User Selection and Codebook Design for NOMA-Based High Altitude Platform Station (HAPS) Communications

İrem Cumalı, Berna Özbek, *Senior Member, IEEE*, Güneş Karabulut Kurt, *Senior Member, IEEE*, Halim Yanikomeroglu, *Fellow, IEEE*

***Abstract*—High altitude platform station (HAPS) communications have made a tremendous impact on recent research into sixth-generation (6G) and beyond wireless networks. The large coverage area and significant computational capability of HAPS systems enable many areas of utilization in 6G and beyond applications, including Internet of Things (IoT) services, augmented reality, and connected autonomous vehicles. In addition, non-orthogonal multiple access (NOMA) is a cutting-edge technology that can be utilized to enhance spectral efficiency in HAPS systems. In this paper, we exploit NOMA-based HAPS communications and multiple antennas to meet the connectivity, reliability, and high-data-rate requirements of 6G and beyond applications. We propose a user selection and correlation-based user pairing algorithm for a NOMA-based multi-user HAPS system. Moreover, we investigate the codebook design for HAPS communication and adapt the polar-cap codebook (PCC) to the HAPS channel which shows Rician fading propagation characteristics dominated by the line-of-sight (LOS) component. Performance evaluations show that the proposed user selection algorithm is perfectly suited to the HAPS channel and that the PCC provides a remarkable spectral efficiency.**

***Index Terms*—High altitude platform station (HAPS), non-orthogonal multiple access (NOMA), user selection, codebook design, sixth-generation (6G) communications.**

I. Introduction

With the ever-increasing data requirements of sixth-generation (6G) and beyond wireless communications, high altitude platform station (HAPS) systems have emerged as effective complementary solutions to satellite and terrestrial communications. A HAPS can be defined as an aerial vehicle that flies about 20 km above ground level. The altitude of a HAPS provides a comparatively large coverage area as well as lower transmission delays. In addition, the wind speed in that altitude is relatively low providing moderate turbulence [1]. HAPS communications systems can provide a superior performance when compared to satellite communications having higher energy efficiency, lower propagation loss and delay, and lower deployment costs [2].

İ. Cumalı and B. Özbek are with the Department of Electrical and Electronics Engineering, Izmir Institute of Technology, 35430, Izmir, Türkiye, e-mail: {iremocal, bernaozbek}@iyte.edu.tr.

G. Karabulut Kurt is with the Poly-Grames Research Center, Department of Electrical Engineering, Polytechnique Montréal, Montréal, Canada, e-mail: gunes.kurt@polymtl.ca

H. Yanikomeroglu is with the Department of Systems and Computer Engineering, Carleton University, Ottawa, Canada, e-mail: halim@sce.carleton.ca

HAPS systems emerge as an essential enabler for integrated communications, computing, sensing, navigation, and positioning. The authors of [3] have envisioned several use-cases for a HAPS-mounted super macro base station, such as delivering Internet of Things (IoT) services, covering temporary and unexpected events, supporting intelligent transportation systems, managing aerial networks, and filling coverage gaps resulting from physical obstacles. Since HAPS systems can provide greater coverage and significant computational capability, they can offload data with low communication delays for IoT devices, augmented reality (AR) applications or connected autonomous vehicles.

HAPS systems can also be utilized as a complementary solution for terrestrial networks in areas without coverage (e.g. forests, oceans, etc.) [3]. In the literature, HAPS systems have been integrated with many existing technologies to meet the requirements of 6G systems. The authors of [4] considered the cooperation of low earth orbit (LEO) satellites and HAPS systems for a massive number of mobile and IoT devices with the aim of providing ubiquitous access for 6G applications, and [5] utilized HAPS in a space-air-ground integrated network (SAGIN), which is also in cooperation with satellites, to support ground-to-space link transmission. Also, [6] and [7] exploited a HAPS system as a relay between a ground station and a satellite in a SAGIN. The use of HAPSs as relays improves the reliability of the link between the ground station and the satellite, and reliability is vital for mission-critical applications. Thus, we utilize a HAPS-based communication system employing both non-orthogonal multiple access (NOMA) and multiple antennas to meet connectivity, reliability, and high data rate requirements of 6G and beyond applications.

NOMA is a promising technology capable of supplying higher spectral efficiency compared to orthogonal multiple access schemes, such as time-division multiple access (TDMA), frequency-division multiple access (FDMA), and code-division multiple access (CDMA) [8]. Conventional orthogonal multiple access schemes allocate different users to orthogonal resources in time, frequency, or code. However, in NOMA, multiple users can share time and frequency resources in the same spatial layer via power-domain or code-domain multiplexing [9]. In other words, NOMA can serve more users through non-orthogonal resource allocation. In power-domain multiplexing, which is also considered in this work, different power levels are assigned to users based on their channel

conditions. In this method, the aim is to maximize the spectral efficiency by using superposition coding in combination with successive interference cancellation (SIC) to eliminate multi-user interference. Thus, NOMA is a promising candidate for enhancing the spectral efficiency in HAPS systems [10].

User selection in HAPS systems is another important issue that needs to be addressed. A single HAPS can cover hundreds of square kilometers [11], and in this coverage area, a significant number of users may require high data rates. However, not all users can be served simultaneously by a limited number of antennas. Therefore, user selection in HAPS systems is crucial in terms of overall system performance. On the side of NOMA, in addition to serving multiple users by each beam, the selection of users and the determination of NOMA user pairs also play critical roles in increasing the spectral efficiency. In the literature, user grouping and clustering methods in HAPS systems were examined in [12]–[14]. The authors of [12] used the average chordal distance between the statistical eigenmodes for user grouping. In [13], a user clustering was presented on the basis of the line-of-sight (LOS) component of a user channel when a uniform linear array was utilized with a HAPS. Moreover, [14] presented a user grouping and beamforming scheme for beamspace HAPS-NOMA. However, it did not consider user selection and limited feedback issues for HAPS-NOMA systems. The authors of [15] investigated a millimeter-wave HAPS-NOMA system from the point of power allocation having only two users, which is very low for a HAPS system. User pairing/clustering for traditional NOMA-based multiple-input single-output (MISO) systems has also been examined in the literature. The sorting algorithm was presented in [16]; an improved user clustering algorithm (IUCA) considering both channel correlation and channel gain difference was presented in [17]; spatial correlation-based clustering that selects clusters on the basis of user spatial channel properties relative to generated orthogonal directions was provided in [18]; and the semi-orthogonal user selection (SUS) with a user-matching algorithm was presented in [19]. Nevertheless, the user selection issue in NOMA-based HAPS systems has not yet been discussed in the literature to the best of our knowledge.

In this work, we propose a channel correlation-based user pairing method for a NOMA-based HAPS system. The propagation characteristics of the HAPS channel show that if two users are located in proximity to each other, their channels are highly correlated. This is because the LOS component predominates the propagation in the HAPS channel. Therefore, pairing the highly correlated users in the concept of NOMA can improve the spectral efficiency.

Furthermore, transmit beamforming is a useful technique to exploit the multiplexing gain in multiple-input multiple-output (MIMO) systems. Yet, the channel state information (CSI) is required at the transmitter to apply beamforming for frequency-division duplex (FDD) systems in which the downlink and uplink channels are not reciprocal, and acquiring the perfect CSI is impossible in practice [20]. To address this problem, the transmitter and receiver use a common set of codewords; the receiver then sends the best codeword index to the transmitter by means of a limited-rate feedback channel [21]. The best codeword refers to the best-fitted codeword to the channel estimated by the receiver; it is also termed as the transmit beamforming vector. In this way, the transmitted signal is adapted to the channel. This is also the main idea behind the feedback system. In the literature, the codebook design for Rayleigh fading channels has been widely discussed [22], including Grassmannian beamforming [23], random vector quantization (RVQ) [24], and polar-cap codebook (PCC) [25]. However, the HAPS channel shows Rician fading propagation characteristics. The authors of [26] presented a PCC design for MISO Rician fading channels. Therefore, we adapt the PCC in [26] to the HAPS channel.

The contributions of this paper can be summarized as follows:

C1 The user selection issue is discussed in NOMA-based HAPS communications to meet high data rate requirements of 6G and beyond applications. To the best of our knowledge, this is the first work of its kind in the area of NOMA-based HAPS communication with user selection and limited feedback link.

C2 A joint algorithm of user selection and correlation-based user pairing is proposed for the NOMA-based multi-user HAPS communications by considering the maximization of the beamforming gain while having reduced computational complexity.

C3 Codebook design is studied for a limited-feedback HAPS system. From this point of view, the polar-cap codebook is adapted to the HAPS channel by utilizing the LOS component. Extensive numerical results are provided to demonstrate the performance of the proposed approaches.

The rest of this paper is organized as follows. In Section II, the system model is described in detail. The channel model for a multi-user HAPS is also provided and approximation of the correlation matrix for non-line-of-sight (NLOS) component is presented. In Section III, the proposed user selection algorithm is explained thoroughly. Moreover, the precoding schemes for the NOMA user pairs and the complexity analysis are given. Section IV describes the PCC design for a HAPS system with a limited feedback channel. Section V provides the simulation results of the proposed user selection algorithm and polar-cap codebook design. Finally, conclusions are provided in Section VI.

Notation. Matrices and vectors are denoted by upper-case and lower-case boldface letters, respectively. Inversion, transpose, conjugate transpose, vector norm, trace, and expectation operations are represented by $(.)^{-1}$, $(.)^{T}$, $(.)^{H}$, $\|.\|$, $\mathrm{tr}(.)$, and $\mathbb{E}\{.\}$, respectively. Also, $|.|$ denotes absolute value for the scalars and cardinality for the sets. $\mathbb{C}$ represents the field of complex numbers. $\mathbf{I}_n$ is an $n\times n$ identity matrix while $\mathbf{0}_{n\times m}$ is an $n \times m$ zero matrix. In addition, $\mathcal{CN}(0, \sigma^2)$ is the complex Gaussian random variable with zero mean and variance σ^2. For the complexity analysis, $\mathcal{O}(.)$ is the big-O notation.

II. System Model

We consider a downlink NOMA-based multi-user HAPS system with N_{T} transmit antennas and L single-antenna users. For the transmitter, HAPS is equipped with a uniform planar

array (UPA) with $M \times N$ antennas utilizing M antennas set vertically and N antennas set horizontally. In total, $N_{\mathrm{T}} = MN$ transmit antennas serve the active user pairs by NOMA as shown in Figure 1. The number of active users is L_u, which corresponds to $L_u/2$ NOMA pairs. In this scenario, there is an equal number of users utilizing applications which demand high data rates and low data rates (i.e., high-rate users and low-rate users).

The received signal is given by

$$\mathbf{r} = \mathbf{H}^H \mathbf{W}\mathbf{s} + \mathbf{n}, \tag{1}$$

where $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_{L_u}] \in \mathbb{C}^{N_{\mathrm{T}} \times L_u}$ is the overall channel matrix, $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{L_u}] \in \mathbb{C}^{N_{\mathrm{T}} \times L_u}$ is the precoding matrix, $\mathbf{s} \in \mathbb{C}^{L_u \times 1}$ denotes the symbol vector such that $\mathbb{E}\{\mathbf{s}\mathbf{s}^H\} = \mathbf{I}_{L_u}$, and $\mathbf{n} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_{L_u})$ is the additive white Gaussian noise (AWGN) vector.

The channel vector for the ℓ^{th} user is defined as [12]

$$\mathbf{h}_\ell = \sqrt{\alpha_\ell} \left(\sqrt{\frac{K_\ell}{K_\ell + 1}} \overline{\mathbf{h}}_\ell + \sqrt{\frac{1}{K_\ell + 1}} \widehat{\mathbf{h}}_\ell \right), \tag{2}$$

where K_ℓ is the Rician factor, $\overline{\mathbf{h}}_\ell$ is the LOS component, and $\widehat{\mathbf{h}}_\ell$ is the NLOS component for the ℓ^{th} user. $\alpha_\ell = (4\pi r_\ell/\lambda)^{-2}$ denotes the large-scale fading factor where r_ℓ is the distance between the HAPS and the ℓ^{th} user, λ is the carrier wavelength. λ is also defined by c/f_c where c is the speed of light and f_c is the carrier frequency. The LOS component of the channel is given by

$$\overline{\mathbf{h}}_\ell = \mathbf{a}_v(\theta_\ell, \varphi_\ell) \otimes \mathbf{a}_h(\theta_\ell, \varphi_\ell), \tag{3}$$

where $\otimes$ denotes the Kronecker product. The array steering vectors along vertical and horizontal directions are denoted by $\mathbf{a}_v(\theta_\ell, \varphi_\ell)$ and $\mathbf{a}_h(\theta_\ell, \varphi_\ell)$, respectively. They are defined by exponentials such that

$$\begin{aligned} \mathbf{a}_v(\theta_\ell, \varphi_\ell) &= \left[1, e^{j2\pi d_v}, \ldots, e^{j2\pi(M-1)d_v}\right]^T \in \mathbb{C}^{M \times 1}, \\ \mathbf{a}_h(\theta_\ell, \varphi_\ell) &= \left[1, e^{j2\pi d_h}, \ldots, e^{j2\pi(N-1)d_h}\right]^T \in \mathbb{C}^{N \times 1}, \end{aligned} \tag{4}$$

where the parameters of the exponentials along the vertical and horizontal directions are defined as $d_v = d_c \sin\theta_\ell \cos\varphi_\ell/\lambda$ and $d_h = d_r \sin\theta_\ell \sin\varphi_\ell/\lambda$, respectively. The antenna spacings in a column and a row are denoted by d_c and d_r, the angle of departure (AoD) along vertical and horizontal directions are $\theta_\ell \in [0, \pi/2)$ and $\varphi_\ell \in [-\pi, \pi]$, respectively. On the other hand, the NLOS component of the channel is defined as follows

$$\widehat{\mathbf{h}}_\ell = \sqrt{\widehat{\mathbf{R}}_\ell}\, [w_{1\ell}, w_{2\ell}, \ldots, w_{N_{\mathrm{T}}\ell}]^T, \tag{5}$$

where $\widehat{\mathbf{R}}_\ell \in \mathbb{C}^{N_{\mathrm{T}} \times N_{\mathrm{T}}}$ is the correlation matrix of the ℓ^{th} user, and $w_{n\ell} \sim \mathcal{CN}(0, 1)$ for $n = 1, 2, \ldots, N_{\mathrm{T}}$.

The $(p, q)^{th}$ element of the correlation matrix can be given by [14]

$$\left[\widehat{\mathbf{R}}_\ell\right]_{p,q} = \int_0^{\pi/2} \int_{-\pi}^{\pi} f(\varphi)\, f(\theta) e^{j2\pi(d_1+d_2)} d\varphi\, d\theta, \tag{6}$$

where $f(\theta) \propto e^{\frac{-\sqrt{2}|\theta-\theta_0|}{\delta}}$, θ_0 is the mean vertical AoD, δ is the vertical angular spread, $f(\varphi) = \frac{e^{\kappa\cos(\varphi-\mu_\ell)}}{2\pi \mathrm{I}_0(\kappa)}$, $\mathrm{I}_0(.)$ is the zeroth-order modified Bessel function of the first kind, $\mu_\ell \in [-\pi, \pi]$ is the horizontal AoD of the ℓ^{th} user, κ is the parameter controlling the horizontal angular spread, $d_1 = (p-q)d_c \sin\theta\cos\varphi/\lambda$, and $d_2 = (p-q)d_r \sin\theta \sin\varphi/\lambda$.

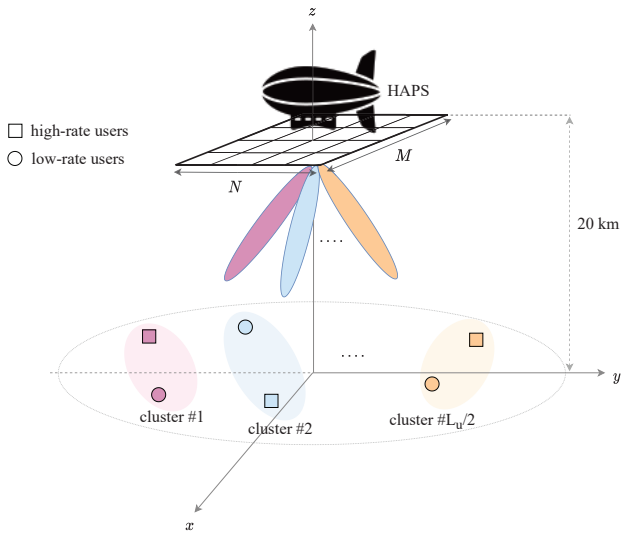


Figure 1: System model of the NOMA-based multi-user HAPS.

A. Approximation of the Correlation Matrix

The correlation matrix is computationally formidable since it has no closed form expression in the literature. Thus, we exploit the approximation of the correlation matrix that was proposed in [27]. This approximation utilizes Simpson's numerical integration method in order to analyze the correlation by a flexible expression. By substituting d_1, d_2, and $f(\varphi)$ in (6), the correlation matrix is rewritten as

$$\left[\widehat{\mathbf{R}}_\ell\right]_{p,q} = \int_0^{\pi/2} \frac{f(\theta)}{2\pi \mathrm{I}_0(\kappa)}\, I_\varphi\, d\theta, \tag{7}$$

where the integral with respect to φ is determined as $I_\varphi = \int_{-\pi}^{\pi} e^{\kappa\cos(\varphi-\mu_\ell)} e^{j\chi(\cos\varphi+\sin\varphi)} d\varphi$ with $\chi = \pi(p-q)\sin\theta$. Then, the integral is simplified to

$$I_\varphi = 2\pi \mathrm{I}_0\left(\sqrt{\kappa^2 - 2\chi^2 + j2\chi\kappa(\cos\mu_\ell + \sin\mu_\ell)}\right). \tag{8}$$

Since (8) depends only on the variable θ, it is expressed as $I_\varphi = g(\theta)$ in the rest of the paper. Thus, the correlation matrix is stated as

$$\left[\widehat{\mathbf{R}}_\ell\right]_{p,q} = \frac{1}{2\pi \mathrm{I}_0(\kappa)} \int_0^{\pi/2} h(\theta)\, d\theta, \tag{9}$$

where $h(\theta) = f(\theta)g(\theta)$. To find the approximation for the definite integral in (9), Simpson's first rule [28] can be applied so that we have

$$\left[\widehat{\mathbf{R}}_\ell\right]_{p,q} \approx \frac{I}{2\pi \mathrm{I}_0(\kappa)}. \tag{10}$$

By Simpson's 1/3 rule for n subintervals, the integral can be approximated by

$$I = \frac{h_1}{3}\left[h(\vartheta_0) + 4\sum_{i=1}^{n/2} h(\vartheta_{2i-1}) + 2\sum_{j=1}^{n/2-1} h(\vartheta_{2j}) + h(\vartheta_n)\right], \tag{11}$$

where $h_1 = \left(\frac{\pi}{2} - 0\right)/n$, $\vartheta_0 = 0$, $\vartheta_k = \vartheta_0 + kh_1$ for $k = 1, 2, \ldots, n$, and n is an even number.

III. Proposed User Selection Algorithm

In this section, we provide the details of the proposed user selection algorithm, described in Algorithm 1. Since the LOS component dominates the HAPS channel, the channels of closely located users can be highly correlated. Therefore, to increase the sum data rate, we utilize a pairing algorithm that selects highly correlated users with different applications for the considered NOMA-based HAPS systems. The proposed algorithm can be efficiently applied in densely populated metropolitan areas as envisioned in [3].

In a conventional NOMA system, a NOMA user pair includes a strong user with higher signal-to-noise ratio (SNR) and a weak user with lower SNR. By contrast, in a HAPS system, the height of HAPS dominates the large-scale fading; therefore, users all have similar SNR values and cannot be classified as strong or weak in terms of large-scale fading. In this work, we consider an application that has $L/2$ users that demand high data rates, which we refer to as high-rate users, and $L/2$ users that demand low data rates, which we refer to as low-rate users. The high-rate users and the low-rate users are randomly distributed over the HAPS coverage area. The set of high-rate users' indexes and the set of low-rate users' indexes are denoted by $\mathcal{K}_{\text{high}}$ and $\mathcal{K}_{\text{low}}$ such that $|\mathcal{K}_{\text{high}}| = |\mathcal{K}_{\text{low}}| = L/2$, respectively. The corresponding channels are then constructed by selecting the channel vectors of corresponding users from the spatial channel matrix $\mathbf{H}$ such that

$$\begin{aligned}\mathbf{H}_{\text{high}} &= [\mathbf{h}_p]_{\forall p \in \mathcal{K}_{\text{high}}} = [\mathbf{h}_{\text{high}_{(1)}}, \mathbf{h}_{\text{high}_{(2)}}, \ldots, \mathbf{h}_{\text{high}_{(L/2)}}],\\ \mathbf{H}_{\text{low}} &= [\mathbf{h}_r]_{\forall r \in \mathcal{K}_{\text{low}}} = [\mathbf{h}_{\text{low}_{(1)}}, \mathbf{h}_{\text{low}_{(2)}}, \ldots, \mathbf{h}_{\text{low}_{(L/2)}}].\end{aligned} \tag{12}$$

The proposed algorithm consists of two parts: 1) the selection of high-rate users and 2) the selection of low-rate users constituting the NOMA pairs. At first, high-rate users to be served are determined based on the channel gains. For hundreds of kilometers coverage area, which is typical for a HAPS, exploiting the channel gains to select the high-rate users can maximize the sum data rate. Therefore, the first part of Algorithm 1 determines the norm of the channel vector for each high-rate user and forms the channel norm vector $\mathbf{v}$. Then, it sorts the elements of $\mathbf{v}$ in a descending order to obtain the vector of corresponding user indices $\mathbf{u}$. In this way, the algorithm finds $L_u/2$ high-rate users having the largest channel gains among the $L/2$ high-rate users. The resulting channel matrix of the selected high-rate users is obtained as $\widehat{\mathbf{H}}_{\text{high}}$.

For the low-rate users, the algorithm selects the most correlated low-rate user for each selected high-rate user to construct NOMA user pairs. The aim behind selecting the most correlated low-rate user is to maximize the beamforming gain. At first, a temporary channel matrix $\mathbf{H}_{\text{temp}}$ is defined as the channel of low-rate users $\mathbf{H}_{\text{low}}$. Then, the $(L_u/2) \times (L/2)$ dimensional correlation matrix $\mathbf{C}$ is defined by

$$\mathrm{c}(i,j) = \frac{|\widehat{\mathbf{h}}^H_{\text{high}_{(i)}} \mathbf{h}_{\text{temp}_{(j)}}|}{\|\widehat{\mathbf{h}}_{\text{high}_{(i)}}\| \, \|\mathbf{h}_{\text{temp}_{(j)}}\|}, \tag{13}$$

where $\widehat{\mathbf{h}}_{\text{high}_{(i)}} \in \mathbb{C}^{N_{\text{T}} \times 1}$ and $\mathbf{h}_{\text{temp}_{(j)}} \in \mathbb{C}^{N_{\text{T}} \times 1}$ are the i^{th} and j^{th} columns of $\widehat{\mathbf{H}}_{\text{high}}$ and $\mathbf{H}_{\text{temp}}$, respectively. The ℓ^{th} low-rate user is provided as the most correlated low-rate user to the i^{th} high-rate user and (i, ℓ) is determined as a NOMA user pair. So, the channel vector of the ℓ^{th} user is set to zero and the ℓ^{th} user is added to the set of selected low-rate users $\mathcal{L}_{\text{low}}$. Finally, the resulting channel of low-rate users $\widehat{\mathbf{H}}_{\text{low}}$ is constructed by selecting the column vectors of $\mathbf{H}_{\text{low}}$. The NOMA user pairs generate the clusters shown in Figure 1, where each cluster includes a high-rate user and a low-rate user.

The precoding scheme is determined on the basis of the resulting channels of the high-rate users after performing the proposed user selection algorithm. Then, the zero forcing (ZF) precoding matrix can be given as

$$\widehat{\mathbf{W}} = \alpha \, \widehat{\mathbf{H}}_{\text{high}} \left(\widehat{\mathbf{H}}^H_{\text{high}} \widehat{\mathbf{H}}_{\text{high}}\right)^{-1}, \tag{14}$$

where $\widehat{\mathbf{W}} = [\widehat{\mathbf{w}}_1, \widehat{\mathbf{w}}_2, \ldots, \widehat{\mathbf{w}}_{L_u/2}] \in \mathbb{C}^{N_{\text{T}} \times (L_u/2)}$. The scaling

Algorithm 1: Proposed user selection algorithm for the NOMA-based HAPS system

Data: $\mathbf{H} \in \mathbb{C}^{N_{\text{T}} \times L}$, $\mathcal{K}_{\text{high}}$, $\mathcal{K}_{\text{low}}$
Result: $\widehat{\mathbf{H}}_{\text{high}} \in \mathbb{C}^{N_{\text{T}} \times (L_u/2)}$, $\widehat{\mathbf{H}}_{\text{low}} \in \mathbb{C}^{N_{\text{T}} \times (L_u/2)}$
Part-1: Selection of the high-rate users:
Initialize $\mathcal{L}_{\text{high}} = \emptyset$, $\mathbf{v} = \mathbf{0}_{(L/2)\times 1}$;
for $i = 1 \to L/2$ **do**
  $\mathbf{v}_{(i)} = ||\mathbf{h}_{\text{high}_{(i)}}||$;
end
$[\mathbf{v}_s, \mathbf{u}] = \text{sort}(\mathbf{v}, \text{'descend'})$;
$\mathcal{L}_{\text{high}} \leftarrow \{\mathbf{u}_{(1)}, \mathbf{u}_{(2)}, \ldots, \mathbf{u}_{(L_u/2)}\}$;
$\widehat{\mathbf{H}}_{\text{high}} = \left[\mathbf{h}_{\text{high}_{(k)}}\right]_{\forall k \in \mathcal{L}_{\text{high}}}$;
Part-2: Selection of the low-rate users:
Initialize $\mathbf{H}_{\text{temp}} = \mathbf{H}_{\text{low}} \in \mathbb{C}^{N_{\text{T}} \times (L/2)}$;
Initialize $\mathbf{C} = \mathbf{0}_{(L_u/2)\times(L/2)}$, $\mathcal{L}_{\text{low}} = \emptyset$;
for $i = 1 \to L_u/2$ **do**
  for $j = 1 \to L/2$ **do**
    $\mathbf{C}_{(i,j)} = \mathrm{c}(i,j)$; /* Eq.(13) */
  end
  $\ell = \arg\max_{j \in \{1,2,\ldots,L/2\}} \mathbf{C}_{(i,j)}$;
  $\mathbf{h}_{\text{temp}_{(\ell)}} = \mathbf{0}_{N_{\text{T}} \times 1}$;
  $\mathcal{L}_{\text{low}} \leftarrow \mathcal{L}_{\text{low}} \cup \{\ell\}$;
end
$\widehat{\mathbf{H}}_{\text{low}} = \left[\mathbf{h}_{\text{low}_{(t)}}\right]_{\forall t \in \mathcal{L}_{\text{low}}}$;

factor is calculated by

$$\alpha = \sqrt{\frac{1}{\mathrm{tr}\left(\left(\widehat{\mathbf{H}}_{\mathrm{high}}^{H}\widehat{\mathbf{H}}_{\mathrm{high}}\right)^{-1}\right)}}. \tag{15}$$

The instantaneous signal-to-interference-plus-noise ratio (SINR) for the high-rate user in cluster i is given by

$$\gamma_{\mathrm{high}}^{(i)} = \frac{\mathrm{P}_{\mathrm{high}}\left|\widehat{\mathbf{h}}_{\mathrm{high}_{(i)}}^{H}\widehat{\mathbf{w}}_i\right|^2}{\mathrm{I}_{\mathrm{high}}^{(i)} + \mathrm{I}_{\mathrm{high,low}}^{(i)} + \mathrm{P}_n}, \tag{16}$$

where $\mathrm{P}_{\mathrm{high}}$ and $\mathrm{P}_{\mathrm{low}}$ are the transmit powers for the high-rate user and the low-rate user, respectively. The transmit powers are then defined by

$$\mathrm{P}_{\mathrm{high}} = \frac{\beta_{\mathrm{high}}\mathrm{P}_t}{(L_u/2)} \quad \text{and} \quad \mathrm{P}_{\mathrm{low}} = \frac{\beta_{\mathrm{low}}\mathrm{P}_t}{(L_u/2)}, \tag{17}$$

where P_t is the total transmit power, β_{high} and β_{low} are the fixed power allocation factors for the high-rate users and low-rate users such that $\beta_{\mathrm{high}} + \beta_{\mathrm{low}} = 1$ with $\beta_{\mathrm{high}} > \beta_{\mathrm{low}}$, respectively. In addition, the noise power over the bandwidth B is $\mathrm{P}_n = N_0 B$ where N_0 is the noise power spectral density. The interference for the high-rate user in cluster i caused by the users in other clusters, namely the inter-cluster interference, is written as

$$\mathrm{I}_{\mathrm{high}}^{(i)} = (\mathrm{P}_{\mathrm{high}} + \mathrm{P}_{\mathrm{low}}) \sum_{j=1, j\neq i}^{L_u/2} \left|\widehat{\mathbf{h}}_{\mathrm{high}_{(i)}}^{H}\widehat{\mathbf{w}}_j\right|^2, \tag{18}$$

where the ZF precoder can completely eliminate it under the perfect CSI. Furthermore, the interference for the high-rate user in cluster i from its NOMA pair, namely the intra-cluster interference, is

$$\mathrm{I}_{\mathrm{high,low}}^{(i)} = \mathrm{P}_{\mathrm{low}}\left|\widehat{\mathbf{h}}_{\mathrm{high}_{(i)}}^{H}\widehat{\mathbf{w}}_i\right|^2. \tag{19}$$

In each cluster, the low-rate user applies SIC so that the NOMA scheme is realized. The low-rate user decodes the signal of the high-rate user first, and then subtracts the signal from the received signal before decoding its own signal. In this way, the decoded signal of the low-rate user becomes free of intra-cluster interference. Therefore, the instantaneous SINR for the low-rate user in cluster i is given as

$$\gamma_{\mathrm{low}}^{(i)} = \frac{\mathrm{P}_{\mathrm{low}}\left|\widehat{\mathbf{h}}_{\mathrm{low}_{(i)}}^{H}\widehat{\mathbf{w}}_i\right|^2}{\mathrm{I}_{\mathrm{low}}^{(i)} + \mathrm{P}_n}, \tag{20}$$

where the interference for the low-rate user in cluster i caused by the users in other clusters (i.e., the inter-cluster interference) is expressed as follows

$$\mathrm{I}_{\mathrm{low}}^{(i)} = (\mathrm{P}_{\mathrm{high}} + \mathrm{P}_{\mathrm{low}}) \sum_{j=1, j\neq i}^{L_u/2} \left|\widehat{\mathbf{h}}_{\mathrm{low}_{(i)}}^{H}\widehat{\mathbf{w}}_j\right|^2. \tag{21}$$

The sum data rates for the high-rate users and low-rate users, which are selected to be served, are respectively defined as

$$\begin{aligned} \mathrm{R}_{\mathrm{high}} &= \mathbb{E}\left\{B\sum_{i=1}^{L_u/2}\log_2\left(1+\gamma_{\mathrm{high}}^{(i)}\right)\right\}, \\ \mathrm{R}_{\mathrm{low}} &= \mathbb{E}\left\{B\sum_{k=1}^{L_u/2}\log_2\left(1+\gamma_{\mathrm{low}}^{(k)}\right)\right\}. \end{aligned} \tag{22}$$

Eventually, the total sum data rate for all selected users is given by

$$\mathrm{R}_{\mathrm{sum}} = \mathrm{R}_{\mathrm{high}} + \mathrm{R}_{\mathrm{low}}. \tag{23}$$

A. Complexity Analysis

Here, we evaluate the complexity of the proposed user selection algorithm and the benchmark methods, namely the modified sorting algorithm of [16] and the modified IUCA of [17]. The proposed algorithm includes two parts, as we can see in Algorithm 1. Part-1 describes the user selection for the high-rate users based on the channel gains. The algorithm computes the norms of the N_{T}-dimensional channel vectors of $L/2$ users. Then, it sorts the elements of $(L/2)$-dimensional channel norm vector. Thus, the complexity of Part-1 in Algorithm 1 is determined as $\mathcal{O}((L/2)N_{\mathrm{T}} + (L/2)\log(L/2))$. Next, Part-2 pairs the users for the NOMA scheme by selecting the most correlated low-rate users for the selected high-rate users. For this purpose, the correlation matrix of dimension $(L_u/2)\times(L/2)$ is computed. So, the complexity of Part-2 in Algorithm 1 is $\mathcal{O}(N_{\mathrm{T}}L_uL)$.

On the other hand, the modified sorting algorithm applies the same procedure as Part-1 in Algorithm 1 for both the high-rate users and low-rate users. Therefore, the complexity of the modified sorting algorithm is calculated as $\mathcal{O}(LN_{\mathrm{T}} + L\log(L))$.

The modified IUCA also selects the high-rate users in the same way; thus, the complexity of this part is the same as the proposed and modified sorting algorithms. However, the complexity of the user pairing is given by $\mathcal{O}(x^2+x)$ where $x = L_u/2$ is the number of pairs. Therefore, the complexity is increased depending on L_u^2, which can be computationally intensive.

IV. Polar-Cap Codebook Design for HAPS Communication

This section provides the polar-cap codebook (PCC) design for HAPS communication. For a limited-feedback system in a HAPS system, the PCC considers the instantaneous LOS channel information in the HAPS channel, which is frequently modeled as Rician fading. In contrast to the conventional PCC in [25], the PCC design in our work is customized to the HAPS channel. The reason is that the radius of the conventional PCC is a predetermined constant, while the PCC design in this paper determines the radius depending on the channel direction information of the LOS component. When the LOS component in the HAPS channel is strong, the radius will be small; otherwise, it will be large. Thus, the PCC design depends on the instantaneous LOS channel. On the other hand, [26] applies the PCC for a MISO Rician fading channel containing

a small-scale fading gain for the LOS path, unlike the HAPS channel considered here. Moreover, we utilize the Householder transformation in [29] to obtain the rotated codewords.

We consider a PCC with $\mathcal{F} = \{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_{N_p}\}$ of cardinality $N_p = 2^b$ where the number of feedback bits b is a positive integer. The codewords have a unit norm such that $\|\mathbf{f}_m\| = 1$ for $m = 1, 2, \ldots, N_p$. The PCC has two types of codewords: namely, a basis codeword and non-basis codewords. Let the basis codeword be $\mathbf{f}_1 \in \mathbb{C}^{N_\mathrm{T} \times 1}$. The other codewords $\mathbf{f}_m \in \mathbb{C}^{N_\mathrm{T} \times 1}$ for $m = 2, 3, \ldots, N_p$ are the non-basis codewords. The design criteria of the PCC can thus be listed as follows [31]:

- The basis codeword is located at the center of the PCC.
- The non-basis codewords are located at the same distance from the basis codeword such that

$$d(\mathbf{f}_1, \mathbf{f}_m) = \epsilon \quad \text{for} \quad m = 2, 3, \ldots, N_p,$$

  where ϵ is the radius of the PCC.
- The minimum distance among non-basis codewords should be maximized such that

$$\max \left\{ \min_{l,m \in \{2,3,\ldots,N_p\}, l \neq m} d(\mathbf{f}_l, \mathbf{f}_m) \right\}.$$

Here, the distance metric is the *chordal distance*, which is defined by $d(\mathbf{a}, \mathbf{b}) = \sqrt{1 - |\mathbf{a}^H \mathbf{b}|^2}$ where $\mathbf{a}$ and $\mathbf{b}$ are N_T dimensional column vectors.

According to the design criteria mentioned above, a conventional PCC with the basis codeword $\mathbf{f}_1$ and radius ϵ, which is a predetermined positive constant, can be generated by

$$\mathcal{F} = \left\{ \mathbf{f}_1, \ \mathbf{U}_{\mathbf{f}_1} \begin{bmatrix} \sqrt{1-\epsilon^2} \\ \epsilon \mathbf{w}_1 \end{bmatrix}, \ldots, \mathbf{U}_{\mathbf{f}_1} \begin{bmatrix} \sqrt{1-\epsilon^2} \\ \epsilon \mathbf{w}_{N_p - 1} \end{bmatrix} \right\}, \tag{24}$$

where $\mathbf{U}_{\mathbf{f}_1} \in \mathbb{C}^{N_\mathrm{T} \times N_\mathrm{T}}$ is the unitary rotation matrix whose first column is the basis codeword $\mathbf{f}_1$ and $\mathcal{G} = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{N_p - 1}\}$ represents the Grassmannian codebook where $\mathbf{w}_n \in \mathbb{C}^{(N_\mathrm{T}-1) \times 1}$ with $\|\mathbf{w}_n\| = 1$ for $n = 1, 2, \ldots, N_p - 1$. For the basis codeword, any unit vector can be selected, such as $\mathbf{f}_1 = [1, 0 \ldots, 0]^T$.

PCC design for HAPS communication, by contrast, necessitates using the instantaneous LOS channel, since the LOS component dominates in HAPS propagation. Therefore, we select the basis codeword $\mathbf{f}_1$ as the unit vector in the direction of the LOS channel, i.e., $\mathbf{o}_{rot} \in \mathbb{C}^{N_\mathrm{T} \times 1}$

$$\mathbf{f}_1 = \mathbf{o}_{rot} = \frac{\overline{\mathbf{h}}}{\|\overline{\mathbf{h}}\|}, \tag{25}$$

where the LOS channel is denoted by $\overline{\mathbf{h}}$ [26]. Also, the radius of the PCC for HAPS communication can be stated as

$$\epsilon_{\mathbf{h}} = d\left(\mathbf{g}, \mathbf{o}_{rot}\right), \tag{26}$$

where $\mathbf{g} = \frac{\mathbf{h}}{\|\mathbf{h}\|}$ is the normalized channel vector (i.e., the maximum ratio transmission beamformer). Unlike conventional PCCs, the radius of the PCC for a HAPS system is not a predetermined constant; it depends on the channel direction information of LOS component. For HAPS systems, we adapt the radius of the PCC according to the $\mathbf{h}$, and the PCC in (24) can be expressed as

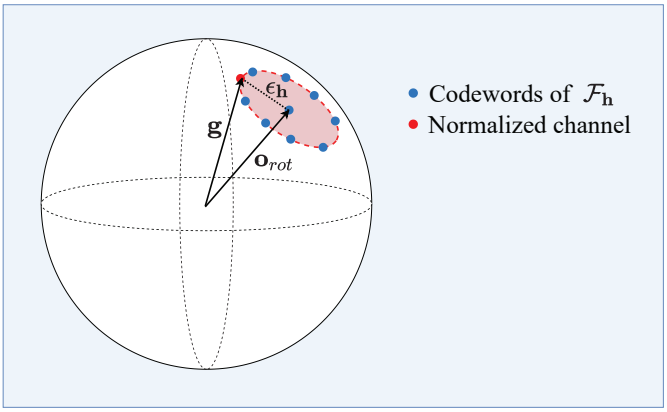


Figure 2: Illustration of the polar-cap codebook for HAPS communication.

$$\mathcal{F}_{\mathbf{h}} = \left\{ \mathbf{o}_{rot}, \ \mathbf{U}_{rot} \begin{bmatrix} \sqrt{1-\epsilon_{\mathbf{h}}^2} \\ \epsilon_{\mathbf{h}} \mathbf{w}_1 \end{bmatrix}, \ldots, \mathbf{U}_{rot} \begin{bmatrix} \sqrt{1-\epsilon_{\mathbf{h}}^2} \\ \epsilon_{\mathbf{h}} \mathbf{w}_{N_p - 1} \end{bmatrix} \right\}, \tag{27}$$

where the unitary rotation matrix $\mathbf{U}_{rot} \in \mathbb{C}^{N_\mathrm{T} \times N_\mathrm{T}}$ can be determined by the following complex Householder transform defined in [29]

$$\mathbf{U}_{rot} = \mathbf{I}_{N_\mathrm{T}} - \frac{1}{\mathbf{u}^H \mathbf{o}} \mathbf{u}\mathbf{u}^H, \tag{28}$$

where $\mathbf{o} = [1, 0 \ldots, 0]^T$ is an N_T dimensional unit vector, $\mathbf{u} = \mathbf{o} - \mathbf{o}_{rot}$, and $\mathbf{o}_{rot}$ is the unit vector in the LOS channel direction. The rotation matrix can also be computed through $\mathbf{o}_{rot}$ as follows

$$\mathbf{U}_{rot} = \mathbf{I}_{N_\mathrm{T}} + \frac{1}{\mathbf{o}_{rot}^H \mathbf{u}} \mathbf{u}\mathbf{u}^H. \tag{29}$$

By using the rotation matrix, the rotated basis codeword can also be written as $\mathbf{o}_{rot} = \mathbf{U}_{rot}\mathbf{o}$. The Grassmannian codebook $\mathcal{G}$ is obtained by the complex Grassmannian line packings in [30]. Finally, the codewords of $\mathcal{F}_{\mathbf{h}}$ in (27) can be written as

$$\mathbf{f}_m = \begin{cases} \mathbf{o}_{rot}, & \text{if } m = 1, \\ \mathbf{U}_{rot} \begin{bmatrix} \sqrt{1-\epsilon_{\mathbf{h}}^2} \\ \epsilon_{\mathbf{h}} \mathbf{w}_{m-1} \end{bmatrix}, & \text{if } m = 2, 3, \ldots, N_p. \end{cases} \tag{30}$$

The receiver constructs the codebook $\mathcal{F}_{\mathbf{h}}$, shown in Figure 2, after estimating the channel vector $\mathbf{h}$. It is assumed that the receiver has the knowledge of the LOS channel vector $\overline{\mathbf{h}}$, which can be obtained by estimating the array steering vectors in vertical and horizontal directions. Then, each user finds the best codeword index m^* given by

$$m^* = \underset{m \in \{1,2,\ldots,N_p\}}{\arg\max} |\mathbf{h}^H \mathbf{f}_m|. \tag{31}$$

Through the feedback channel, the receiver sends the best codeword index m^* and the radius $\epsilon_{\mathbf{h}}$ to the transmitter. Then, the transmitter constructs the best codeword $\mathbf{f}_{m^*}$ based on the index m^* by using (30) and the Grassmannian codebook $\mathcal{G}$, which is also known by the transmitter.

The codebook design depends substantially on the strength

of the LOS channel component. If the LOS component predominates significantly over the NLOS component, the radius $\epsilon_{\mathbf{h}}$ will be small and the non-basis codewords will be close to the basis codeword $\mathbf{o}_{rot}$.

V. Simulation Results

In this section, we evaluate the performance of the proposed user selection algorithm for the NOMA-based HAPS system in Section V-A and that of the polar-cap codebook design for HAPS communications in Section V-B.

A. Proposed User Selection Algorithm

We now analyze the sum data rate performance of the proposed user selection algorithm under different circumstances. The simulation environment is constructed such that the L users are uniformly distributed over a cell with a radius of r on the ground, and the HAPS is located 20 km above the cell. Then, L_u users are selected among the L users such that $L_u/2$ users are among the high-rate users and $L_u/2$ users are among the low-rate users. The NOMA pairs are determined and the NOMA-based HAPS system is constructed via power-domain multiplexing. The simulation parameters are shown in Table I.

Table I: Simulation parameters for the NOMA-based HAPS system.

Parameter	Value
f_c	2.4 GHz
B	10 MHz
N_0	-169 dBm/Hz
K_ℓ, $\forall \ell$	10 dB
d_r, d_c	$\lambda/2$
θ_0	30°
δ	10°
μ_ℓ, $\forall \ell$	0°
κ	5
β_{high}	0.7
β_{low}	0.3
n	12

The sorting-based user pairing algorithm in [16] is utilized as a benchmark for the proposed algorithm in this scenario. The sorting-based user selection and pairing algorithm is called *modified sorting* and is described as follows: The set of high-rate users $\mathcal{K}_{\text{high}}$ and the set of low-rate users $\mathcal{K}_{\text{low}}$ are determined before utilizing the sorting algorithm. Then, the users in $\mathcal{K}_{\text{high}}$ are sorted on the basis of their channel norms, and $L_u/2$ users with the highest channel norms are selected to serve. Similarly, $L_u/2$ low-rate users are selected among $\mathcal{K}_{\text{low}}$ by sorting. The NOMA pairs are constructed by pairing the high-rate user with the i^{th} highest channel norm in $\mathcal{K}_{\text{high}}$ with the low-rate user with the i^{th} highest channel norm in $\mathcal{K}_{\text{low}}$.

On the other hand, the improved user clustering algorithm (IUCA) in [17] is utilized as another benchmark. For the considered scenario, the IUCA-based user selection and pairing algorithm is called *modified IUCA* and is described as follows: The selection of the high-rate users is performed exactly the same as in the modified sorting algorithm. Then, the low-rate users are selected in two steps as in [17]. In the first step, all channel gain correlations and differences between the selected high-rate users and all low-rate users are obtained. In the second step, the set of candidate low-rate users $\mathcal{P}(m)$ for the high-rate user m is determined based on a channel gain correlation metric, ς and step size, $\Delta \in (0, 1)$. Then, the user q from $\mathcal{P}(m)$ is selected by maximizing the channel gain difference. Thus, the NOMA pair is constructed as (m, q).

In Figure 3, the sum data rate results of the proposed algorithm are provided with respect to the total transmit power for two different cell radius ranges, i.e., $r = 50$ km and $r = 100$ km, and two different antenna dimensions, i.e., 8×8 and 8×16. The system includes $L = 200$ users in total, and $L_u = 32$ active users are served by applying the proposed algorithm. The figure shows that utilizing a higher number of antennas in the HAPS increases the sum data rate by achieving antenna gain. The results also show that the smaller the cell radius, the higher the sum data rate since the effect of large-scale fading decreases as the distance between the user and the HAPS decreases.

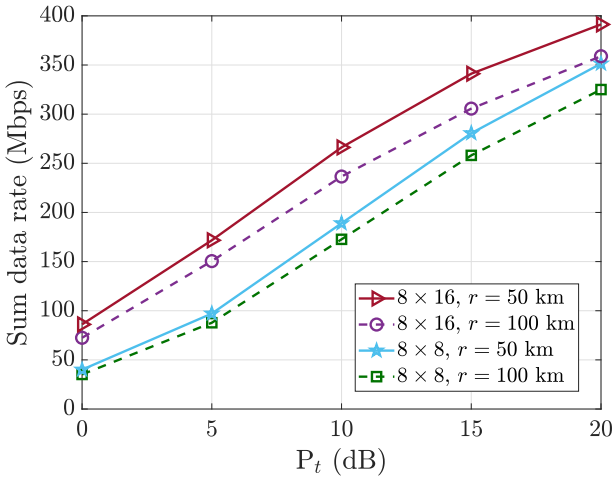


Figure 3: Sum data rate results of the proposed algorithm for $L = 200$ and $L_u = 32$.

Figure 4 demonstrates the sum data rate comparison of different user selection algorithms for an 8×8 array and $r = 100$ km. For the modified IUCA in [17], ς is set as 0.5 or 0.2 when $\Delta = 0.01$. The figure shows that the proposed user selection algorithm outperforms the modified sorting algorithm and the modified IUCA for both ς values. The modified sorting algorithm takes into account the channel gains only. However, considering the correlation between the high-rate and low-rate users instead of only the channel gains enhances the beamforming gain, improving the sum data rate. On the other hand, the modified IUCA in [17] iteratively forms the clusters by considering both the channel correlation and the channel gain difference between users. However, the clustering strictly depends on the set of candidate users, which is constructed based on ς. If the low-rate users in the candidate set of the m^{th} user are already assigned to the other high-rate users,

This article has been accepted for publication in IEEE Transactions on Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVT.2022.3220647

then the m^{th} user will not be paired with the adequate low-rate user, which degrades the overall performance.

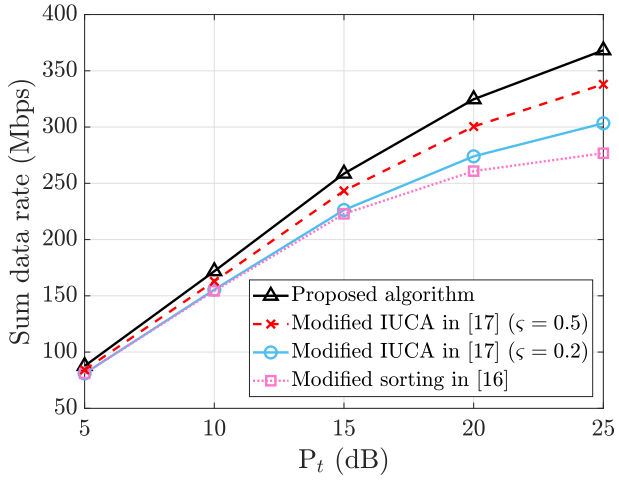


Figure 4: Sum data rate comparison of user selection schemes for 8×8 array, $L = 200$, $L_u = 32$, and $r = 100$ km.

In Figure 5, the effect of the number of users (L) is investigated for a fixed L_u. For the modified IUCA, ς and Δ are set as 0.5 and 0.01, respectively. The simulation results demonstrate that the higher number of users provides more degrees of freedom for all algorithms. In other words, selecting a fixed number of users from a set of users with higher cardinality improves the performance. However, the modified sorting algorithm can not reach the performance of the proposed algorithm for any L value. On the other hand, the proposed algorithm achieves a 8% and 24.4% higher sum data rate compared to the modified IUCA and modified sorting algorithm, respectively, when $L = 200$. The results confirm the feasibility and efficiency of the proposed algorithm especially when a large number of users exists in the coverage area of the HAPS.

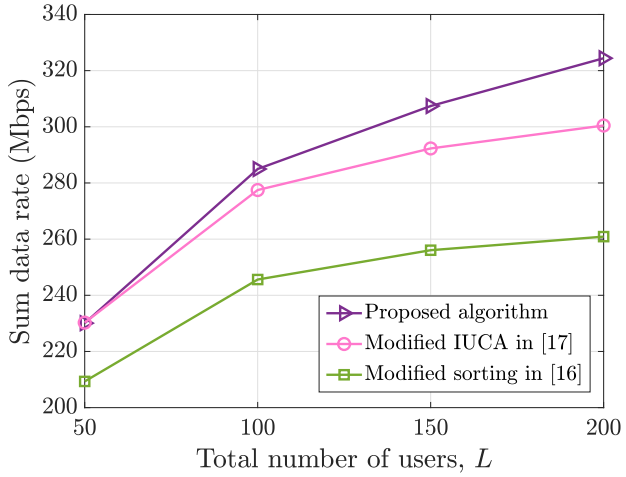


Figure 5: Sum data rate versus L for 8×8 array, $L_u = 32$, $r = 100$ km, and P$_t = 20$ dB.

In Figure 6, the effect of the number of active users (L_u) is discussed when the total number of users (L) is fixed. The results are provided in terms of data rate per user to clarify the effect of L_u on the algorithms. The modified IUCA algorithm is performed for $\varsigma = 0.5$ and $\Delta = 0.01$. The figure shows that the data rate per user results for all algorithms improve when L_u decreases, that is, when $L \gg L_u$. Because small L_u for a fixed L ensures the highly correlated high-rate and low-rate user's pairing in the proposed algorithm. Similarly, the correlation between the candidate users and the selected high-rate users can increase in the modified IUCA algorithm when L_u decreases. On the other hand, small L_u enables the modified sorting algorithm to select only the users with higher channel gain, and the users with poorer channels are excluded. Therefore, it ensures the pairing of the users having higher channel gains. Nevertheless, the proposed algorithm provides almost 2.5 Mbps and 10.5 Mbps higher data rate per user compared to the modified IUCA and modified sorting algorithm, respectively, when $L_u = 8$.

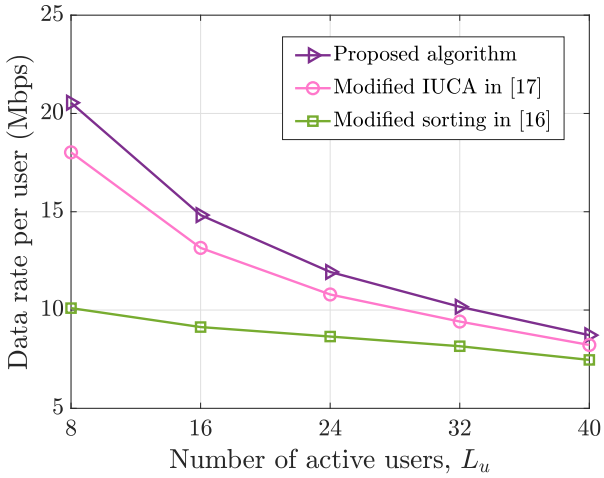


Figure 6: Data rate per user versus L_u for 8×8 array, $L = 200$, $r = 100$ km, and P$_t = 20$ dB.

In Table II, the average data rate per user results for the low-rate users are given for the proposed, modified IUCA ($\varsigma = 0.5$ and $\Delta = 0.01$) and modified sorting algorithms. According to the results, the proposed algorithm significantly improves the average data rate of the low-rate users. The average data rate of the high-rate users is 15.91 Mbps for all three algorithms. Because the way of selecting the high-rate users is exactly the same for these three algorithms. On the other hand, pairing the selected high-rate users with the low-rate users is performed differently. Since the proposed algorithm purely considers the correlation information, the low-rate users do not suffer from very low data rates, contrary to the modified IUCA and modified sorting algorithms.

B. Polar-Cap Codebook (PCC) Design

Here, we present the performance results of the PCC design referred to in Section IV compared to the Grassmannian

Table II: Average data rate per user results for the low-rate users, 8×8 array, $L = 200$, $L_u = 32$, $r = 100$ km, and $\mathrm{P}_t = 20$ dB.

Parameter	Algorithm	Value (Mbps)
$\overline{\mathrm{R}}_{\mathrm{low}}$	proposed	4.28
$\overline{\mathrm{R}}_{\mathrm{low}}$	modified IUCA	2.82
$\overline{\mathrm{R}}_{\mathrm{low}}$	modified sorting	0.37

codebook in [30]. The average beamforming gain is calculated by

$$\eta = \mathbb{E}\left\{ \frac{\left|\mathbf{h}^H \mathbf{f}_{m^*}\right|^2}{\|\mathbf{h}\|^2} \right\}, \tag{32}$$

where $\mathbf{f}_{m^*}$ is the best codeword of either the PCC or the Grassmannian codebook.

In Figure 7, we plot the average beamforming gain with respect to the Rician factor (K) for 4×8 and 8×8 arrays when the number of codewords is $N_p = 2^7 = 128$. In the presence of perfect CSI knowledge, unity average beamforming gain is obtained for all K values where $\mathbf{f}_{m^*} = \mathbf{h}$. However, the performance of the PCC depends on the Rician factor (i.e., the strength of the LOS channel). We observed that a large Rician factor, which is the case for HAPS-to-ground channels, improves the performance of the PCC since the PCC design depends on the strength of the LOS component. For large K values, the LOS component dominates relative to the NLOS component, and the radius $\epsilon_{\mathbf{h}}$ is small. Therefore, the quantization of the channel is more effective. By contrast, the less dominant LOS component leads to a greater $\epsilon_{\mathbf{h}}$, which means that the correlation between $\mathbf{g}$ and $\mathbf{o}_{rot}$ is decreased. Also, the PCC considerably outperforms the Grassmannian codebook due to its adaptation to the LOS channel. On the other hand, the performances of both codebooks enhance when the number of antennas decreases, and the number of feedback bits is the same.

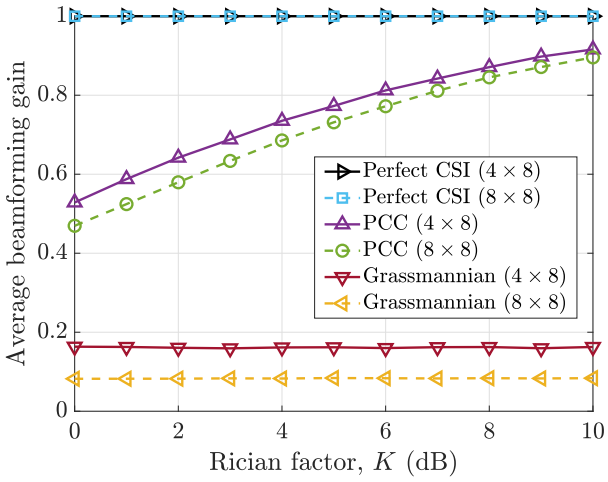


Figure 7: Average beamforming gain comparison of the PCC and the Grassmannian codebook for $N_p = 128$, 4×8 and 8×8 arrays.

Figure 8 demonstrates the sum data rate performance of the limited-feedback NOMA-based multi-user HAPS system applying the proposed user selection algorithm. Part-1 in the proposed algorithm is performed while assuming that the perfect channel quality information (CQI) is available at the transmitter. In part-2, the correlation coefficients are calculated based on the quantized channel direction information (CDI) through the PCC and Grassmannian codebook. The results show that the PCC achieves remarkable performance in the NOMA-based multi-user HAPS system.

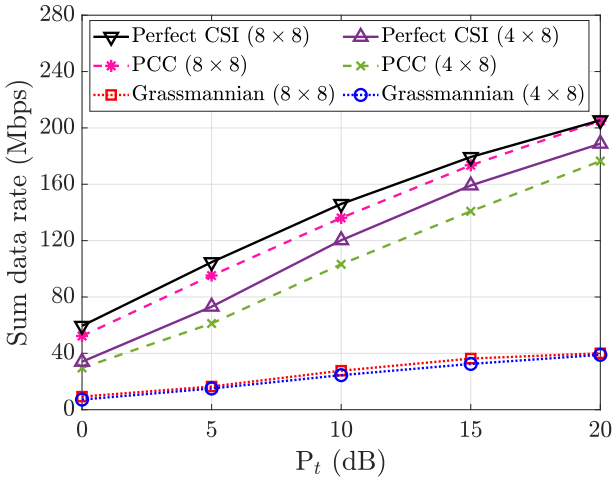


Figure 8: Sum data rate results of the limited-feedback NOMA-based multi-user HAPS system for $N_p = 128$, $L = 100$, $L_u = 16$, $r = 100$ km, 4×8 and 8×8 arrays.

VI. Conclusion

This paper proposed a joint user selection and user pairing algorithm for NOMA-based multi-user HAPS systems. The scenario involved an equal number of high-rate and low-rate users to apply the power-domain NOMA scheme. In the proposed algorithm, we selected the high-rate users having greater channel gains. The channel gain-based selection can maximize the sum data rate with considerably lower complexity for the large coverage areas. Accordingly, we presented a correlation-based user pairing algorithm. Since the HAPS channel experiences highly correlated users due to the predominance of the LOS component, the proposed algorithm maximizes the beamforming gain by considering a correlation-based user pairing. The simulation results confirmed the superiority of the proposed algorithm for the NOMA-based multi-user HAPS system under various conditions. By contrast with the sorting-based user selection, the proposed algorithm considered both channel gains and channel correlations. The proposed algorithm also outperforms the IUCA-based user selection, which considers both the channel correlation and the channel gain difference. In addition, we investigated the polar-cap codebook design for HAPS systems with a limited-rate feedback channel. The polar-cap codebook, which was adapted to the HAPS channel, utilized the LOS component of the HAPS channel. The simulation results demonstrated

that the performance of the limited-feedback NOMA-based multi-user HAPS system utilizing the PCC is fairly close to the performance of the perfect CSI when the proposed user selection algorithm is applied. The PCC significantly outperforms the Grassmannian codebook, which reveals the feasibility of the PCC with the proposed algorithm. The NOMA-based HAPS system conducted in this paper can be extended to a NOMA-based multiple HAPS system as a future work.

REFERENCES

[1] S. Karapantazis and F. Pavlidou, "Broadband Communications via High-Altitude Platforms: A Survey," *IEEE Communications Surveys & Tutorials*, vol. 7, no. 1, pp. 2-31, First Qtr. 2005.

[2] A. Malinowski and R. J. Zieliński, "High Altitude Platform: Future of Infrastructure," *International Journal of Electronics and Telecommunications*, vol. 56, no. 2, pp. 191-196, June 2010.

[3] G. Karabulut Kurt, M. G. Khoshkholgh, S. Alfattani, A. Ibrahim, T. S. J. Darwish, M. S. Alam, H. Yanikomeroglu, and A. Yongacoglu, "A Vision and Framework for the High Altitude Platform Station (HAPS) Networks of the Future," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 729-779, Second Qtr. 2021.

[4] Z. Jia, M. Sheng, J. Li, D. Zhou, and Z. Han, "Joint HAP Access and LEO Satellite Backhaul in 6G: Matching Game-Based Approaches," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 4, pp. 1147-1159, April 2021.

[5] X. Cao, B. Yang, C. Yuen, and Z. Han, "HAP-Reserved Communications in Space-Air-Ground Integrated Networks," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 8, pp. 8286-8291, August 2021.

[6] S. Shah, M. Siddharth, N. Vishwakarma, R. Swaminathan, and A. S. Madhukumar, "Adaptive-Combining-Based Hybrid FSO/RF Satellite Communication With and Without HAPS," *IEEE Access*, vol. 9, pp. 81492-81511, June 2021.

[7] S. R, S. Sharma, N. Vishwakarma, and A. S. Madhukumar, "HAPS-Based Relaying for Integrated Space–Air–Ground Networks With Hybrid FSO/RF Communication: A Performance Analysis," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 3, pp. 1581-1599, June 2021.

[8] 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 Communications Magazine*, vol. 53, no. 9, pp. 74-81, September 2015.

[9] S. M. R. Islam, N. Avazov, O. A. Dobre, and K. Kwak, "Power-Domain Non-Orthogonal Multiple Access (NOMA) in 5G Systems: Potentials and Challenges," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 721-742, Second Qtr. 2017.

[10] M. S. Alam, G. K. Kurt, H. Yanikomeroglu, P. Zhu, and N. D. Đào, "High Altitude Platform Station Based Super Macro Base Station Constellations," *IEEE Communications Magazine*, vol. 59, no. 1, pp. 103-109, January 2021.

[11] K. Tashiro, K. Hoshino, and A. Nagate, "Cylindrical Massive MIMO System for HAPS: Capacity Enhancement and Coverage Extension," in *Proc. IEEE Vehicular Technology Conference (VTC2021-Spring)*, pp. 1-6, April 2021.

[12] Z. Lian, L. Jiang, C. H. He, and D. He, "User Grouping and Beamforming for HAP Massive MIMO Systems," *IEEE Wireless Communications Letters*, vol. 8, no. 3, pp. 961-964, June 2019.

[13] Q. Zhang, Q. Xi, C. He, and L. Jiang, "User Clustered Opportunistic Beamforming for Stratospheric Communications," *IEEE Communications Letters*, vol. 20, no. 9, pp. 1832-1835, September 2016.

[14] P. Ji, L. Jiang, C. He, Z. Lian, and D. He, "Energy-Efficient Beamforming for Beamspace HAP-NOMA Systems," *IEEE Communications Letters*, vol. 25, no. 5, pp. 1678-1681, May 2021.

[15] M. A. Azzahra and Iskandar, "NOMA Signal Transmission over Millimeter-wave Frequency for Backbone Network in HAPS with MIMO Antenna," in *Proc. IEEE 13th International Conference on Telecommunication Systems, Services, and Applications (TSSA)*, pp. 186-189, October 2019.

[16] H. Zhang, D. Zhang, W. Meng, and C. Li, "User Pairing Algorithm with SIC in Non-Orthogonal Multiple Access System," in *Proc. IEEE International Conference on Communications (ICC)*, pp. 1-6, May 2016.

[17] Z. Liu, L. Lei, N. Zhang, G. Kang, and S. Chatzinotas, "Joint Beamforming and Power Optimization With Iterative User Clustering for MISO-NOMA Systems," *IEEE Access*, vol. 5, pp. 6872-6884, May 2017.

[18] C. Göztepe, B. Özbek, and G. K. Kurt, "Design and Implementation of Spatial Correlation-Based Clustering for Multiuser MISO-NOMA Systems," *IEEE Communications Letters*, vol. 25, no. 1, pp. 254-258, January 2021.

[19] S. Liu, C. Zhang, and G. Lyu, "User Selection and Power Schedule for Downlink Non-Orthogonal Multiple Access (NOMA) System," in *Proc. IEEE International Conference on Communication Workshop (ICCW)*, pp. 2561-2565, June 2015.

[20] D. J. Love, R. W. Heath, V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An Overview of Limited Feedback in Wireless Communication Systems," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 8, pp. 1341-1365, October 2008.

[21] P. Xia and G. B. Giannakis, "Design and Analysis of Transmit-Beamforming based on Limited-Rate Feedback," *IEEE Transactions on Signal Processing*, vol. 54, no. 5, pp. 1853–1863, May 2006.

[22] B. Özbek and D. Le Ruyet, *Feedback Strategies for Wireless Communication.* New York, NY, USA: Springer, 2014.

[23] D. J. Love, R. W. Heath, and T. Strohmer, "Grassmannian Beamforming for Multiple-Input Multiple-Output Wireless Systems," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2735-2747, October 2003.

[24] C. K. Au-yeung and D. J. Love, "On the Performance of Random Vector Quantization Limited Feedback Beamforming in a MISO System," *IEEE Transactions on Wireless Communications*, vol. 6, no. 2, pp. 458-462, February 2007.

[25] V. Raghavan, R. W. Heath, and A. M. Sayeed, "Systematic Codebook Designs for Quantized Beamforming in Correlated MIMO Channels," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 7, pp. 1298-1310, September 2007.

[26] S. H. Hong, S. Kim, J. Choi, and W. Choi, "Polar-Cap Codebook Design for MISO Rician Fading Channels With Limited Feedback," *IEEE Wireless Communications Letters*, vol. 10, no. 4, pp. 730-734, April 2021.

[27] A. Z. Cengiz, S. T. Basaran, B. Özbek, G. K. Kurt, and H. Yanikomeroglu, "Approximation of Correlation Matrix for High Altitude Platform Stations (Yüksek İrtifa Platform İstasyonları için Korelasyon Matrisinin Yaklaşımı)," in *Proc. IEEE Conference on Signal Processing and Communications Applications (SIU)*, pp. 1-4, June 2021.

[28] E. Süli and D. F. Mayers, *An Introduction to Numerical Analysis*. Cambridge: Cambridge University Press, 2003.

[29] K.-L. Chung and W.-M. Yan, "The Complex Householder Transform," *IEEE Transactions on Signal Processing*, vol. 45, no. 9, pp. 2374-2376, September 1997.

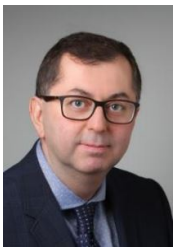
[30] A. Medra and T. N. Davidson, "Flexible Codebook Design for Limited Feedback Systems Via Sequential Smooth Optimization on the Grassmannian Manifold," *IEEE Transactions on Signal Processing*, vol. 62, no. 5, pp. 1305-1318, March 2014.

[31] J. Choi, B. Clerckx, N. Lee, and G. Kim, "A New Design of Polar-Cap Differential Codebook for Temporally/Spatially Correlated MISO Channels," *IEEE Transactions on Wireless Communications*, vol. 11, no. 2, pp. 703-711, February 2012.

İrem Cumalı received her B.Sc. and M.Sc. degrees in electronics and communication engineering from Izmir Institute of Technology, Izmir, Türkiye, in 2017 and 2020, respectively. She is currently pursuing a Ph.D. degree in electronics and communication engineering. She has been a Research Assistant at the Department of Electrical and Electronics Engineering at Izmir Institute of Technology since 2018. She is also working as a researcher in an international project supported by European Union under Horizon 2020-MSCA-RISE Programme. Her research interests include wireless communications, 5G/B5G, cell-free massive MIMO networks, high altitude platform station communications, and non-orthogonal multiple access.

Berna Özbek (Senior Member, IEEE) is currently an Associate Professor in telecommunication with the Department of Electrical and Electronics Engineering, Izmir Institute of Technology, Türkiye, and working in the field of wireless communication systems for more than 15 years. Under her supervision, 15 master's thesis and two Ph.D. dissertations have been completed. She has been awarded as Marie-Curie Intra-European (EIF) Fellow by the European Commission for two years in 2010. She has coordinated one international and four national projects, served as a Consultant for three Eureka-Celtic projects and two national industry driven projects. Currently, she is supervising three Ph.D. and one master's students, conducting one international project under H2020-MSCA-RISE Programme from 2018 to 2023, and coordinating one international project under Horizon Europe MSCA-DN Programme from 2022 to 2027. She has published more than 100 peer-reviewed articles, one book, one book chapter, and two patents. Her research interests include interference management, resource allocation, limited feedback links, device-to-device communications, physical layer security, massive MIMO systems, and mmWave communications.

Güneş Karabulut Kurt received the B.S. degree with high honors in electronics and electrical engineering from Bogazici University, Istanbul, Türkiye, in 2000 and the M.A.Sc. and the Ph.D. degrees in electrical engineering from the University of Ottawa, ON, Canada, in 2002 and 2006, respectively. From 2000 to 2005, she was a Research Assistant with the CASP Group, University of Ottawa. Between 2005 and 2006, she was with TenXc Wireless, Canada. From 2006 to 2008, Dr. Karabulut Kurt was with Edgewater Computer Systems Inc., Canada. From 2008 to 2010, she was with Turkcell Research and Development Applied Research and Technology, Istanbul. Between 2010 and 2021, she was with Istanbul Technical University. She is currently an Associate Professor of Electrical Engineering at Polytechnique Montréal, Montréal, QC, Canada. She is a Marie Curie Fellow and has received the Turkish Academy of Sciences Outstanding Young Scientist (TÜBA-GEBIP) Award in 2019. In addition, she is an adjunct research professor at Carleton University. She is currently serving as an associate technical editor of the *IEEE Communications Magazine*, an associate editor of *IEEE Communication Letters*, an associate editor of *IEEE Wireless Communications Letters*, and an area editor of *IEEE Transactions on Machine Learning in Communications and Networking*. She is a member of the IEEE WCNC Steering Board. She is serving as the secretary of IEEE Satellite and Space Communications Technical Committee and also the chair of the IEEE special interest group entitled "Satellite Mega-constellations: Communications and Networking". She is a Distinguished Lecturer of Vehicular Technology Society Class of 2022.



Halim Yanikomeroglu (Fellow, IEEE) is currently a Professor with the Department of Systems and Computer Engineering, Carleton University, Ottawa, Canada. His primary research interest includes wireless communications and networks. His research group has made substantial contributions to 4G and 5G wireless technologies. His group's current focus is the aerial (UAV and HAPS) and satellite networks for the 6G and beyond-6G era. His extensive collaboration with industry resulted in 39 granted patents. He is a fellow of Engineering Institute of Canada (EIC) and Canadian Academy of Engineering (CAE). He has received several awards for his research, teaching, and service, including the IEEE ComSoc Fred W. Ellersick Prize, in 2021; the IEEE VTS Stuart Meyer Memorial Award, in 2020; and the IEEE ComSoc Wireless Communications Technical Committee Recognition Award, in 2018. He is serving as the Chair for the IEEE ComSoc Technical Committees Board and the Chair for the IEEE Wireless Communications and Networking Conference (WCNC) Steering Committee. He has served as the General Chair or the TP Chair for several conferences, including three WCNCs and two VTCs. He has also served as the Chair for the IEEE's Technical Committee on Personal Communications. He is a Distinguished Speaker for both IEEE Communications Society and IEEE Vehicular Technology Society.