatersrand, Johannesburg 2020, South Africa

Corresponding author: Waheed M. Audu (waheed.audu1@students.wits.ac.za)

This work was supported in part by the National Research Foundation (NRF) of South Africa through NRF Competitive Program for Rated Researchers under Grant 118547; and in part by NRF, South Africa/National Centre for Research and Development [Naradowe Centrum Badan I Rozwoju (NCBR)], Poland Joint Science, and Technology Research Collaboration under Grant 118678.

ABSTRACT The fifth generation (5G) and beyond 5G race covers both spectral-efficient and energy-efficient mechanisms such as simultaneous wireless information decoding and energy harvesting (SWIDEH). Structured as full or sub-connection of hybrid-precoded mm-wave massive multiple-input multiple-output non-orthogonal multiple access (NOMA) systems. However, the more practical hardware-impaired (HI) functionality yet to be considered for point-to-point (p-to-p) MIMO is conceived from an existing system. A hardware impairment-based power splitting is proposed for energy harvesting. The energy of the HI signal is harnessed through the joint optimization of power allocation (PA) for information decoding (ID) and the power splitting (PS) factor for energy harvesting (EH) using zero-forcing (ZF) precoding and minimum mean square error (MMSE) downlink (DL) detection. Based on the simulation results, an adaptive cluster head selection (CHS) criterion is proposed for analog precoding to mitigate intracluster interferences at a high signal-to-noise ratio (SINR). A successive interference cancellation (SIC) is carried out to enable dynamic channel overhead estimation and user signal detection. An investigation is done by simulating the system performance at varied numbers of users, the number of radio-frequency (RF) chains, and channel characteristics for modelling the user group threshold coupled with channel size reduction via an antenna selection scheme. Observing the execution time and some graphical plots shows that the spectral and energy efficiency can be improved by a random (R) or maximum norm average (MNA) channel vector formation from each user's channel matrix combined with the best adaptive correlation threshold during CHS and a regularized ZF (RZF) precoder.

INDEX TERMS Adaptive correlation threshold, hardware-impaired (HI), hybrid-precoding, massive multiple-input multiple-output (mMIMO), mm-wave (mmWave), non-orthogonal multiple access (NOMA), point-to-point (p-to-p) MIMO, simultaneous wireless information decoding and energy harvesting (SWIDEH).

I. INTRODUCTION

Simultaneous wireless information decoding and energy harvesting (SWIDEH) is otherwise known as simultaneous wireless information and power transfer (SWIPT) [1]. SWIDEH plays a major role in current and future technology to ensure reliable mobile stations in wireless communication networks. Therefore, in addition to information decoding, radio frequency (RF) energy from a base station (BS) is harnessed by a mobile station (MS). Four major architectures for the realization proposed in the literature are a separate receiver, time-switching (TS), antenna switching (AS), and power

The associate editor coordinating the review of this manuscript and approving it for publication was Marco Martalo.

splitting (PS) based [2]. The PS option provides the best theoretical guarantees to enable SWIDEH.

Some work has been done to jointly optimize the transmitter at the BS, the PS factor (PSF), and the information received at the MS under predefined energy and spectrum efficiency constraints. This bears root in full or sub-connection of hybrid-precoded mm-wave massive multiple-input multiple-output (HP-mmWave-mMIMO) non-orthogonal multiple access (NOMA) based technologies among others [3], [4]. The work of [5] showed that residual interference is useful for boosting harvested energy which is detrimental to the achievable sum rate in training-based hybrid TS/PS relay-assisted mMIMO downlink (DL) operating SWIPT mode. This energy-rate trade-off is phenomenal with practical viability to boost SWIDEH performance when DL spatial multiplexing is exploited in mMIMO. A similar relaying protocol for the orthogonal multiple access (OMA) systems is found in [6] to optimize throughput performance with partially activated relay nodes that can be fixed or optimized. Apart from OMA relay networks, SWIDEH has found equivalent diversity gain in fixed power and cognitive radio NOMA which is effective in reducing outage probability [7]. The closed-form analytical results derived from the two power allocation strategies revealed the different trade-offs between user fairness, reception reliability, and system complexity.

Efforts have been made to improve wireless power transfer (WPT) without degrading wireless information transfer (WIT) in code and power domain NOMA, however, the focus has been on the power domain for easy integration into current systems [8]. One power domain solution is the joint design of an energy interleaver and constellation rotation-based modulator in the symbol block level to constructively superimpose the symbols destined for WIT users thereby guaranteeing the symbol-error-rate (SER) of all users [9]. Energy maximization at minimum rate requirement with two user grouping strategies and a novel power allocation scheme of [10] showed very high spectral efficiency from a large number of BS antenna to radio frequency (RF) chain ratio (256:8) at the expense of placing a lower bound on the signal-to-noise ratio (SNR$\geq$5 dB) to which SWIDEH can be realized compared to a better value in [11]. In [12] the joint impact of hardware impaired (HI) and imperfect channel state information (ICSI) on cooperative SWIPT NOMA in muti-relay systems was studied. The authors used partial relay selection (PRS), closed-form analytic expressions verified with numerical results for outage probability performance of users to reveal the negative impact of transceiver HIs and ICSI that can be significantly reduced by increasing the number of relays.

A. CONTRIBUTIONS

While most works referenced herein, and many others focused on multi-users (MU) single antenna and zero-forcing (ZF) precoders except for few others that considered point-to-point MIMO without HI or SWIDEH options [13], [14], [15], this work intends to re-imagine the work of [11] in a non-relaying form as follows.

- We show how point-to-point MIMO channel dimension is reduced into multiple user single antenna mm-wave massive MIMO NOMA channel.
- We document mathematical expressions relating HI to PS factor (PSF).
- Development of an adaptive thresholding model for the cluster head selection (CHS) scheme during user grouping.
- Demonstration of the robustness of regularized zero-forcing (RZF) over conventional ZF precoders in managing HIs constituting ICSI for OMA and NOMA via simulation results.

B. ORGANIZATION AND NOTATION

The rest of this paper is organized in the following manner. Section II centers on the system model and conceptual CSI dimension reduction. Section III presents the mathematical relationship between PSF and HI in the system model. Section IV provides details on the iterative procedure for precoder design, and user grouping with PA and PSF to attain an optimized achievable sum rate (ASR). In section V, simulation results are presented with and without a derived adaptive user group threshold for ZF precoders in comparison with the RZF for discussion. The conclusion of the work can be found in section VI.

Notation: For the rest of this paper, bold upper and lower case signify matrices ($e.g.$, $\mathbf{X}$, $\mathbf{Y}$) and vectors ($e.g.$, $\mathbf{x}$, $\mathbf{y}$), respectively. The transpose and Hermitian transpose operators are designated as $(.)^T$ and $(.)^H$, respectively. The diag($\mathbf{x}$) is the vector-to-matrix conversion with elements of $\mathbf{x}$ on its diagonal while for an array of vectors $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$ the blkdiag$[\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N]$ is used. $E[y]$ denotes the expectation of y. The complex normal random variable with scalar mean x and variance y is $\mathbb{CN}(x, y)$. A random variable W with uniform distribution bounded by (x, y) is represented as $W \sim \mathcal{U}(x, y)$. The set of complex and real matrices or vectors of dimension $N_x - by - N_y$ is $\in \mathbb{C}^{(N_x \mathrm{x} N_y)}$ and $\in \mathbb{R}^{(N_x \mathrm{x} N_y)}$, respectively. The real part of a complex number X is Re$[X]$. The union and intersection of set X and Y are $X \bigcup Y$ and $X \bigcap Y$. The null set is Φ, and the number of elements in a set X is $|X|$ (magnitude or size of X). The p-norm of Y is $||Y||_p$. The symbol # is used to represent the number of items (e.g., #pins connote the number of pins).

II. SYSTEM MODEL WITH THE PROPOSED CSI DIMENSION REDUCTION

Consider the hybrid precoded mm-wave massive point-to-point [16] multiple-input multiple-output (MIMO) non-orthogonal multiple access (NOMA) system of Fig.1 which requires precoder design and power allocation scheduling based on channel state information (CSI) [17], [18]. There exist heavily correlated channels of users in the same

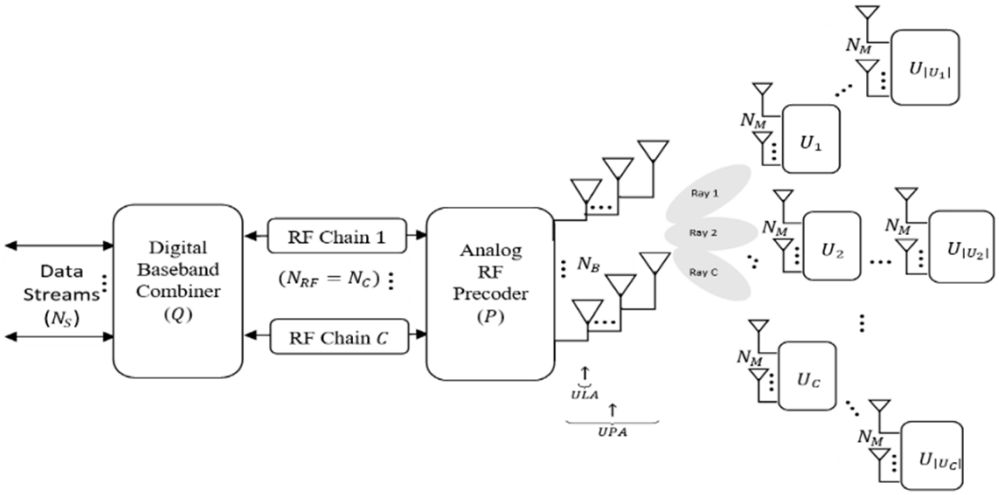

FIGURE 1. Hybrid precoded (HP) mmWave Massive MIMO NOMA transceiver System.

direction and the highly directed beams of the large-scale antenna in mm-wave enable NOMA deployment in each beam for larger beamforming gains and lower inter-beam interference compared to their OMA counterparts. Fully-connected and sub-connected hybrid architectures with energy harvesting (EH) capabilities for downlink communication with uplink channel estimation have been proposed by [11]. Careful consideration of the paper portends the possibility of adopting concepts from [7], [19], and [20], to expand the work for users with multiple antennas, simultaneous (complicated) cluster head selection for user grouping, and analog/digital precoding.

From Fig.1, the transceiver system is modelled as a single cell such that the number of radio-frequency (RF) chains N_{RF} are fully utilized by an equal number of beams/rays, groups, or clusters (N_C). The set of users in a cluster is denoted as $U_c, c = 1, 2, \ldots, C$ satisfying $U_i \bigcap U_j = \Phi, i \neq j$, $|U_c| \geq 1$, and $\sum_{c=1}^{C} |U_c| = K$. An uplink initial channel state information (CSI) is assumed for downlink user detection such that the received signal vector $\mathbf{y}_c \in \mathbb{C}^{(N_M \text{x} 1)}$ at an arbitrary cluster c is represented as

$$\mathbf{y}_c = \mathbf{H}_c^H \mathbf{P} \sum_{i=1}^{C} \sum_{j=1}^{|U_i|} \mathbf{q}_i \sqrt{v_{i,j}} s_{i,j} + n_c [11], \tag{1}$$

where $\mathbf{H}_c$ is the channel matrix of N_B (number of base station transmitter antenna) by N_M (number of mobile station receiver antenna) per cluster. The corresponding downlink Gaussian noise vector is $\mathbf{n}_c \in \mathbb{C}^{(N_M \text{x} 1)}$. The analog precoding matrix $\mathbf{P} \in \mathbb{C}^{(N_B \text{x} N_C)}$ and digital precoding vector $\mathbf{q} \in \mathbb{C}^{(N_C \text{x} 1)}$ for the cluster, c are related in the form $||\mathbf{P}\mathbf{q}_c||_2 = 1$. The transmitted signal $s_{i,j}$ is such that $E[|s_{i,j}|^2] = 1$. The transmitted power to the ith cluster and jth user is $v_{i,j}$. The precoder architecture is two folds: The fully-connected (FC) type $\mathbf{P}^{(F)}$ which is spectral efficient (SE) and the sub-connected (SC) type $\mathbf{P}^{(S)}$ which is energy efficient (EE) in terms of the number of radio frequencies (RFs) mapped to the transmitting antenna. So while the SE $\mathbf{P}^{(F)} = [\mathbf{p}_1^{(F)}, \mathbf{p}_2^{(F)}, \ldots, \mathbf{p}_C^{(F)}]$ with $\mathbf{p}_c^{(F)} \in \mathbb{C}^{(N_B \text{x} 1)}$, the EE $\mathbf{p}_C^{(S)} = \text{blkdiag}[\mathbf{p}_{\text{D}_1}^{(S)}, \mathbf{p}_{\text{D}_2}^{(S)}, \ldots, \mathbf{p}_{\text{D}_C}^{(S)}]$ for $\mathbf{p}_{\text{D}_c}^{(S)} \in \mathbb{C}^{(N_{\text{D}_c} \text{x} 1)}$, $c = 1, 2, \ldots, C$ being vectors with $N_{\text{D}_c}(\frac{N_B}{N_C})$ rows/antennas spanned across N_C or N_{RF} columns/RF of an N_BxN_C block diagonal matrix [11]. All clusters for any of the two connection types share similar amplitude but different phases, However, similar to [11] and [18], channel models can be derived as [21]:

$$\mathbf{H}_c = \sqrt{\frac{N_B N_M}{\rho_c}} \sum_{i=1}^{L_c} \alpha_c^{(l)} \mathbf{a}_B(\theta_B^{(l)}) \mathbf{a}_M(\phi_M^{(l)}), \tag{2}$$

where the number of paths for users in a cluster c is L_c. Inspired by [11], each path is characterized by a complex gain $\alpha_c^{(l)}$, angle of departure (AOD), $\theta_B^{(l)}$, an azimuth angle of arrival (AOA), $\phi_M^{(l)}$, and the associated average path loss ρ_c [21] can be represented by the number of paths, L_c. Clustering simplifies the expression such that the N_BxN_M channel matrix of (2) and the N_Mx1 noise vector of (1) is reduced through common size co-located N_M per user k in cluster c by reshaping the matrix $\mathbf{H}_c$ and assuming the same element-wise noise distribution in $\mathbf{n}_c$ to obtain $\mathbf{h}_{c,k}$ and $n_{c,k}$ respectively. The resulting representative channel model can then be expressed as

$$\begin{aligned} \mathbf{h}_{c,k} &= \sqrt{\frac{N_B N_M}{L_{c,k}}} \sum_{i=1}^{L_{c,k}} \alpha_{c,k}^{(l)} \mathbf{a}^{(m)}(\theta_{c,k}^{(l)}, \phi_{c,k}^{(l)}) \\ &= \sum_{i=1}^{L_{c,k}} z_{c,k}^{(l)} \mathbf{a}^{(m)}(\theta_{c,k}^{(l)}, \phi_{c,k}^{(l)}), \end{aligned} \tag{3}$$

where $z_{c,k}^{(l)} = \sqrt{\frac{N_B N_M}{L_{c,k}}} \sum_{i=1}^{L_{c,k}} \alpha_{c,k}^{(l)}$ and $\mathbf{a}^{(m)}(\theta_{c,k}^{(l)}, \phi_{c,k}^{(l)})$ is the steering vector projected onto one antenna out of N_M antennas that are chosen at random, indexed m for all K users when implemented as detailed in [21] for both uniform linear array (ULA) and uniform planner

array (UPA). The received signal at the kth user is then written as

$$\mathrm{y}_{c,k} = \mathbf{h}_{c,k}^H \mathbf{P} \sum_{i=1}^{C} \sum_{j=1}^{|U_i|} \mathbf{q}_i \sqrt{v_{i,j}} s_{i,j} + n_{c,k}[11], \tag{4}$$

the expanded form is the same as (see page 195 of [22])

$$\begin{aligned}\mathrm{y}_{c,k} &= \mathbf{h}_{c,k}^H \mathbf{P} \mathbf{q}_c \sqrt{v_{c,k}} s_{c,k} \\ &+ \mathbf{h}_{c,k}^H \mathbf{P} \mathbf{q}_c \left(\sum_{j=1}^{k-1} \sqrt{v_{c,j}} s_{c,j} + \sum_{j=k+1}^{|U_c|} \sqrt{v_{c,j}} s_{c,j} \right) \\ &+ \mathbf{h}_{c,k}^H \mathbf{P} \sum_{i \neq c}^{C} \sum_{j=1}^{|U_i|} \mathbf{q}_i \sqrt{v_{i,j}} s_{i,j} + n_{c,k},\end{aligned} \tag{5}$$

where the expressions on the right of the equality sign represent the desired, intra-beam interference, inter-beam interference, and noise signals respectively. The noise signal $n_{c,k}$ is $\mathbb{CN}(0, \sigma^2)$.

III. THE MATHEMATICAL RELATIONSHIP BETWEEN HI AND PS FACTOR

The power splitting factor for simultaneous information decoding and energy harvesting [11] can be seen as a function of residual hardware impairment (RHI) of signal in mmWave massive MIMO transmitter sub-system [21], see page 433 of [22]. This is incorporated into (4) by assuming $s_{i,j} = \bar{s}_{i,j} + e_B$, where e_B accounts for the transmitter RHI (TRHI) signal. Let $e_B = \bar{e}_{i,j}$ so that

$$s_{i,j} = \bar{s}_{i,j} + \bar{e}_{i,j}, \tag{6}$$

The desired signal for the kth user in cluster c becomes

$$\mathrm{y}_{\text{desired}} = \mathbf{h}_{c,k}^H \mathbf{P} \mathbf{q}_c \sqrt{v_{c,k}} (\bar{s}_{c,k} + \bar{e}_{c,k}), \tag{7}$$

where $\bar{s}_{c,k}$ is the pure signal which can be normalized by $s_{c,k}$ in (6) to get $\frac{s_{c,k}}{s_{c,k}} = \frac{\bar{s}_{c,k}}{s_{c,k}} + \frac{\bar{e}_{c,k}}{s_{c,k}}$, $1 = \frac{\bar{s}_{c,k}}{s_{c,k}} + \frac{\bar{e}_{c,k}}{s_{c,k}}$, $\frac{\bar{e}_{c,k}}{s_{c,k}} = 1 - \frac{\bar{s}_{c,k}}{s_{c,k}}$, and the RHI factor ($\bar{e}_{c,k}^{\mathrm{EH}} = \frac{\bar{e}_{c,k}}{s_{c,k}}$) that can be harvested given a pure signal-splitting factor ($\beta_{c,k} = \frac{\bar{s}_{c,k}}{s_{c,k}}$) is $\bar{e}_{c,k}^{\mathrm{EH}} = 1 - \beta_{c,k}$ or $\sqrt{\bar{e}_{c,k}^{\mathrm{EH}}} = \sqrt{1 - \beta_{c,k}}$, where $\beta_{c,k} : 0 \leq \beta_{c,k} \leq 1$. $\beta_{c,k}$ becomes the power splitting factor for which the energy harvesting signal in [11] is expressed:

$$\mathrm{y}_{c,k}^{\mathrm{EH}} = \sqrt{1 - \beta_{c,k}} \mathrm{y}_{c,k}. \tag{8}$$

Similarly, when $\bar{\mathbf{h}}_{c,k}^H = \mathbf{h}_{c,k}^H \mathbf{P}$ from (4) the harvested energy in terms of signal power [11]:

$$\mathrm{P}_{c,k}^{\mathrm{EH}} = \eta(1 - \beta_{c,k}) \left(\sum_{i=1}^{C} \sum_{j=1}^{|U_i|} ||\bar{\mathbf{h}}_{c,k}^H \mathbf{q}_i||_2^2 v_{i,j} + \sigma_n^2 \right), \tag{9}$$

where the energy conversion efficiency is given as $\eta : 0 \leq \eta \leq 1$. The signal for information decoding (ID) at the kth user in cluster c is

$$\mathrm{y}_{c,k}^{\mathrm{ID}} = \sqrt{\beta_{c,k}} \mathrm{y}_{c,k} + u_{c,k}[11]. \tag{10}$$

where the power-splitting noise, $u_{c,k}$ is equivalent to the receiver RHI [21], e_M which follows the distribution $u_{c,k} = \mathbb{CN}(0, \sigma_u^2)$ and suites the concerns of [23] as it affects NOMA systems.

IV. ITERATIVE HP DESIGN, USER GROUPING WITH PA AND PSF OPTIMIZATION

The NOMA concept is applied in each cluster where intra-beam interference is circumvented by successive interference cancellation (SIC) for multi-user signal detection according to [24]. The enabling assumption, without loss of generality, is that $||\bar{\mathbf{h}}_{c,1}^H \mathbf{q}_i||_2^2 \geq ||\bar{\mathbf{h}}_{c,2}^H \mathbf{q}_i||_2^2 \geq \ldots \geq ||\bar{\mathbf{h}}_{c,|U_i|}^H \mathbf{q}_i||_2^2$ for $i \triangleq c = 1, 2, \ldots, C$. This user ordering due to near-far effect enables the kth user in cluster c to cancel interference from decodeable intra-beam users ($i > k$) with lower effective channel gain while CSI overhead can be avoided using zero-forcing (ZF) detection scheme [25]. Besides, SIC is insensitive to imperfect CSI by 88% in mm-wave MIMO systems [26]. The resulting received signal for ID from (10) is

$$\begin{aligned}\mathrm{y}_{c,k}^{\mathrm{ID}} = \sqrt{\beta_{c,k}} \Bigg(&\bar{\mathbf{h}}_{c,k}^H \mathbf{q}_c \sqrt{v_{c,k}} s_{c,k} + \bar{\mathbf{h}}_{c,k}^H \mathbf{q}_c \sum_{j=1}^{k-1} \sqrt{v_{c,j}} s_{c,j} \\ &+ \bar{\mathbf{h}}_{c,k}^H \sum_{i \neq c}^{C} \sum_{j=1}^{|U_i|} \mathbf{q}_c \sqrt{v_{i,j}} s_{i,j} + n_{c,k} \Bigg) + u_{c,k}[11],\end{aligned} \tag{11}$$

where $s_{c,k} = \bar{s}_{c,k} + \bar{e}_{c,k}$ and the signal-to-interference-plus-noise ratio (SINR) at the kth user in cluster c is expressed as

$$\mathrm{SINR}_{c,k} = \frac{||\bar{\mathbf{h}}_{c,k}^H \mathbf{q}_c||_2^2 v_{c,k}}{\zeta_{c,k}}[11], \tag{12}$$

where

$\zeta_{c,k} = (||\bar{\mathbf{h}}_{c,k}^H \mathbf{q}_i||_2^2 \sum_{j=1}^{k-1} v_{c,j} + \sum_{i \neq c}^{C} ||\bar{\mathbf{h}}_{c,k}^H \mathbf{q}_i||_2^2 \sum_{j=1}^{|U_i|} v_{i,j} + \sigma_n^2 + \frac{\sigma_u^2}{\beta_{c,k}})$. The matching achievable rate is expressed as

$$R_{c,k} = \log(1 + \mathrm{SINR}_{c,k})[11]. \tag{13}$$

Furthermore, the achievable sum rate (ASR) is

$$R_{\mathrm{sum}} = \sum_{c=1}^{C} \sum_{k=1}^{|U_i|} R_{c,k}[11]. \tag{14}$$

The ASR is one way to control the maximal admissible users per cluster with a lower bound of $1 + \Delta$, the Δ is dependent on the interference tolerance capability [27]. It gets better with careful design of the five parameters involved. The parameter design can be ordered as follows. User grouping (c) before analog precoding matrix (P), then digital precoder (q_c) followed by power allocation ($v_{c,k}$) and splitting ($\beta_{c,k}$), where $c = 1, 2, \ldots, C$ and $k = 1, 2, \ldots, |U_c|$. Optimizing all five parameters at the same time could be challenging. However, a modular approach is made possible by addressing the user grouping and hybrid precoding before jointly optimizing the

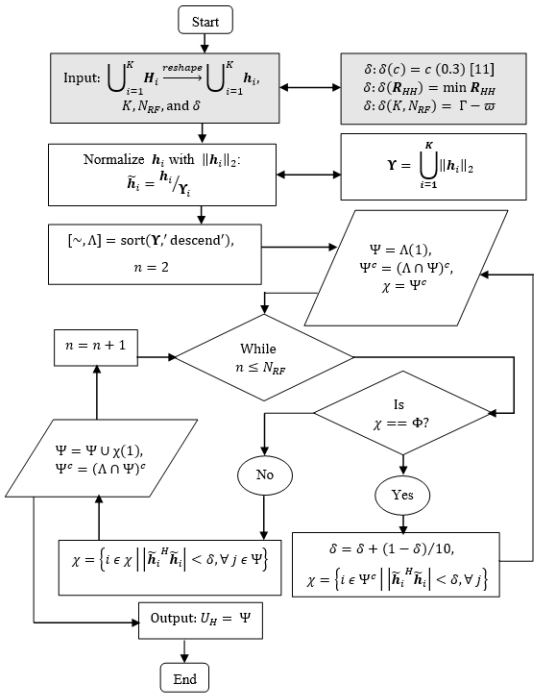

FIGURE 2. Flow chart for the proposed cluster head selection (CHS).

power allocation and splitting. In NOMA, the number of users K is more than the number of RF chains N_{RF} or beams (clusters, N_c) accessed at the same frequency and time. The implication is that only $N_{\text{RF}} = N_c$ precoding vectors are available to serve the K users. This motivates hybrid precoding schemes such as cluster head selection (CHS) algorithm that satisfies $1 + \Delta$, the lower bound for ASR [11]. One user is selected as a cluster head for each one of the ($N_{\text{RF}} = N_c$) beams first before designing a suitable analog precoder based on their representative antenna array gains. Other users are grouped with the most correlated cluster head based on the representative CSI. So, while analog precoding is designed to obtain all user antenna array gain using a selected cluster head in each beam, digital precoding is designed to cancel inter-user interference by selecting users with the strongest representative channel gain per beam. The channel estimation method in [21] may be assumed before clustering.

A detailed CHS algorithm according to [11] begins with minimizing the channel correlation between selected cluster heads for inter-beam interference cancellation. All user channels below the set correlation threshold become candidates for the next cluster selection of which the user with the highest gain becomes the next cluster head. The process is repeated for all RFs with successive thresholds updated for every cluster head search to form a set of cluster heads, U_H. The entire procedure is summarized in the flow chart of Fig.2. It is interesting to note that [11] suggested that alternative algorithms that will perform better must exceed the polynomial complexity of the CHS algorithm. However, the simulation results of section V shows that significant improvements can still be achieved with the same algorithm by making the initial correlation threshold (δ) adaptive in finding the first cluster head before updating it to get subsequent cluster heads. The shaded blocks represent areas of contribution. The first shaded block constitutes the CSI dimension reduction, explained earlier in section II. This is suitable for comparison with [11] in the second shaded block where (δ) is revised from a constant value to a dependent variable. It is discussed later in section V.

More sophisticated hybrid precoding schemes are proposed in the literature [11], and the reason is stated as follows. A known bit quantized phase shifter is practical for analog precoding yet serves as a constraint to accuracy. The elements of the analog precoding matrix will therefore take values limited to quantized forms:

$$\mathbf{P}^{(F)} = \frac{1}{\sqrt{N_B}} e^{-j\frac{2\pi n}{2^{\#\text{bits}}}} \text{(Fully connected)[11]}, \tag{15}$$

$$\mathbf{P}^{(S)} = \frac{1}{\sqrt{N_{\text{D}_c}}} e^{-j\frac{2\pi n}{2^{\#\text{bits}}}} \text{(Sub-connected)[11]}, \tag{16}$$

where $N_{\text{D}_c} = \frac{N_B}{N_C}$ and $n = 0, 1, 2, \ldots, 2^{\#\text{bits}} - 1$. Analog precoding can be designed based on the channel vector of cluster heads $\mathbf{h}_{U_H(c)}$ in the set U_H such that the element-wise angle differs slightly from the precoder quantized angle $\bar{\mathbf{p}}_c^{(F)}$ or $\bar{\mathbf{p}}_c^{(S)}$ for each cluster. That is, $\mathbf{h}_{U_H(c)}$ is orthogonal to $\bar{\mathbf{p}}_c^{(F)}$ or $\bar{\mathbf{p}}_c^{(S)}$ which gets maximized in the antenna array gain $|\mathbf{h}_{U_H(c)}\bar{\mathbf{p}}_c^{(F)}|^2$ or $|\mathbf{h}_{U_H(c)}\bar{\mathbf{p}}_c^{(S)}|^2$ respectively. The expression for the respective element-wise analog precoding vectors are

$$\bar{\mathbf{p}}_c^{(F)}(i) = \frac{1}{\sqrt{N_B}} e^{-j\frac{2\pi \hat{n}}{2^{\#\text{bits}}}} [11], \text{ and} \tag{17}$$

$$\bar{\mathbf{p}}_c^{(S)}(i) = \frac{1}{\sqrt{N_{\text{D}_c}}} e^{-j\frac{2\pi \hat{n}}{2^{\#\text{bits}}}} [11], \tag{18}$$

where $i = 1, 2, \ldots, N_B$ for a fully connected case, $i = (c-1)N_{\text{D}_c} + 1, (c-1)N_{\text{D}_c} + 2, \ldots, cN_{\text{D}_c}$ for sub-connected cases and $\hat{n}$ is the minimization function represented as

$$\hat{n} = \underset{n \in (n=0,1,2,\ldots,2^{\#\text{bits}}-1)}{\text{argmin}} |\angle(\mathbf{h}_{U_H(c)}(i)) - e^{-j\frac{2\pi \hat{n}}{2^{\#\text{bits}}}}| [11]. \tag{19}$$

The kth user in any cluster after obtaining the analog precoding vectors can now have an equivalent channel vector $\bar{\mathbf{h}}_k^H = \mathbf{h}_k^H \mathbf{P}$. A user $k \notin U_H$ is then grouped by a maximization function $\hat{c}$:

$$\hat{c} = \underset{n \in (0,1,2,\ldots,C)}{\text{argmax}} \frac{|\bar{\mathbf{h}}_k^H - \bar{\mathbf{h}}_{U_H(c)}^H|}{||\bar{\mathbf{h}}_k^H||_2 - ||\bar{\mathbf{h}}_{U_H(c)}^H||_2} [11], \tag{20}$$

where $\hat{c}$ is the beam or cluster that the equivalent channel correlation between the user u and any one of the cluster head U_H is highest.

The digital precoding can be derived using the low-complexity zero-forcing (ZF) precoding when the equivalent channel vector of the k_cth user given as $\bar{\mathbf{h}}_{k_c}$ is assumed to be the highest in the cluster. That is $\bar{\mathbf{h}}_{c,k} = \bar{\mathbf{h}}_{k_c}$, so that for all kth

users in each cluster $c = 1, 2, \ldots, C$ the equivalent channel matrix is $\bar{\mathbf{H}} = [\bar{\mathbf{h}}_{k_1}, \bar{\mathbf{h}}_{k_2}, \ldots, \bar{\mathbf{h}}_{k_C}]$. The digital precoding matrix can then be generated as

$$\bar{\mathbf{Q}} \triangleq [\bar{\mathbf{q}}_1, \bar{\mathbf{q}}_2, \ldots, \bar{\mathbf{q}}_C] = [\bar{\mathbf{H}}(\bar{\mathbf{H}}^H\bar{\mathbf{H}})^{-1}][11], \quad (21)$$

while the digital precoding vector is obtained by normalizing $\bar{\mathbf{q}}_c$:

$$\mathbf{q}_c = \frac{\bar{\mathbf{q}}_c}{||\mathbf{P}\bar{\mathbf{q}}_c||_2}[11], \quad (22)$$

Since the kth user is made the cluster head user reordering of the form earlier described is carried out:

$$||\bar{\mathbf{h}}_{c,1}^H\mathbf{q}_c||_2^2 \geq ||\bar{\mathbf{h}}_{c,2}^H\mathbf{q}_c||_2^2 \geq \ldots \geq ||\bar{\mathbf{h}}_{c,|U_c|}^H\mathbf{q}_c||_2^2[11]. \quad (23)$$

This marks the end of the user grouping and hybrid precoding design. A look at the joint power splitting and allocation problem from the ASR expression of (14) presents the challenge of coupling both power allocation and splitting factors for different users. This intractable problem receives an iterative algorithmic approach to a sub-optimal solution [11]. It follows a formulation as

$$\begin{aligned}
&\underset{(v_{c,k}),(\beta_{c,k})}{\text{argmax}} \sum_{c=1}^{C}\sum_{k=1}^{|U_i|} R_{c,k} \\
&s.t.\ C_1:\ v_{c,k} \geq 0,\ \forall\, c, k, \\
&\qquad C_2:\ \sum_{c=1}^{C}\sum_{k=1}^{|U_i|} v_{c,k} \leq V_{\text{th}}, \\
&\qquad C_3:\ R_{c,k} \geq R_{c,k}^{\min},\ \forall\, c, k, \\
&\qquad C_4:\ P_{c,k}^{\text{EH}} \geq P_{c,k}^{\min},\ \forall\, c, k[11],
\end{aligned} \quad (24)$$

where the power allocation to each user is constrained to be positive by C_1, the maximum total transmitted power at the BS is defined by the constraint C_2, each user data rate is constrained to a minimum value $R_{c,k}^{\min}$ by C_3, and the minimum energy harvesting constraint is defined by C_4. The maximization of the objective makes (24) a non-convex problem that can be optimized iteratively by employing the extended Sherman-Morrison-Woodbury formula [28], [29] for convexification as in [11] and [17]:

$$(\mathbf{A}+\mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{I}+\mathbf{CDA}^{-1}\mathbf{B})^{-1}\mathbf{CDA}^{-1}, \quad (25)$$

where for $\mathbf{A} = 1, \mathbf{C} = v_{c,k}, \mathbf{D} = ||\bar{\mathbf{h}}_{c,k}^H\mathbf{q}_c||_2^2$, and $\mathbf{B} \triangleq \zeta_{c,k}^{-1}$ (12) can be transformed as

$$(1+\text{SINR}_{c,k})^{-1} = 1 - \Big\{(\zeta_{c,k} + v_{c,k}||\bar{\mathbf{h}}_{c,k}^H\mathbf{q}_c||_2^2)^{-1} \times v_{c,k}||\bar{\mathbf{h}}_{c,k}^H\mathbf{q}_c||_2^2\Big\}, \quad (26)$$

where c and k remain the cluster and user index respectively. Estimate (11) as

$$\begin{aligned}
\tilde{\mathrm{y}}_{c,k} = \bar{\mathbf{h}}_{c,k}^H\mathbf{q}_c\sqrt{v_{c,k}}s_{c,k} + \bar{\mathbf{h}}_{c,k}^H\mathbf{q}_c\sum_{j=1}^{k-1}\sqrt{v_{c,j}}s_{c,j} \\
+ \bar{\mathbf{h}}_{c,k}^H\sum_{i\neq c}^{C}\sum_{j=1}^{|U_i|}\mathbf{q}_c\sqrt{v_{i,j}}s_{i,j} + n_{c,k} + \frac{u_{c,k}}{\sqrt{\beta_{c,k}}}[11].
\end{aligned} \quad (27)$$

This is necessary to aid the formulation of the detection problem which involves getting $s_{c,k}$ from $\tilde{\mathrm{y}}_{c,k}$ using minimum mean square error (MMSE) detection.

$$\varsigma_{c,k}^* = \underset{\varsigma_{c,k}}{\text{argmin}}\, E\{|s_{c,k} - \varsigma_{c,k}\tilde{\mathrm{y}}_{c,k}|^2\}[11]. \quad (28)$$

where $E\{|s_{c,k} - \varsigma_{c,k}\tilde{\mathrm{y}}_{c,k}|^2\} = \text{MSE}_{c,k}$ is the mean square error (MSE) and $\varsigma_{c,k}$ is the channel equalization coefficient that balances channel differences based on initial assumptions made. The MSE can be expressed with reduced complex addition:

$$\begin{aligned}
\text{MSE}_{c,k} &= E\{|s_{c,k} - \varsigma_{c,k}\tilde{\mathrm{y}}_{c,k}|^2\}, \\
&= E\{(s_{c,k})^2 - 2\text{Re}(s_{c,k}\varsigma_{c,k}\tilde{\mathrm{y}}_{c,k}) + (\varsigma_{c,k}\tilde{\mathrm{y}}_{c,k})^2\}, \\
&= E\{(s_{c,k})^2\} - 2\text{Re}E\{(s_{c,k}\varsigma_{c,k}\tilde{\mathrm{y}}_{c,k})\} \\
&\quad + E\{(\varsigma_{c,k}\tilde{\mathrm{y}}_{c,k})^2\},
\end{aligned} \quad (29)$$

Recall $E\{(s_{c,k})^2\} = 1$ for the desired signal can be used to simplify (29) assuming $E\{(s_{c,k}s_{i\neq c,j\neq k})\} = 0$, $E\{(n_{c,k})\} \equiv E\{(u_{c,k})\} = 0$, $E\{(u_{c,k})^2\} = \sigma_u^2$, $E\{(n_{c,k})^2\} = \sigma_n^2$, and compacting other terms in the form of $\zeta_{c,k}$ in (12) to obtain

$$\begin{aligned}
\text{MSE}_{c,k} = 1 - 2\text{Re}(\varsigma_{c,k}\bar{\mathbf{h}}_{c,k}^H\mathbf{q}_c\sqrt{v_{c,k}}) \\
+ |\varsigma_{c,k}|^2(||\bar{\mathbf{h}}_{c,k}^H\mathbf{q}_c||_2^2 v_{c,k} + \zeta_{c,k})[11].
\end{aligned} \quad (30)$$

To achieve (28), a partial derivative of (30) related to $\varsigma_{c,k}$ is obtained as:

$$\begin{aligned}
\frac{\partial \text{MSE}_{c,k}}{\partial \varsigma_{c,k}} = -2(\bar{\mathbf{h}}_{c,k}^H\mathbf{q}_c\sqrt{v_{c,k}}) \\
+ 2\varsigma_{c,k}(||\bar{\mathbf{h}}_{c,k}^H\mathbf{q}_c||_2^2 v_{c,k} + \zeta_{c,k}).
\end{aligned} \quad (31)$$

The optimum equalization co-efficient $\varsigma_{c,k}^*$, is obtained by finding the value of $\varsigma_{c,k}$ in RHS of (31) which satisfies $(\frac{\partial \text{MSE}_{c,k}}{\partial \varsigma_{c,k}}) = 0$. That is

$$\varsigma_{c,k}^* = (\bar{\mathbf{h}}_{c,k}^H\mathbf{q}_c\sqrt{v_{c,k}})(||\bar{\mathbf{h}}_{c,k}^H\mathbf{q}_c||_2^2 v_{c,k} + \zeta_{c,k})^{-1}[11]. \quad (32)$$

Substitute (32) into (30) to get the MMSE.

$$\begin{aligned}
\text{MSE}_{c,k}^* = 1 - 2\Big\{\frac{(\bar{\mathbf{h}}_{c,k}^H\mathbf{q}_c\sqrt{v_{c,k}})}{(||\bar{\mathbf{h}}_{c,k}^H\mathbf{q}_c||_2^2 v_{c,k} + \zeta_{c,k})}\Big\}(\bar{\mathbf{h}}_{c,k}^H\mathbf{q}_c\sqrt{v_{c,k}}) \\
+ \Big\{\Big\{\frac{(\bar{\mathbf{h}}_{c,k}^H\mathbf{q}_c\sqrt{v_{c,k}})}{(||\bar{\mathbf{h}}_{c,k}^H\mathbf{q}_c||_2^2 v_{c,k} + \zeta_{c,k})}\Big\}^2 \\
\times (||\bar{\mathbf{h}}_{c,k}^H\mathbf{q}_c||_2^2 v_{c,k} + \zeta_{c,k})\Big\}
\end{aligned}$$

$$= 1 - 2\frac{(\bar{\mathbf{h}}_{c,k}^H \mathbf{q}_c||_2^2 v_{c,k})}{(||\bar{\mathbf{h}}_{c,k}^H \mathbf{q}_c||_2^2 v_{c,k} + \zeta_{c,k})} + \frac{(\bar{\mathbf{h}}_{c,k}^H \mathbf{q}_c||_2^2 v_{c,k})}{(||\bar{\mathbf{h}}_{c,k}^H \mathbf{q}_c||_2^2 v_{c,k} + \zeta_{c,k})} = 1 - \frac{(||\bar{\mathbf{h}}_{c,k}^H \mathbf{q}_c||_2^2 v_{c,k})}{(||\bar{\mathbf{h}}_{c,k}^H \mathbf{q}_c||_2^2 v_{c,k} + \zeta_{c,k})}[11] \quad (33)$$

Observe that, by comparing (26) with (33) based on (28) we have the expressions for MMSE.

$$(1 + \text{SINR}_{c,k})^{-1} = \min_{\varsigma_{c,k}} E\{|s_{c,k} - \varsigma_{c,k}\tilde{y}_{c,k}|^2\}[11], \quad (34)$$

where

$$\min_{\varsigma_{c,k}} \text{MSE}_{c,k} = \text{MSE}_{c,k}^*. \quad (35)$$

To maximize ASR in (24) the achievable rate (AR) of the kth user in cluster c, that is, $R_{c,k}$ is maximized from (34)-(35):

$$(1 + \text{SINR}_{c,k})^{-1} = \min_{\varsigma_{c,k}} \text{MSE}_{c,k} \quad (36)$$

$$\frac{1}{(1 + \text{SINR}_{c,k})} = \min_{\varsigma_{c,k}} \text{MSE}_{c,k} \quad (37)$$

$$\frac{1}{\min_{\varsigma_{c,k}} \text{MSE}_{c,k}} = (1 + \text{SINR}_{c,k}) \quad (38)$$

$$(1 + \text{SINR}_{c,k}) = \max_{\varsigma_{c,k}} \text{MSE}_{c,k}^{-1} \quad (39)$$

The logarithm of (39) gives the AR in (13):

$$R_{c,k} \triangleq \log_2(1 + \text{SINR}_{c,k}) = \log_2 \max_{\varsigma_{c,k}} \text{MSE}_{c,k}^{-1} = \max_{\varsigma_{c,k}} (\log_2 \text{MSE}_{c,k}^{-1}) = \max_{\varsigma_{c,k}} (-\log_2 \text{MSE}_{c,k})[11] \quad (40)$$

A proposition is made at this point to remove the log function as follows [11].

Proposition 1: Given $f(a) = -\frac{ab}{\ln 2} + \log_2 a + \frac{1}{\ln 2}$, and a is a positive real number, then

$$\max_{a>0} f(a) = -\log_2 b, \quad (41)$$

where the optimal value of a is $a^* = \frac{1}{b}$. This allows for $R_{c,k}$ and the maximization term in (40) to be re-written as

$$R_{c,k} = \max_{\varsigma_{c,k}} \max_{a_{c,k}>0} (-\frac{a_{c,k}\text{MSE}_{c,k}}{\ln 2} + \log_2 a_{c,k} + \frac{1}{\ln 2})[11] \quad (42)$$

It follows that the optimization problem of (24) can be re-written as

$$\max_{\varsigma_{c,k}} \max_{a_{c,k}>0} \sum_{c=1}^{C} \sum_{k=1}^{|U_c|} (-\frac{a_{c,k}\text{MSE}_{c,k}}{\ln 2} + \log_2 a_{c,k}) \quad s.t. \quad C_1, C_2, C_3, C_4[11]. \quad (43)$$

The solution is preserved without the constant term in (42). The iterative optimization algorithm of (43) is carried out to optimize $v_{c,k}$, $\beta_{c,k}$, $\varsigma_{c,k}$, and $a_{c,k}$ disjointly. It begins with a given optimal power allocation $v_{c,k}^{\text{iter-1}}$ and the power-splitting factor $\beta_{c,k}^{\text{iter-1}}$ in the previous iteration, say iter-1, continues to find $\varsigma_{c,k}^{\text{iter}}$ and $a_{c,k}^{\text{iter}}$ in the next iteration based on (32) and (41):

$$\varsigma_{c,k}^{\text{iter}} = (\bar{\mathbf{h}}_{c,k}^H \mathbf{q}_c \sqrt{v_{c,k}^{\text{iter-1}}})(||\bar{\mathbf{h}}_{c,k}^H \mathbf{q}_c||_2^2 v_{c,k}^{\text{iter-1}} + \zeta_{c,k}^{\text{iter-1}})^{-1}[11], \quad (44)$$

where

$$\zeta_{c,k}^{\text{iter-1}} = ||\bar{\mathbf{h}}_{c,k}^H \mathbf{q}_i||_2^2 \sum_{j=1}^{k-1} v_{c,j}^{\text{iter-1}} + \left\{ \sum_{i\neq c}^{C} ||\bar{\mathbf{h}}_{c,k}^H \mathbf{q}_i||_2^2 \times \sum_{j=1}^{|U_i|} v_{i,j}^{\text{iter-1}} \right\} + \sigma_n^2 + \frac{\sigma_u^2}{\beta_{c,k}^{\text{iter-1}}}[11], \quad (45)$$

while

$$a_{c,k}^{\text{iter}} = \frac{1}{\text{MSE}_{c,k}^{*\text{iter}}}[11], \quad (46)$$

from (42), and from (33)

$$\text{MSE}_{c,k}^{*\text{iter}} = 1 - \frac{(||\bar{\mathbf{h}}_{c,k}^H \mathbf{q}_c||_2^2 v_{c,k}^{\text{iter-1}})}{(||\bar{\mathbf{h}}_{c,k}^H \mathbf{q}_c||_2^2 v_{c,k}^{\text{iter-1}} + \zeta_{c,k}^{\text{iter-1}})}[11]. \quad (47)$$

A simplified form of the optimization problem in (43) would then be to minimize the negative term in the objective function as

$$\min_{(v_{c,k}),(\beta_{c,k})} \sum_{c=1}^{C} \sum_{k=1}^{|U_c|} a_{c,k}^{\text{iter}} \text{MSE}_{c,k}^{\text{iter}} \quad s.t.\ C_1^{\text{iter}}: v_{c,k}^{\text{iter}} \geq 0,\ \forall\, c, k, \quad C_2^{\text{iter}}: \sum_{c=1}^{C} \sum_{k=1}^{|U_i|} v_{c,k}^{\text{iter}} \leq V_{\text{th}}, \quad C_3^{\text{iter}}: R_{c,k}^{\text{iter}} \geq R_{c,k}^{\min},\ \forall\, c, k, \quad C_4^{\text{iter}}: P_{c,k}^{\text{EH (iter)}} \geq P_{c,k}^{\min},\ \forall\, c, k, [11] \quad (48)$$

where

$$\text{MSE}_{c,k}^{\text{iter}} = 1 - 2\text{Re}(\varsigma_{c,k}^{\text{iter}} \bar{\mathbf{h}}_{c,k}^H \mathbf{q}_c \sqrt{v_{c,k}^{\text{iter}}}) + |\varsigma_{c,k}^{\text{iter}}|^2 (||\bar{\mathbf{h}}_{c,k}^H \mathbf{q}_c||_2^2 v_{c,k}^{\text{iter}} + \zeta_{c,k}^{\text{iter}})[11]. \quad (49)$$

The objective in (48) is made convex by adding the constrain that $\tau_{c,k}^{\text{iter}} \geq \frac{1}{\beta_{c,k}^{\text{iter}}}$ in the expanded form of $\zeta_{c,k}^{\text{iter}}$ in (49):

$$\tilde{\text{MSE}}_{c,k}^{\text{iter}} = 1 - 2\text{Re}(\varsigma_{c,k}^{\text{iter}} \bar{\mathbf{h}}_{c,k}^H \mathbf{q}_c \sqrt{v_{c,k}^{\text{iter}}}) + |\varsigma_{c,k}^{\text{iter}}|^2 \left(||\bar{\mathbf{h}}_{c,k}^H \mathbf{q}_c||_2^2 \sum_{j=1}^{k} v_{c,k}^{\text{iter}} + \sum_{i\neq c}^{C} ||\bar{\mathbf{h}}_{c,k}^H \mathbf{q}_i||_2^2 \sum_{j=1}^{|U_i|} v_{i,j}^{\text{iter}} + \sigma_n^2 + \sigma_u^2 \tau_{c,k}^{\text{iter}} \right)[11], \quad (50)$$

to obtain

$$\min_{(v_{c,k}),(\beta_{c,k})} \sum_{c=1}^{C}\sum_{k=1}^{|U_c|} a_{c,k}^{\text{iter}} \tilde{\text{MSE}}_{c,k}^{\text{iter}}[11], \tag{51}$$

with an added constraint

$$C_5^{\text{iter}}: \tau_{c,k}^{\text{iter}} \geq \frac{1}{\beta_{c,k}^{\text{iter}}}, \ \forall\, c,k[11]. \tag{52}$$

The third constraint in (48) is also made convex by the following transformation.

$$\begin{aligned}\tilde{C}_3^{\text{iter}}: & ||\bar{\mathbf{h}}_{c,k}^H \mathbf{q}_c||_2^2 v_{c,k}^{\text{iter}} - \omega_{c,k}||\bar{\mathbf{h}}_{c,k}^H \mathbf{q}_c||_2^2 \sum_{j=1}^{k-1} v_{c,k}^{\text{iter}} \\ & - \omega_{c,k}\sum_{i\neq c}^{C}||\bar{\mathbf{h}}_{c,k}^H \mathbf{q}_i||_2^2 \sum_{j=1}^{|U_i|} v_{i,j}^{\text{iter}} - \omega_{c,k}\sigma_n^2 \\ & - \omega_{c,k}\sigma_u^2 \tau_{c,k}^{\text{iter}} \geq 0, \\ : & ||\bar{\mathbf{h}}_{c,k}^H \mathbf{q}_c||_2^2 v_{c,k}^{\text{iter}} - \omega_{c,k}||\bar{\mathbf{h}}_{c,k}^H \mathbf{q}_c||_2^2 \sum_{j=1}^{k-1} v_{c,k}^{\text{iter}} \\ & - \omega_{c,k}\sum_{i\neq c}^{C}||\bar{\mathbf{h}}_{c,k}^H \mathbf{q}_i||_2^2 \sum_{j=1}^{|U_i|} v_{i,j}^{\text{iter}} \\ & - \omega_{c,k}\sigma_u^2 \tau_{c,k}^{\text{iter}} \geq \omega_{c,k}\sigma_n^2[11], \end{aligned} \tag{53}$$

where $\omega_{c,k}$ is a function of the minimum achievable rate $R_{c,k}^{\min}$ weighted against all interference and noise terms by the expression $\omega_{c,k} = 2^{R_{c,k}^{\min}} - 1$ to ensure the desired signal power is always dominant. The multivariable coupled nature of the fourth and fifth constraints of (48) makes a non-convex mixture that can be addressed by the introduction of another variable $\mu_{c,k}^{\text{iter}}$:

$$C_6^{\text{iter}}: \mu_{c,k}^{\text{iter}} \geq \frac{P_{c,k}^{\min}}{\eta(1-\beta_{c,k}^{\text{iter}})}, \ \forall\, c,k[11]. \tag{54}$$

This allows C_4^{iter} in (49) to be rewritten convex-wise from (9) as

$$\begin{aligned}\tilde{C}_4^{\text{iter}}: & \ \eta(1-\beta_{c,k}^{\text{iter}})\Big(\sum_{i=1}^{C}\sum_{j=1}^{|U_i|}||\bar{\mathbf{h}}_{c,k}^H \mathbf{q}_i||_2^2 v_{i,j} + \sigma_n^2\Big) \\ & \geq \eta(1-\beta_{c,k}^{\text{iter}})\mu_{c,k}^{\text{iter}}, \ \forall\, c,k, \\ & : \Big(\sum_{i=1}^{C}\sum_{j=1}^{|U_i|}||\bar{\mathbf{h}}_{c,k}^H \mathbf{q}_i||_2^2 v_{i,j} + \sigma_n^2\Big) \geq \mu_{c,k}^{\text{iter}} \ \forall\, c,k[11]. \end{aligned} \tag{55}$$

The constraints C_5^{iter} in (52) and C_6^{iter} in (54) are non-convex. By employing Schur's complement lemma suggested in [1] through matrix representation applied in [11], both constraints are transformed into

$$\tilde{C}_5^{\text{iter}}: \begin{bmatrix} \beta_{c,k}^{\text{iter}} & 1 \\ 1 & \tau_{c,k}^{\text{iter}} \end{bmatrix} \geq \mathbf{0}, \ \forall\, c,k. \tag{56}$$

$$\tilde{C}_6^{\text{iter}}: \begin{bmatrix} \mu_{c,k}^{\text{iter}} & \sqrt{\frac{P_{c,k}^{\min}}{\eta}} \\ \sqrt{\frac{P_{c,k}^{\min}}{\eta}} & 1-\beta_{c,k}^{\text{iter}} \end{bmatrix} \geq \mathbf{0}, \ \forall\, c,k. \tag{57}$$

Finally, the optimization problem of (48) is fully convexified into

$$\begin{aligned}\min_{(v_{c,k}),(\beta_{c,k})} & \sum_{c=1}^{C}\sum_{k=1}^{|U_c|} a_{c,k}^{\text{iter}} \tilde{\text{MSE}}_{c,k}^{\text{iter}} \\ & s.t.\ C_1, C_2, \tilde{C}_3^{\text{iter}}, \tilde{C}_4^{\text{iter}}, \tilde{C}_5^{\text{iter}}, \tilde{C}_6^{\text{iter}}[11]. \end{aligned} \tag{58}$$

This standard convex optimization problem is resolved using numerical convex program solvers [30]. The entire process is summarized in Algorithm 1. One out of the multiple antennas at the user is selected in step 3 and indexed for all users in step 4 at random (R), by maximum norm average (MNA), or a combination of both. Channels are generated for all users assuming a fully digital (FD) or HP system followed by the design of an analog precoder (AP) and digital precoder (DP) in step 5. The equivalent channel vector ($\bar{\mathbf{h}}_{c,k}^H$) and power (V_{th}) allocation for each user ($v_{c,k}^0$) are then computed to obtain the minimum achievable rate ($R_{c,k}^{\min}$) from the user achievable rate ($R_{c,k}^0$) based on the power splitting factor ($\beta_{c,k}^0$) in steps 6-9. In step 10, zero-forcing (ZF) AP and DP are designed wherein the power allocation and the equivalent channel vector per user are computed for hybrid precoding (HP), HP with NOMA (HP-NOMA), and HP with OMA (HP-OMA) accordingly. Step 10 covers what appears to be the RF based user group thresholding (i.e. $\delta = \frac{1}{N_{\text{RF}}}$) of the cluster head selection (CHS) algorithm in [11] hinted by [31]. In step 13, the channel overhead estimate ($\varsigma_{c,k}^0$) due to quantization that minimizes the MSE with the corresponding inverse ($a_{c,k}^0$) is initialized for the joint optimization of the power allocation ($v_{c,k}^{\text{iter}}$) and power spitting factor ($\beta_{c,k}^{\text{iter}}$) in step 15. Spectrum efficiency (SE) and energy efficiency (EE) are expected outputs for performance evaluation under full and sub-connected HP. Given the ASR (R_{sum}), total transmit power ($V_B = \sum_{c=1}^{C}\sum_{k=1}^{|U_c|} v_{c,k}$), number of RF chains ($N_{\text{RF}}$), power consumed by each RF chain (V_{RF}), number of phase shifters (N_{PS}), power consumed by each phase shifter (V_{PS}), and baseband power consumption (V_{BB}), the energy efficiency in bps/Hz/W is computed as in [11]:

$$\text{EE} = \frac{R_{\text{sum}}}{V_B + N_{\text{RF}}V_{\text{RF}} + N_{\text{PS}}V_{\text{PS}} + V_{\text{BB}}} \tag{59}$$

V. SIMULATION RESULTS AND DISCUSSION

In this section, MATLAB R2022a is used for the simulation. The results have been presented in the form of tables and graphs. The spectral efficiency (SE) and energy efficiency (EE) performance at various signal-to-noise ratios is evaluated. The parameters used for the computer simulations are similar to those in [11] and presented in Table 1.

A. RESULTS

To adaptively improve the system performance in [11], we conducted simulation test on sample values of the input's

Algorithm 1 Proposed Modified HP Design With Joint PA and PS Optimization for ID and EH

Input: $N_B, N_M, N_{\text{RF}}/N_c, K, L, \lambda, V_{\text{th}}, \eta, P_{c,k}^{\min}$, bit, cvx_ iter, N_iter, δ
1: **for** SNR$= -20, -15, \ldots, 15, 20$ dB
2: **while** $i \leq$ N_iter
3: Generate $\mathbf{H}_c^k \in \mathbb{C}^{N_B \times N_M \times k}$ for $k = 1, 2, \ldots, K$ users by (2)
4: Reduce $\mathbf{H}_c^k$ to $\mathbf{h}_{c,k}^H$ by random/MNA single (m) antenna selection in (3)
5: Initialize $\mathbf{P}$(AP), $\mathbf{Q}$(DP), $v_{c,k}^0$, and $\bar{\mathbf{h}}_{c,k}$ for FD
6: **for** $k = 1, 2, \ldots, K$
7: Generate $R_{c,k}^0$ and $\beta_{c,k}^0$
8: Calculate $R_{c,k}^{\min}$
9: **end for** (in 6)
10: Design AP (15)-(18), ZF DP (21)/(22), and get $v_{c,k}^0$, $\mathbf{h}_{c,k}^H$ for HP, HP-NOMA, and HP-OMA respectively
11: **for** iter $= 1, 2, \ldots,$ cvx_iter
12: **for** $k = 1, 2, \ldots, K$ **do**
13: Initialize $\varsigma_{c,k}^0 = \varsigma_{c,k}^{\text{iter-1}}$, $\text{MSE}_{c,k}^{*0} = \text{MSE}_{c,k}^{*\text{iter-1}}$, and $a_{c,k}^0 = a_{c,k}^{\text{iter-1}}$ by (44), (47), & (46)
14: **end for** (in 12)
15: Optimize $v_{c,k}^{\text{iter}} \leftarrow v_{c,k}^{\text{iter}}$ and $\beta_{c,k}^{\text{iter}} \leftarrow \beta_{c,k}^{\text{iter}}$ by (58)
16: Calculate ASR with the objective in (42)/(43) at feasible points
17: **end for** (in 11)
Output: SE (ASR)
18: $i = i + 1$
19: **end while** (in 2)
Output: EE
20: **end for** (in 1)

TABLE 1. System parameters.

Parameter	Value
System bandwidth (B_w)	1 Hz (13)
# of antennas ($N_B = N_M$)	64
# of RF chain ($N_{\text{RF}} = C$)	2 and 4
# of users (K)	3,4,5,6, and 8
# of multipath ($L_c = 3$)	l_1 $\mathcal{CN}(0,1)$: l_2/l_3 $\mathcal{CN}(0,0.1)$
Azimuth angles ($\theta_{c,k}^{(l)}, \phi_{c,k}^{(l)}$)	$\sim \mathcal{U}(-\pi, \pi)$
# of bits (#bits)	4
Maximum transmit power (V_{th})	30 mW
Minimum each user AR ($R_{c,k}^{\min}$)	0.1 min$[R_{c,1}^{0(\text{FD-ZF})}, \ldots, R_{c,k}^{0(\text{FD-ZF})}]$
Minimum each user EH ($P_{c,k}^{\min}$)	0.1 mW
V_{RF}	300 mW
V_{PS}	40 mW
V_{BB}	200 mW
N_{PS}	$N_{\text{PS}}^{(F)} = N_B N_{\text{RF}}, N_{\text{PS}}^{(S)} = N_B$
N_iter	10
cvx_iter	20
SNR	$[-20, 20]$]
η	0.9

initial correlation threshold ($\delta = 0.3$) for CHS within reasonable time. This led to the development of approximate run time values in Table 2 coupled with several graphical

TABLE 2. Approximate execution time (seconds) for system loading under specified schemes.

Antenna index (m) type/ δ	$N_{\text{RF}}: 2, K: 3,$ under-load	$N_{\text{RF}}: 2, K: 4,$ full-load	$N_{\text{RF}}: 2, K: 5,$ over-load	$N_{\text{RF}}: 4, K: 6,$ under-load	$N_{\text{RF}}: 4, K: 8,$ full-load
[11]	30529	122420	550726	94621	779536
(R_{HH})	28748	82977	390964	73866	455277
(K, N_{RF})	32312	125064	520416	131823	854871
Rand. (R)	36162	111547	595134	72556	864648
MNA	33826	107391	593690	78123	736828
R_MNA	32510	130439	563726	71997	733387
R/(R_{HH})	29347	95063	394365	69831	541363
R/(K, N_{RF})	43500	172729	705472	97445	812419

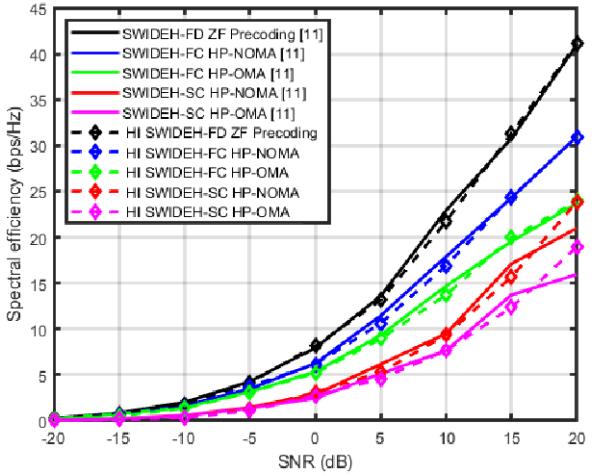

FIGURE 3. SE Vs SNR comparison between existing threshold $\delta(c)$ [11] and sample threshold $\delta(N_{\text{RF}})$ for proposed HI based system ($N_{\text{RF}} = 4, K = 6$).

plots for comparison. The first sample considered making the threshold a function of the number of RF chains such that $\delta(N_{\text{RF}}) = 1/N_{\text{RF}}$. This represents a difference of 0.05 from the original $\delta(c) = c$, where c is a constant (0.3) and N_{RF} is 4. It has a significant effect as seen in the SE plot of Fig. 3 without a noticeable change in the EE plot of Fig.4. Disparities are better noticed in the 10 to 20 dB SNR range of Fig. 3, therefore, we concentrate on [5, 20] SNR. The change in the SE value of the sub-connected (SC) system configuration is more visible for both NOMA and OMA than the fully-connected (FC) type. The fully digital (FD) configuration is included to serve as a reference for performance evaluation. Other samples of δ were tried based on the perceived relationship between N_{RF}, K, and $\mathbf{h}_{c,k}$. The R_{HH} represents the case where δ is chosen as the minimum equivalent channel correlation between K users (that is, $\delta(R_{HH}) = \min R_{HH}$). The K, N_{RF} represents

$$\delta(K, N_{\text{RF}}) = \Gamma - \varpi \tag{60}$$

where $\Gamma = \frac{(K+N_{\text{RF}})}{(K \times N_{\text{RF}})}$. and $\varpi = \frac{(1-\Gamma)}{10}$. Another significant yet positive impact is observed in the SE plot when this derived value of $\delta(K, N_{\text{RF}})$ is compared with existing work $\delta(c)$ shown in Fig. 5 which supersedes the performance

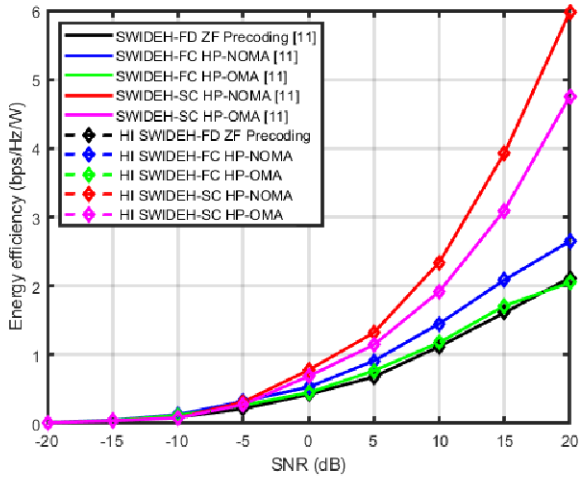

FIGURE 4. EE Vs SNR comparison between existing threshold $\delta(c)$ [11] and sample threshold $\delta(N_{RF})$ for proposed HI based system ($N_{RF} = 4, K = 6$).

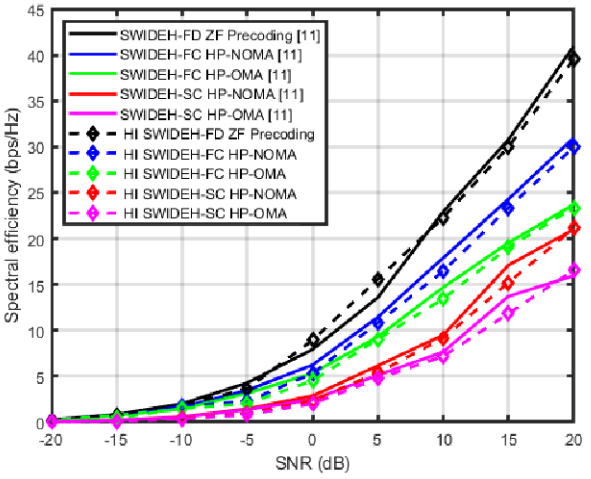

FIGURE 6. SE Vs SNR comparison between existing threshold $\delta(c)$ [11] and proposed threshold $\delta(R_{HH})$ for HI based system ($N_{RF} = 4, K = 6$).

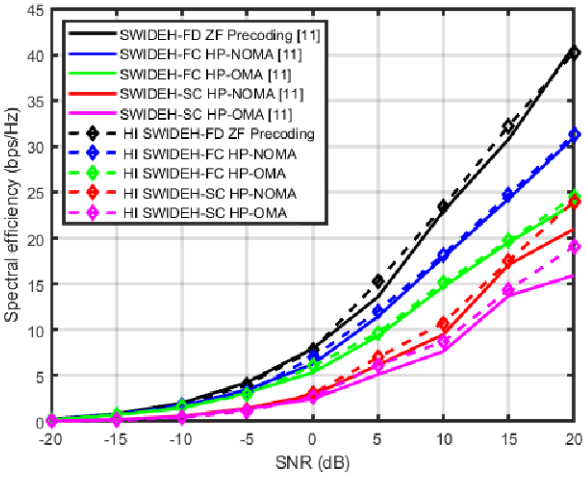

FIGURE 5. SE Vs SNR comparison between existing threshold $\delta(c)$ [11] and proposed threshold $\delta(K, N_{RF})$ for HI based system ($N_{RF} = 4, K = 6$).

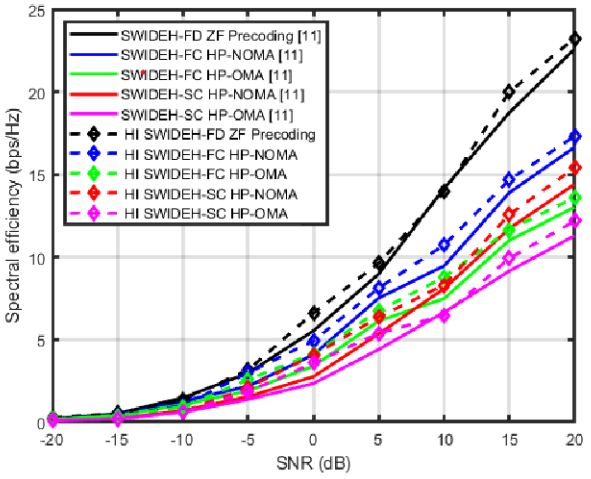

FIGURE 7. SE Vs SNR comparison between existing threshold $\delta(c)$ [11] and proposed threshold $\delta(K, N_{RF})$ for HI based system ($N_{RF} = 2, K = 3$).

of the proposed $\delta(R_{HH})$ in Fig. 6. While Figs. 5 and 6 are plotted for $N_{RF} = 4$, $K = 6$, a similar performance is noticed with the second proposed $\delta(K, N_{RF})$ in Figs. 7 and 8 for most of the SNR range when $N_{RF} = 2$, $K = 3$. This provides some level of consistent improvement compared to that of the first proposed $\delta(R_{HH})$ for the two underload cases tested. The trade-off is in the increased evaluation time for higher system loading relative to the number of RF chains.

At $N_{RF} = 2$, in the full-load ($K = 4$) and overload ($K = 5$) cases, the SEs of the proposed $1 - \delta(R_{HH})$ and $2 - \delta(K, N_{RF})$ are compared in Figs. 9 and 10 per system loading condition. The second proposed $\delta(K, N_{RF})$ remains a better choice model. Generally, it is observed that the SE remains somewhat unchanged at specific N_{RF} and any system loading condition. The FC/SC NOMA and FC/SC OMA curve parallel each other closely with increasing SNR thereby giving both FC and SC NOMA the needed advantage over their OMA counterparts. Specifically, the SE value of SC NOMA will eventually be greater than that of the FC OMA within smaller SNR as the number of users increases. The early interception between SE values of SC NOMA and FC OMA before exceeding the FC OMA added to the deviations from the objective of achieving a near FD SE performance arising from intra-cluster interference and SIC error propagation [31] can be seen as a function of the radio frequency resource N_{RF} and the number of users K.

The EE plot when $N_{RF} = 2, K = 3$ for threshold values $\delta(c)$ [11], $1 - \delta(R_{HH})$ and $2 - \delta(K, N_{RF})$ is shown in Fig. 11 while that of when $N_{RF} = 4, K = 6$ is shown in Fig. 12. The plots further validate the motivation to use the second proposed $\delta(K, N_{RF})$ as the best and preferred adaptive model for the initial correlation threshold in the proposed system. In addition, justification is provided by the performance of the SE and EE plots comparing the two proposed thresholds for the second full-load case ($N_{RF} = 4, K = 8$) of

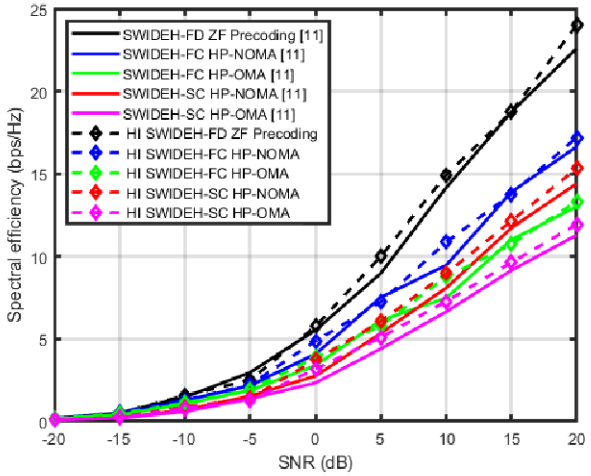

FIGURE 8. SE Vs SNR comparison between existing threshold $\delta(c)$ [11] and proposed threshold $\delta(R_{HH})$ for HI based system ($N_{\mathrm{RF}} = 2, K = 3$).

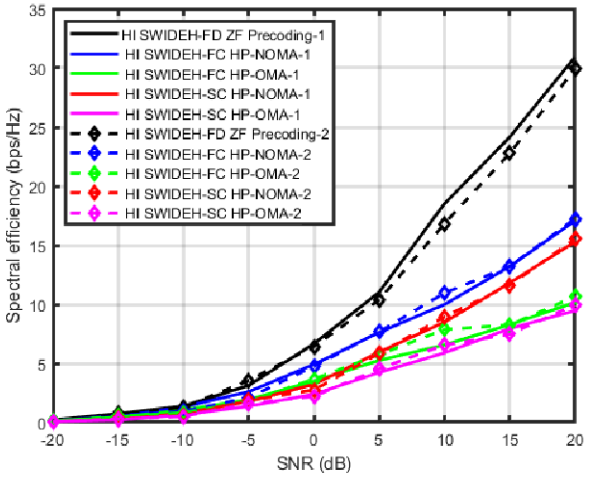

FIGURE 9. SE Vs SNR comparison between proposed threshold $1 - \delta(R_{HH})$ and $2 - \delta(K, N_{\mathrm{RF}})$ for HI based system ($N_{\mathrm{RF}} = 2, K = 4$).

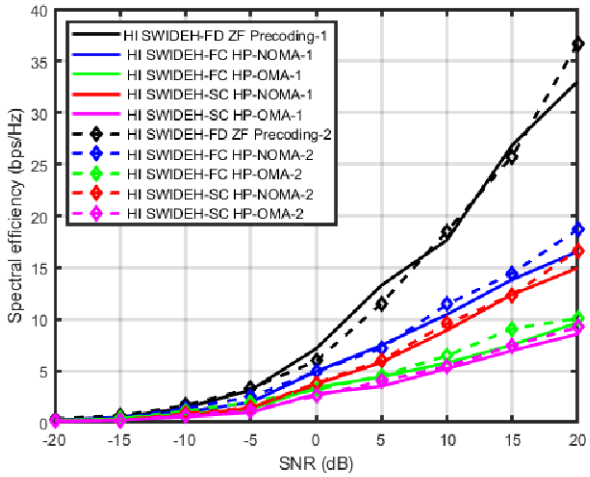

FIGURE 10. SE Vs SNR comparison between proposed threshold $1 - \delta(R_{HH})$ and $2 - \delta(K, N_{\mathrm{RF}})$ for HI based system ($N_{\mathrm{RF}} = 2, K = 5$).

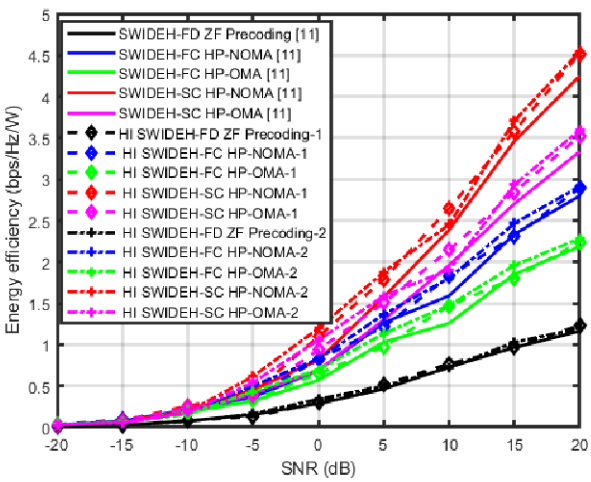

FIGURE 11. EE Vs SNR comparison between existing threshold,proposed threshold $1 - \delta(R_{HH})$ and $2 - \delta(K, N_{\mathrm{RF}})$ for HI based system ($N_{\mathrm{RF}} = 2, K = 3$).

Figs. 13 and 14, respectively. The SE plot reveals a similar trend where the SC NOMA exceed FC OMA at a lower SNR value as the number of users increases. However, the EE plot also reveals the limit of FC OMA below the reference FD system performance. It is observed that the second proposed threshold value of $\delta(K, N_{\mathrm{RF}})$ may not be as effective as the first proposed threshold value of $\delta(R_{HH})$ at some higher SNR points beyond 15 dB.

The assumption that the channel matrix for each user is known provides room for proposing the user channel matrix to vector selection schemes based on random/rand. (R) indexed (m) user vector, maximum norm average (MNA) for all such indexed (m) users per time, or a combination of the two (R_MNA). Figs. 15 and 16 show the SE and EE performance of the three schemes (S1, S2, and S3, respectively) compared with the reference paper [11]. The combination policy for the R_MNA scheme is to use R for $K > 2N_{\mathrm{RF}}$ and MNA, otherwise. The SNR range of focus remains [5], [20] dB. Both S2-MNA and S3-R_MNA are similar but have smaller run-time to implement than the S1-R (see Table 2). However, the gains of the R scheme are seen in the SE plot of Fig. 17 which shows SC NOMA can match and possibly outperform FC OMA at 20 dB SNR and beyond. In Fig. 16, the EE plot of the R scheme is better than the other schemes, especially for the SC type of systems. The R scheme provides a significant EE to SE tradeoff beyond [11] than the other proposed schemes. For that reason, the energy-efficient R scheme is selected for modified HP despite its long run time.

So far, the modified user group and hybrid precoding for information decoding and energy harvesting in hardware-impaired point-to-point mm-wave massive MIMO NOMA is implemented by combining the proposed adaptive initial thresholding $\delta(K, N_{\mathrm{RF}})$ and the matrix to vector pilot signaling R scheme with zero-forcing (ZF) digital precoding. The SE and EE plots of the combination feature C-R/(K, N_{RF}) in contrast with $\delta(c)$ [11] as shown in Figs. 17 and 18. Only the

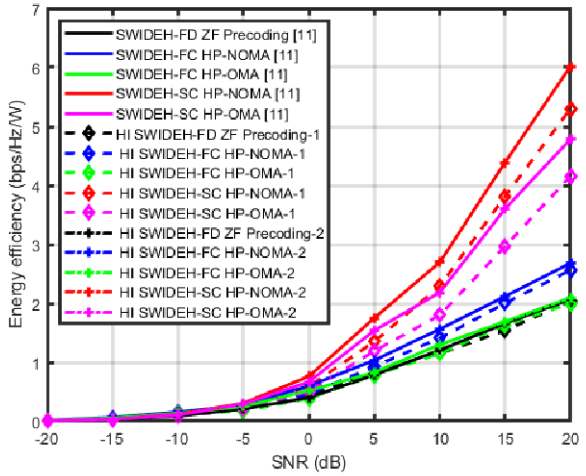

FIGURE 12. EE Vs SNR comparison between existing threshold,proposed threshold $1-\delta(R_{HH})$ and $2-\delta(K, N_{RF})$ for HI based system ($N_{RF}=4, K=6$).

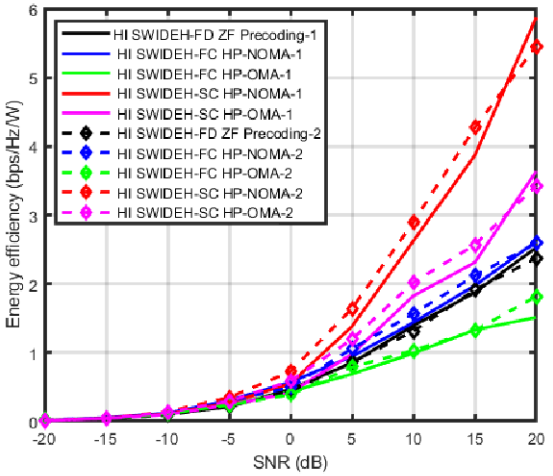

FIGURE 14. EE Vs SNR comparison between proposed threshold $1-\delta(R_{HH})$ and $2-\delta(K, N_{RF})$ for HI based system ($N_{RF}=4, K=8$).

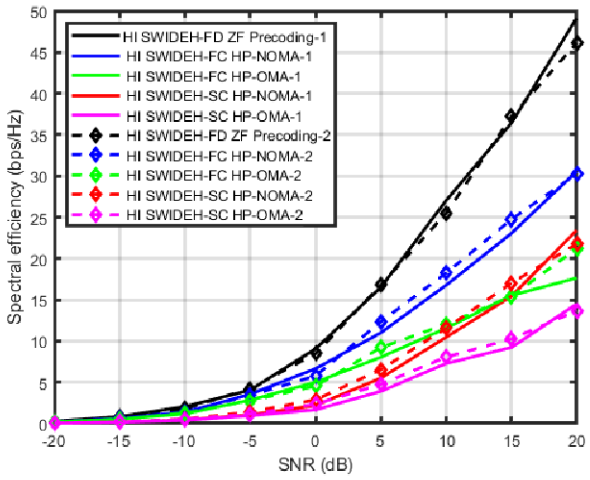

FIGURE 13. SE Vs SNR comparison between proposed threshold $1-\delta(R_{HH})$ and $2-\delta(K, N_{RF})$ for HI based system ($N_{RF}=4, K=8$).

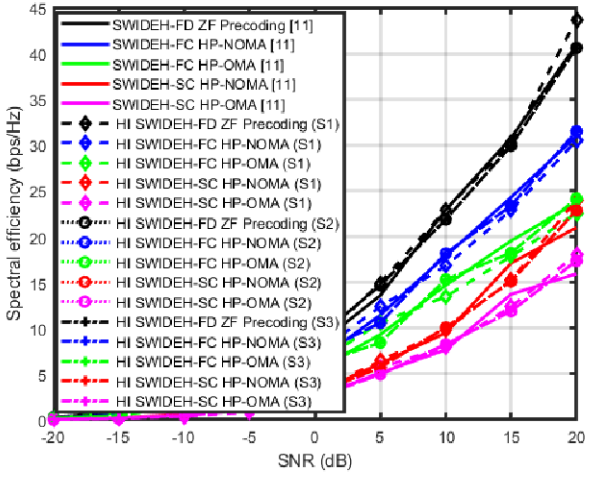

FIGURE 15. SE Vs SNR performance between existing multiple user single antenna channel vector [11] and proposed multiple users multiple antenna equivalent channel vector selection schemes (S1-R, S2-MNA, and S3-R_MNA) at $N_{RF}=4$ and $K=6$.

SC systems appear to have significantly improved at 10 and 20 dB SNR in the SE plot. In the 15-20 dB SNR region of the EE plot, the FC system performance relative to the FD system is below expectation. Hence, the need for the proposed regularized ZF (RZF). The proposed RZF is similar to [32] but differs by the regularizing factor such that from (21)

$$\bar{\mathbf{Q}} = (\bar{\mathbf{H}}(\bar{\mathbf{H}}^H\bar{\mathbf{H}})^{-1} + \xi\mathbf{I}), \tag{61}$$

where $\xi = \sqrt{(\text{diag}_{1\le j\le n}\sum_{i=1}^{m}\bar{h}_{(i,j)}^H)}$, m is the cardinality of cluster heads in the set U_H, and n is the number of users K. ξ is simply the square root of the resultant diagonalization after a column sum of $\bar{\mathbf{H}}^H$ is carried out. This concludes the proposed modification PM-C/RZF compared with $\delta(c)$ [11] in the SE and EE plots of Figs. 19 and 20. Generally, in the SE plot, the RZF is most effective at 5 dB SNR and below while the ZF performance is appreciable above 5 dB SNR. In the EE plot, both NOMA and OMA curves perform well above the set limit at the expense of lowered energy efficiency values for most of the SNR range. Especially at [0, 20] SNR.

The normalized mean square error (NMSE) of signal detection relative to the channel estimate for the proposed methods along the best-performing in each of the three (3) stages: 2-$\delta(K, N_{RF})$, C-R/(K, N_{RF}), and PM-C/RZF are plotted in Figs. 21-23. It is computed with (47) and we observed that the accuracy is between 0.01 and 10 which degrades with increasing SNR when compared with the FD system.

B. DISCUSSION OF RESULTS

This work assumes multiple antennas at both BS and MS operating UL channel estimation for mm-wave massive MIMO OMA with known channel estimates and DL mm-wave massive MIMO NOMA where user grouping, hybrid precoding, and power allocation for SWIDEH are conceived as hardware-impaired. The performance is

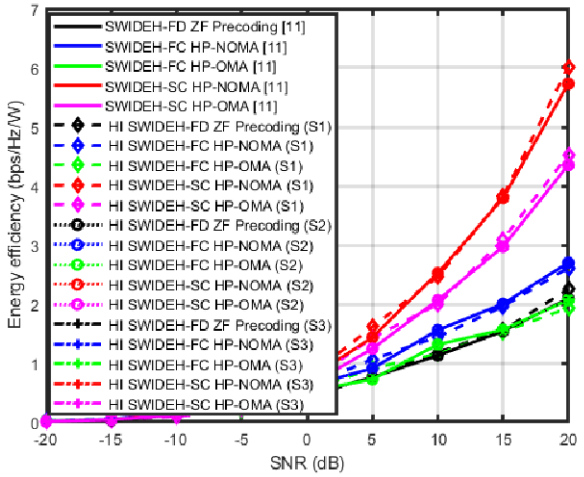

FIGURE 16. EE Vs SNR performance between existing multiple user single antenna channel vector [11] and proposed multiple users multiple antenna equivalent channel vector selection schemes (S1-R, S2-MNA, and S3-R_MNA) at $N_{RF} = 4$ and $K = 6$.

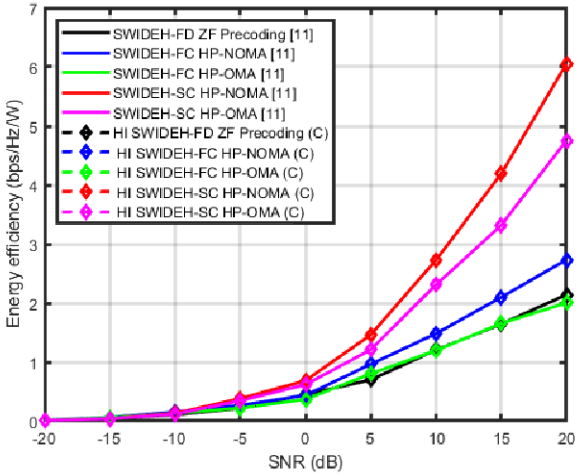

FIGURE 18. EE Vs SNR comparison between existing threshold $\delta(c)$ [11] and proposed method combination C-R/(K, N_{RF}) for HI based system ($N_{RF} = 4, K = 6$).

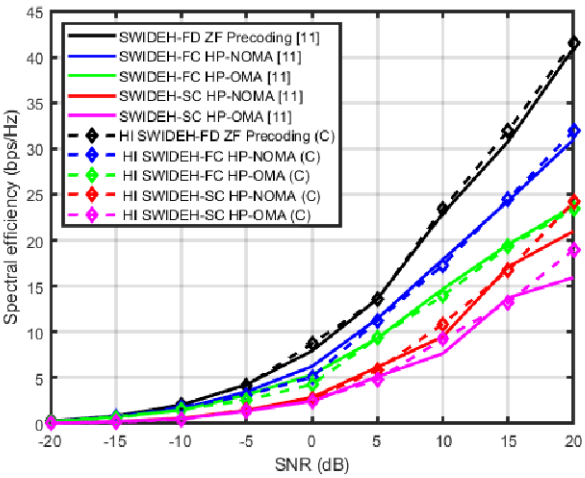

FIGURE 17. SE Vs SNR comparison between existing threshold $\delta(c)$ [11] and proposed method combination C-R/(K, N_{RF}) for HI based system ($N_{RF} = 4, K = 6$).

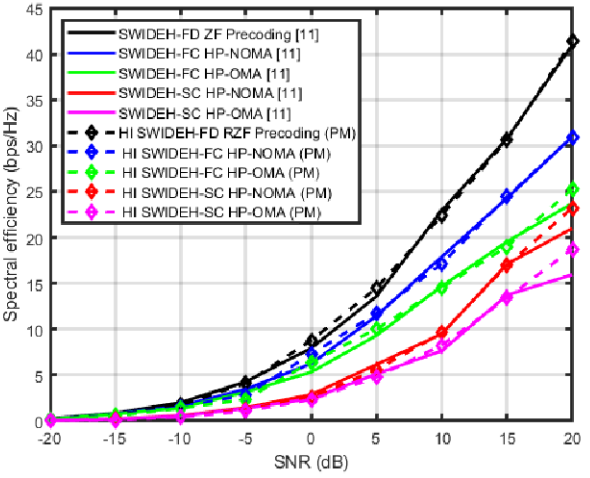

FIGURE 19. SE Vs SNR comparison between existing threshold $\delta(c)$ [11] and proposed method PM/C-RZF for HI based system ($N_{RF} = 4, K = 6$).

benchmarked with a fully digital (FD) system in [11] compared to DL mm-wave massive MIMO OMA forming the basis on which further development is explored. The results support the claim that the capacity of mm-wave massive MIMO with NOMA outperforms that of OMA at high SNR, especially under a cell-edge scenario [33]. Another possible supported claim is the limiting power of co-channel interferences from neighbouring users on the SE performance that is evident at SNRs where $\delta(R_{HH})$ is better than $\delta(K, N_{\mathrm{RF}})$ because it prioritizes the channel correlation above the relationship between the number of users and the radio frequency resource. However, the greater effect is short lived for $N_{\mathrm{RF}} = 2$ based on the simulation results. The SE result of the proposed initial threshold value (δ) compared to existing work [11] shows that the sensitivity of the sub-connected (SC) system (NOMA and OMA) may be higher than that of the fully-connected (FC) type with increasing SNR. This can also be attributed to the factored number of users and radio frequency resources as the basis for thresholding.

The random (R) antenna selection scheme shows consistent improvement in SE at the cost of the longest computer simulation time for initial value $\xi = 0$ (ZF) in (61). The weakness of the ZF precoder ($\xi = 0$) is revealed by cross-over curves at high SNR point(s) thereby showing the sensitivity of the hybrid precoder in meeting design specifications. The combination of a random (R) antenna selection scheme and any of the adaptive initial thresholds $\delta(R_{HH})$ or $\delta(K, N_{\mathrm{RF}})$ reveals the superiority of $\delta(K, N_{\mathrm{RF}})$ further, and better improvement in SE/EE for fully connected HP systems. The choice of N_{RF} and K values are inspired by [34] such that the system can be underloaded, fully loaded, and overloaded. An N_{RF} to K ratio of 2:6 (not included) would be four (4) times computation intensive than 2:5, the results are consistent with the design requirements of

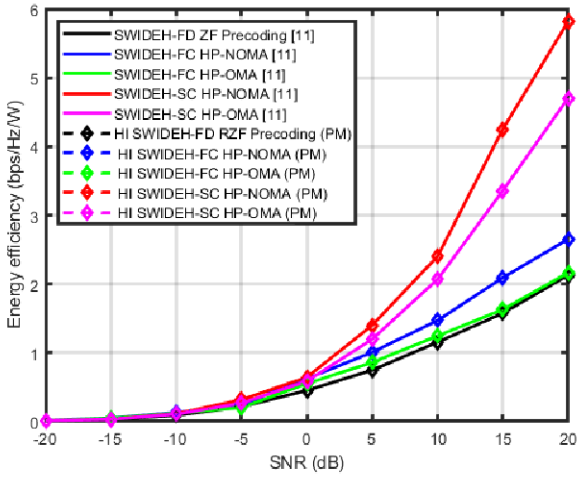

FIGURE 20. EE Vs SNR comparison between existing threshold $\delta(c)$ [11] and proposed method PM/C-RZF for HI based system ($N_{RF} = 4, K = 6$).

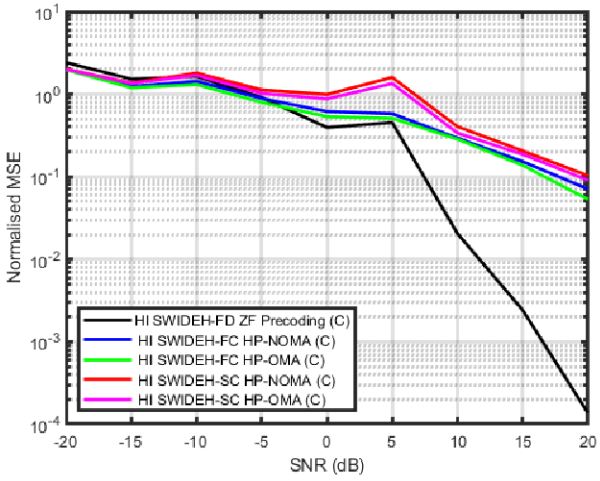

FIGURE 22. NMSE Vs SNR of best stage 2 method C-R/(K, N_{RF}) for HI based system ($N_{RF} = 4, K = 6$).

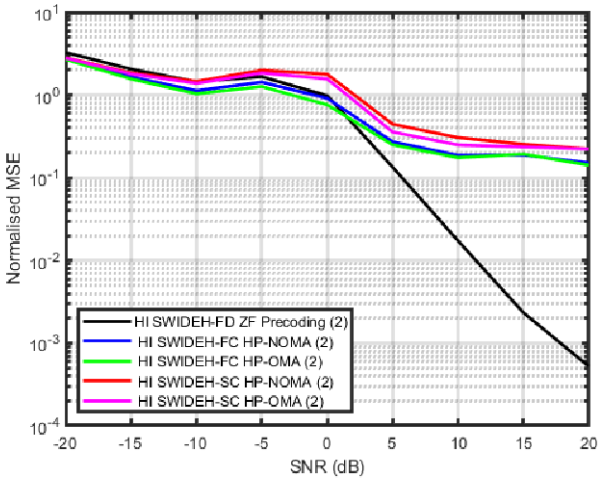

FIGURE 21. NMSE Vs SNR of best stage 1 method (K, N_{RF}) for HI based system ($N_{RF} = 4, K = 6$).

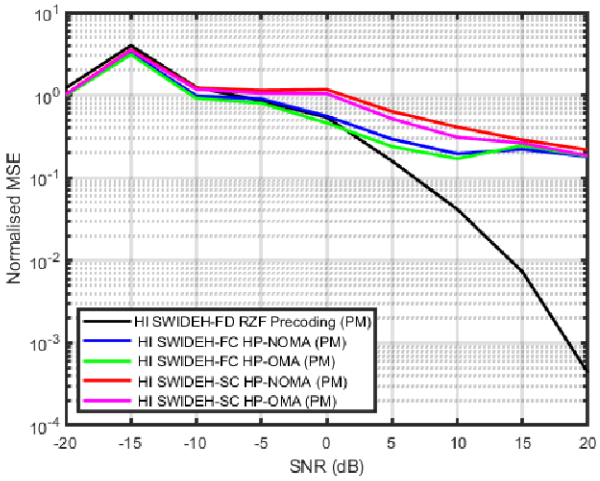

FIGURE 23. NMSE Vs SNR of best stage 3 method PM/C-RZF for HI based system ($N_{RF} = 4, K = 6$).

having 2-3 users per RF for best performance [35]. For higher loading like 10 users on 4 RF chains (that is, 4:10), the computation time gets unbearable making it unsuitable for delay-intolerant applications. The computation time factor is common with iterative methods [36]. This problem can be addressed using rate splitting for adding low data rate (LDR) users to the primary users [37].

Finally, the NMSE performance of the three stages observed reveals the impact of the first stage where only the initial correlation threshold $\delta(K, N_{RF})$ is updated in the work of [11]. This notable improvement does not go beyond the polynomial computational complexity of the CHS algorithm. This is because the first (K, N_{RF}) stage computation is done once for the entire number of iterations. The same goes for the second (R) stage. The third (RZF) stage can also be approximated as a scaled version of the ZF used in [11]. The robustness of RZF is felt in keeping the EE within the performance limit. An in-depth analysis that is out of the scope of this work can be conducted for further evaluation.

VI. CONCLUSION

In this work, a modified user grouping and HP design with joint power allocation (PA) and power splitting (PS) optimization for SWIDEH has been proposed for the HI HP-mmWave-mMIMO NOMA system interacting with common multi-antenna users. The point-to-point MIMO channel matrix realization for each user is reduced to a channel vector by a single antenna representative selection. The base station uses the assumed dimension-reduced uplink channel estimates from each user to design a hybrid (analog and digital) precoder. An adaptive initial threshold-based cluster head selection scheme is introduced by transforming a constant initial correlation threshold into a function of varying identified influencing parameters. This improves the quantized analog precoder, equivalent channel, and correlation-based

user grouping for a zero-forcing (ZF) digital precoder design to cancel inter-user interference. Thereafter, the joint power allocation at the base station and power splitting factor at the mobile station is optimized through the MMSE user detection technique. Such that the achievable sum rate is maximized at the base station translating into higher spectrum efficiency and energy efficiency as in the case of full and sub-connection types respectively. The computer simulation results from using the best EE antenna selection and adaptive initial threshold for CHS show significant SE/EE performance improvement as SNR increases compared to similar work in the literature.

For future work, it would be interesting to develop an actual channel estimation scheme for the system. Because it was observed that equivalent channel estimates, like the bit error rate (BER) pattern of [38], become less accurate compared to FD system performance with increasing SNR. Therefore, it is expected that machine learning [39] with more accurate equivalent channel estimates at [-20,0] dB SNR can be trained for improving the performance at [-5,30] dB SNR. One particular interest is the use of transfer learning (TL) because, unlike most machine learning (ML) which requires large data sets, transfer learning does not [40], [41], [42], [43], and [44]. Alternatively, the SNR can be factored among parameters that determine the correlation threshold for cluster head selection (CHS) or ZF. Another direction will be to explore the analytical performance of regularized zero-forcing in hardware and non-hardware impaired-based mm-wave massive MIMO.

REFERENCES

[1] H. Zhang, A. Dong, S. Jin, and D. Yuan, "Joint transceiver and power splitting optimization for multiuser MIMO SWIPT under MSE QoS constraints," *IEEE Trans. Veh. Technol.*, vol. 66, no. 8, pp. 7123–7135, Aug. 2017.

[2] A. N. Uwaechia and N. M. Mahyuddin, "A comprehensive survey on millimeter wave communications for fifth-generation wireless networks: Feasibility and challenges," *IEEE Access*, vol. 8, pp. 62367–62414, 2020.

[3] M. Vaezi, G. A. A. Baduge, Y. Liu, A. Arafa, F. Fang, and Z. Ding, "Interplay between noma and other emerging technologies: A survey," *IEEE Trans. Cognit. Commun. Netw.*, vol. 5, no. 4, pp. 900–919, Dec. 2019.

[4] Y. Liu, Z. Qin, M. Elkashlan, Z. Ding, A. Nallanathan, and L. Hanzo, "Nonorthogonal multiple access for 5G and beyond," *Proc. IEEE*, vol. 105, no. 12, pp. 2347–2381, Dec. 2017.

[5] R. Shrestha and G. Amarasuriya, "SWIPT in hybrid relay-assisted massive MIMO downlink," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–7.

[6] W. Qu, X. Cheng, and L. Yang, "Hybrid spatial-modulation based virtual MIMO relaying protocol with SWIPT," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–6.

[7] Z. Yang, Z. Ding, P. Fan, and N. Al-Dhahir, "The impact of power allocation on cooperative non-orthogonal multiple access networks with SWIPT," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 900–919, Jul. 2017.

[8] M. Aldababsa, M. Toka, S. Gökçeli, G. K. Kurt, and O. Kucur, "A tutorial on nonorthogonal multiple access for 5G and beyond," *Wireless Commun. Mobile Comput.*, vol. 2018, pp. 1–24, Jun. 2018, doi: 10.1155/2018/9713450.

[9] Y. Zhao, J. Hu, Z. Ding, and K. Yang, "Joint interleaver and modulation design for multi-user SWIPT-NOMA," *IEEE Trans. Commun.*, vol. 67, no. 10, pp. 7288–7301, Oct. 2019.

[10] S. Li, Z. Wan, L. Jin, and J. Du, "Energy harvesting maximizing for millimeter-wave massive MIMO-NOMA," *Electronics*, vol. 9, no. 1, p. 32, Dec. 2019.

[11] L. Dai, B. Wang, M. Peng, and S. Chen, "Hybrid precoding-based millimeter-wave massive MIMO-NOMA with simultaneous wireless information and power transfer," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 1, pp. 131–141, Jan. 2019.

[12] X. Li, J. Li, P. T. Mathiopoulos, D. Zhang, L. Li, and J. Jin, "Joint impact of hardware impairments and imperfect CSI on cooperative SWIPT NOMA multi-relaying systems," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Aug. 2018, pp. 95–99.

[13] Y. Mao, B. Clerckx, and V. O. K. Li, "Rate-splitting for multi-user multi-antenna wireless information and power transfer," in *Proc. IEEE 20th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jul. 2019, pp. 1–5.

[14] M. R. G. Aghdam, R. Abdolee, F. A. Azhiri, and B. M. Tazehkand, "Random user pairing in massive-MIMO-NOMA transmission systems based on mmWave," in *Proc. IEEE 88th Veh. Technol. Conf. (VTC-Fall)*, Aug. 2018, pp. 1–6.

[15] J. Zhu, Q. Li, Z. Liu, H. Chen, and H. V. Poor, "Enhanced user grouping and power allocation for hybrid mmWave MIMO-NOMA systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 3, pp. 2034–2050, 2022.

[16] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Jun. 2014.

[17] D. Fan, F. Gao, A. L. Swindlehurst, A. Ashikhmin, G. Wang, Z. Zhong, and A. Nallanathan, "Channel estimation and transmission strategy for hybrid mmWave NOMA systems," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 3, pp. 584–596, Jun. 2019.

[18] M. Zeng, W. Hao, O. A. Dobre, and H. V. Poor, "Energy-efficient power allocation in uplink mmWave massive MIMO with NOMA," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 3000–3004, Mar. 2019.

[19] Y. Liu, L.-L. Yang, and L. Hanzo, "Spatial modulation aided sparse code-division multiple access," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1474–1487, Mar. 2018.

[20] Z. Gao, L. Dai, Z. Wang, S. Chen, and L. Hanzo, "Compressive-sensing-based multiuser detector for the large-scale SM-MIMO uplink," *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 8725–8730, Oct. 2016.

[21] W. M. Audu and O. O. Oyerinde, "Iteratively reweighted super-resolution channel estimation in hardware-impaired hybrid-precoded MM-wave massive MIMO systems using dual SVD and Marquardt's global search," *Digit. Signal Process.*, vol. 120, Jan. 2022, Art. no. 103308.

[22] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Found. Trends Signal Process.*, vol. 11, nos. 3–4, pp. 154–655, 2017.

[23] B. Selim, S. Muhaidat, P. C. Sofotasios, A. Al-Dweik, B. S. Sharif, and T. Stouraitis, "Radio-frequency front-end impairments: Performance degradation in nonorthogonal multiple access communication systems," *IEEE Veh. Technol. Mag.*, vol. 14, no. 1, pp. 89–97, Mar. 2019.

[24] L. Dai, B. Wang, Y. Yuan, S. Han, C.-L. I, and Z. Wang, "Non-orthogonal multiple access for 5G: Solutions, challenges, opportunities, and future research trends," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 74–81, Sep. 2015.

[25] Z. Ding, F. Adachi, and H. V. Poor, "The application of MIMO to non-orthogonal multiple access," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 537–552, Jan. 2015.

[26] X. Gao, L. Dai, S. Han, I. Chih-Lin, and R. W. Heath, "Energy-efficient hybrid analog and digital precoding for mmWave MIMO systems with large antenna arrays," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 998–1009, Apr. 2016.

[27] J. Ni, J. Zuo, S. Han, S. Wang, and C.-L. I, "On the number of admissible users in NOMA," *IEEE Wireless Commun. Lett.*, vol. 8, no. 3, pp. 793–796, Jun. 2019.

[28] F. Wu, "Sherman-Morrison-Woodbury formula for linear integrodifferential equations," *Math. Probl. Eng.*, vol. 2016, Oct. 2016, Art. no. 9418730, doi: 10.1155/2016/9418730.

[29] W. W. Hager, "Updating the inverse of a matrix," *SIAM Rev.*, vol. 31, no. 2, pp. 221–239, Jun. 1989.

[30] M. C. Grant and S. T. Boyd. (2012). *CVX: MATLAB Software for Disciplined Convex Programming, Version 2.2*. [Online]. Available: http://cvxr.com/cvx

[31] H. Tabassum, M. S. Ali, E. Hossain, M. J. Hossain, and D. I. Kim, "Non-orthogonal multiple access (NOMA) in cellular uplink and downlink: Challenges and enabling techniques," 2016, *arXiv:1608.05783*.
[32] V. C. Rodrigues, A. Amiri, T. Abrao, E. D. Carvalho, and P. Popovski, "Low-complexity distributed XL-MIMO for multiuser detection," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Jun. 2020, pp. 1–6.
[33] D. Zhang, Z. Zhou, C. Xu, Y. Zhang, J. Rodriguez, and T. Sato, "Capacity analysis of NOMA with mmWave massive MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1606–1618, Jul. 2017.
[34] O. O. Oyerinde, "Modified orthogonal matching pursuit (MOMP)-based multiuser detector for uplink grant free NOMA systems," in *Proc. 13th Int. Conf. Signal Process. Commun. Syst. (ICSPCS)*, Dec. 2019, pp. 1–6.
[35] D. Zhai, R. Zhang, L. Cai, B. Li, and Y. Jiang, "Energy-efficient user scheduling and power allocation for NOMA-based wireless networks with massive IoT devices," *IEEE Internet Things J.*, vol. 5, no. 3, pp. 1857–1868, Jun. 2018.
[36] M. Vaezi, Z. Ding, and H. V. Poor, Eds., *Multiple Access Techniques for 5G Wireless Networks and Beyond*. Switzerland: Springer, 2018, doi: 10.1007/978-3-319-92090-0.
[37] N. M. Balasubramanya, A. Gupta, and M. Sellathurai, "Combining code-domain and power-domain NOMA for supporting higher number of users," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2018, pp. 1–6, doi: 10.1109/GLOCOM.2018.8647770.
[38] A. Singh, K. K. Naik, and C. R. S. Kumar, "Novel pilot-aided channel estimation scheme for power domain NOMA-UFMC system in fading scenarios," in *Proc. Int. Conf. Range Technol. (ICORT)*, Feb. 2019, pp. 1–5, doi: 10.1109/ICORT46471.2019.9069626.
[39] H. Jiang, M. Cui, D. W. K. Ng, and L. Dai, "Accurate channel prediction based on transformer: Making mobility negligible," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 9, pp. 2717–2732, Sep. 2022.
[40] S. J. Nawaz, S. K. Sharma, S. Wyne, M. N. Patwary, and M. Asaduzzaman, "Quantum machine learning for 6G communication networks: State-of-the-art and vision for the future," *IEEE Access*, vol. 7, pp. 46317–46350, 2019, doi: 10.1109/ACCESS.2019.2909490.
[41] C. Qing, L. Dong, L. Wang, G. Ling, and J. Wang, "Transfer learning-based channel estimation in orthogonal frequency division multiplexing systems using data-nulling superimposed pilots," *PLoS ONE*, vol. 17, no. 5, May 2022, Art. no. e0268952.
[42] N. Van Huynh and G. Y. Li, "Transfer learning for signal detection in wireless networks," *IEEE Wireless Commun. Lett.*, vol. 11, no. 11, pp. 2325–2329, Nov. 2022.
[43] C. T. Nguyen, N. Van Huynh, N. H. Chu, Y. M. Saputra, D. T. Hoang, D. N. Nguyen, Q.-V. Pham, D. Niyato, E. Dutkiewicz, and W.-J. Hwang, "Transfer learning for wireless networks: A comprehensive survey," *Proc. IEEE*, vol. 110, no. 8, pp. 1073–1115, Aug. 2022.
[44] B. Lim, W. J. Yun, J. Kim, and Y.-C. Ko, "Joint pilot design and channel estimation using deep residual learning for multi-cell massive MIMO under hardware impairments," *IEEE Trans. Veh. Technol.*, vol. 71, no. 7, pp. 7599–7612, Jul. 2022.

WAHEED M. AUDU (Member, IEEE) received the M.Eng. degree in communication engineering from the Federal University of Technology (FUT), Minna, Nigeria, in 2014. He is currently pursuing the Ph.D. degree in electrical engineering with the University of the Witwatersrand, Johannesburg, South Africa. He was an IEEEX treme 16.0 Ambassador with the University of the Witwatersrand. He has been with the Department of Telecommunication Engineering, FUT, since 2010, where he is currently a Lecturer. He is also with National Research Foundation (NRF), South Africa, Tertiary Education Trust Fund (TETFund), a Nigeria Awardee, and a Registered Engineer (R.Eng.) with the Council for the Regulation of Engineering in Nigeria (COREN). His current research interests include wireless communication system impediments, modeling, and channel estimation strategies, with a major focus on new and future-generation radio technologies.

OLUTAYO O. OYERINDE (Senior Member, IEEE) received the Ph.D. degree in electronic engineering from the University of Kwazulu-Natal, Durban, South Africa, in 2011. He has been with the School of Electrical and Information Engineering, University of the Witwatersrand, Johannesburg, South Africa, since 2013, where he is currently an Associate Professor. He is also a National Research Foundation (NRF) C2 Rated Scientist, a Registered Professional Engineer (Pr.Eng.) with the Engineering Council of South Africa (ECSA), a Registered Engineer (R.Eng.) with COREN, and a Corporate Member of NSE. His current research interests include wireless communications with specific interests in 5G and beyond 5G technologies and signal processing techniques for wireless communication systems. He is an Associate Editor of the IEEE Access and an Editorial Board Member of the *International Journal of Sensors, Wireless Communications & Control*.