

System Design and Parameter Optimization for Remote Coverage from NOMA-based High-Altitude Platform Stations (HAPS)

Sidrah Javed, *Member, IEEE* and Mohamed-Slim Alouini, *Fellow, IEEE*

Abstract—Stratospheric solar-powered high-altitude platform station (HAPS) can provide line-of-sight (LoS) communications to the ground users in its ultra-wide coverage area. This paper addresses the challenge of HAPS communication system design especially the access link. We propose to divide the ground users into multiple user-groups and serve each group by a high-density dynamically steerable spotbeam, generated by the phased array antennas mounted on HAPS. We employ time-division multiplexing (TDM) to serve different user groups and non-orthogonal multiple access (NOMA) to simultaneously serve all users within a usergroup. We formulate user grouping problem as an equivalent geometric disk cover (GDC) problem and beam optimization problem as a minimum enclosing circle (MEC) problem. We present the optimization framework to jointly design user grouping, user association, beam optimization, and power allocation aiming at sum rate maximization while guaranteeing the quality-of-service (QoS) with limited power budget. System performance is assessed using the key metrics such as signal-to-interference noise ratio (SINR), achievable data rate, average energy efficiency (AEE), average spectral efficiency (ASE), user fairness and outage probability. We observe upto 42% reduction in required groups, 5dB increase in received SINR, 37.5% improvement in energy efficiency, 57.9% rise in spectral efficiency, 22% enhanced user fairness, 65% surge in achievable data rates and ten-folds reduction in outage with the proposed optimization framework over conventional schemes using system-level simulations. Our findings reveal the significance of joint design of system parameters for enhanced performance, optimum energy utilization, and resource allocation.

Index Terms—Non-orthogonal multiple access, unmanned aerial vehicles, high-altitude pseudo-satellites, 6G, user grouping, user association, beam location, beam optimization, resource allocation, sum rate maximization, and outage performance.

I. INTRODUCTION

In our increasingly interconnected world, the demand for seamless, high-capacity, secure, and cost-effective wireless communication systems continues to grow [1], [2]. With billions of people in remote, rural, and under-served areas facing limited or no connectivity, the challenge of bridging this digital divide has never been more pressing [3]. Traditional terrestrial and satellite systems, though instrumental, have limitations in coverage, latency, and deployment costs, especially for sparsely populated regions [4]. Here, aerial communications present a transformative solution, promising to extend connectivity with improved coverage, low latency, and efficient deployment [5]. They have the capability to augment/complement the existing terrestrial and satellite communication infrastructure, aligning with the UN Sustainable Development Goals 2030 for global connectivity [6].

Aerial communications can be realized as low-altitude platform station (LAPS) or high-altitude platform station (HAPS). LAPS typically operate in lower troposphere, whereas, HAPS fly in the stratosphere positioned above weather systems and air traffic. LAPS are the preferred choice for quick deployment in emergency scenarios with limited temporal and spatial coverage based on the battery capacity and flying altitudes, respectively. On the other hand, HAPS are preferred for long-endurance, ultra-wide coverage, ubiquitous connectivity, and resilience along with other dividends of aerial communication platforms [7]. Solar-powered HAPS offer green communications while cruising in a station-keeping trajectory for several months given favorable conditions in lower stratosphere [8]. These unmanned aircrafts maintain a quasi-stationary position by flying in a controlled, small circular pattern around a specific point on the ground. This ensures continuous coverage over a target geographical region, with navigation systems compensating for any atmospheric disturbances. To ground-based users, the HAPS appears stationary, effectively maintaining a fixed service area even while in motion [9], [10].

To fully leverage the potential of HAPS in providing high-quality connectivity, several technical challenges must be addressed. Significant progress has been made in system design for HAPS, such as utilizing steerable adaptive antenna arrays in multi-cell HAPS communication systems to enhance user coverage and service adaptability [11]. Energy-efficient beamforming strategies over Rician fading channels have also been proposed for HAPS-based NOMA systems, demonstrating improvements in spectral efficiency for beamspace communications [12]. Meanwhile, using NOMA in multiple-input multiple-output (MIMO)-enabled HAPS systems over millimeter-wave frequencies has emerged as a promising method for effective resource management, particularly relevant in backbone networks with high demand for concurrent user connections [13]. Additionally, secure communication has been examined through deterrence strategies that account for environmental and operational risks in aerial networks [14]. Another work focuses on resource optimization in integrated satellite-airborne-terrestrial networks to enable seamless global connectivity [6]. User grouping and beamforming methods continue to evolve to accommodate the unique demands of HAPS, as seen in schemes leveraging average chordal distances and reduced-dimensional statistical eigenmodes, which enhance the reliability and scalability of user-grouping mechanisms for massive MIMO-HAPS applications [15]. Another study introduces joint user association and beamforming for integrated satellite-aerial-ground networks, achieving improved network coordination and capacity [16]. Recently, elevation-angle-based user association and power

S. Javed is with Department of Engineering, Durham University, United Kingdom. This research was conducted when she was a post-doctorate researcher at King Abdullah University of Science and Technology (KAUST) and M.-S. Alouini is with CEMSE Division, KAUST, Saudi Arabia. E-mail: sidrah.javed@durham.ac.uk and slim.alouini@kaust.edu.sa

allocation strategies for NOMA have been proposed to mitigate inter-cell interference in LAPS, particularly in scenarios such as cellular-connected UAV networks [17], [18]. Although significant progress has been made on individual challenges such as error rate minimization, power efficiency, and interference management in HAPS-based communication systems [19], [20], a comprehensive approach remains under-explored. Specifically, there is a gap in research addressing the joint optimization of user grouping, association, beamforming, and multiple access schemes for stratospheric HAPS—a holistic strategy that could substantially enhance network throughput, improve system efficiency, and reduce outage rates.

This research focuses on refining HAPS communication systems through effective user grouping and association, beam optimization, and advanced resource allocation techniques. We present a comprehensive HAPS communication system model, where users are clustered into groups and phased array antennas dynamically steer beams towards different user groups [11]. The hybrid multiplexing employs TDM to steer beams towards different user groups whereas utilizes NOMA superposition coding to serve all the users in one user group. Inter-beam interference is mitigated using TDM between spot beams. Effective user grouping balances resource distribution while maintaining high-quality service across vast geographic areas. Similarly, robust user association techniques are necessary to ensure that users are connected through the optimal serving beam. Tailoring beam patterns enhances targeted coverage accommodating both densely populated and dispersed areas. Moreover, efficient access techniques and resource allocation will provide improved spectrum utilization and seamless connectivity. This research aims to develop and optimize these aspects of HAPS systems, creating innovative solutions for joint optimization of user grouping, association, beam optimization, and resource allocation. We present a novel optimization framework with suitable algorithms to address these challenges, while ensuring user QoS, fairness, and efficiency. The key contributions of this research include:

- User grouping is equivalent to the GDC problem in computational geometry. We reformulate this problem and present Algorithm 1 to determine the minimum number of groups and their locations.
- The user association problem identifies the serving spot beam for each user. A greedy algorithm is proposed to maximize data rates by improving the received SINR.
- Subsequently, we carry out beam optimization to determine the optimal beams to serve each user group, enhancing directivity, power density, and antenna gain. By framing this problem in analogy with MEC problem, we can employ Welzl's algorithm for efficient solution.
- Next, we present closed-form solution to the NOMA power allocation problem assuming successful successive interference cancellation (SIC) and decoding at all users.
- Finally, the outage probability, energy efficiency, spectral efficiency and user fairness of the HAPS communication system are analyzed to study the effects of the proposed approaches and algorithms on system performance as opposed to the existing schemes.

TABLE I: Frequently used symbols

Definition	Symbol	Definition	Symbol
HAPS coverage radius	R_H	HAPS altitude	H
Total ground users	K	User locations	\mathbf{u}_k
Number of usergroups	M	One-sided HPBW	θ_m^{3dB}
Beamradius	r_m	Beam center	\mathbf{w}_m
Diameter of antenna arrays	D	Diameter of circular trajectory	d
User beam angle	θ_i^m	User elevation angle	ψ_i^m
HAPS-User distance	d_i^m	User association indicator	x_k^m
Transmission power budget	P_t	Power allocation factor	α_k^m
Information signal	s_k^m	Superposed transmit signal	v_m
Channel coefficient	h_i^m	Receiver thermal noise	w_m^m
Thermal noise variance	σ_f^2	Small-scale fading	g_i^m
Transmit antenna gain	G_t^m	FSPL b/w User and HAPS	$L(d_i^m)$
Peak antenna gain	G_0^m	Antenna aperture efficiency	η
Career wavelength	λ	Shadowing	$\chi_{\sigma_s}^{CIPL}$
User's received SINR	γ_i^m	Shadowing standard deviation	σ_s
Channel bandwidth	B	Noise figure	NF
User achievable rate	R_i^m	User target rate	R_j^m
Transmit SNR	ρ_m	Circuit power	P_c
User energy efficiency	EE $_T^m$	User spectral efficiency	SE $_T^m$
Average energy efficiency	AEE	Average spectral efficiency	ASE
User outage probability	OP m	OMA outage probability	OMA_OP

The presented innovations in HAPS technology have the potential to extend reliable, high-quality connectivity to millions worldwide, creating transformative impacts on sustainable and accessible global communication.

The rest of the report is organized as: Section II describes different aspects of the system under consideration i.e., user grouping, user association, and NOMA. In section III, we present the HAPS communication propagation model for link budgeting while incorporating fading and shadowing in the stratospheric communications. Next, we evaluate the system performance in terms of SINR, sum rate, system efficiency, and outage probability in section IV. Moreover, we formulate the optimization problem to maximize system throughput with the given QoS and power constraints. This section V details the optimization framework, proposed algorithms, and design guidelines to achieve the desired system performance. Later, insightful numerical analysis is carried out in section VI, to quantify the gains obtained with the proposed scheme over the conventional schemes, followed by the comprehensive conclusion in Section VII.

In this paper, x , \mathbf{x} , and \mathbf{X} denote scalar, vector, and matrix, respectively. Capital unbolded symbols represent significant scalar variables, such as R_H , H , and D . The inverse and absolute value of a scalar x are denoted as x^{-1} and $|x|$, respectively. The symbols \sum_k , $\|\mathbf{w}\|$, and $[G]_{dB}$ represent summation over index k , the l_2 -norm of vector \mathbf{w} , and the decibel value of gain G . Events are illustrated as $\{R \leq R_{th}\}$, with the probability given by $\Pr\{R \leq R_{th}\}$, while the intersection of events is denoted by \cap . The scalar-valued function $f(x, y | \nu, \sigma)$ depends on the independent variables x and y , conditioned on ν and σ . The exponential and logarithmic functions with base b for variable z are represented as $\exp(z)$ and $\log_b(z)$. A set is represented as $m \in \{1, 2, \dots, M\}$, indicating that m can take any value from 1 to M . The notation $\{m \setminus n\}$ denotes any value of m in this range, excluding n . The notation $R[n]$ indicates the value of the function R at time instance n . Additionally, \tilde{R} and R^* represent the temporary variable and the optimal value of R , respectively.

II. SYSTEM DESCRIPTION

We consider a typical unmanned solar-powered quasi-stationary HAPS at an altitude H over the desired coverage area with radius R_H ranging from 60km to 400km as shown in Fig. 1. The HAPS can operate at altitudes ranging from 18 km to 50 km in the stratosphere, with the preferred range being 18–24 km [7], [8], [21]. This preference is based on favorable environmental conditions [22], such as minimal turbulence for stable flight, lower wind speeds for enhanced energy efficiency, suitable air density to minimize drag, and unobstructed LoS for communication operations. HAPS provides communication services to K ground users over the fourth-generation (4G) long-term evolution or fifth-generation (5G) new radio (NR) air interface via service link and backhaul to the gateway through the feeder link [11]. We suppose that the ground users are on the same horizontal plane with coordinates $\mathbf{u}_k \in \mathbb{R}^2 \ \forall k$ and their locations are known. In rural and remote areas, where clear line-of-sight (LoS) is available, Global Navigation Satellite System (GNSS)-based localization [23] and beam scanning [24] can effectively determine user locations. HAPS assisted GNSS receivers can triangulate user signals or perform beam scanning with phased array antennas to identify locations based on received signal strengths [25]. We further assume that the phased array antennas are mounted at the bottom of HAPS communication panel to serve the ground users with the high-density flexible narrow spot beams. This section describes the user grouping, user association, NOMA scheme and propagation model with the aim to achieve ubiquitous connectivity with effective resource allocation, minimal power consumption and enhanced system performance.

A. User Grouping

The ground users in the coverage area are distributed into M groups and each group is served by a high-density narrow spot beam using time-division. This ensures that all users are located in the main lobe of the beam with zero inter-beam interference. The phased array antennas are capable of generating flexible beams with beamwidth $\theta_m^{3dB} \in \mathbb{R}$, beamradius $r_m \in \mathbb{R}^+$, and beam centers $\mathbf{w}_m \in \mathbb{R}^2$ for all $m \in \{1, 2, \dots, M\}$ as detailed in Fig. 1. It is worth noting that the beam steering may result in slightly elliptical coverage [26], however, our assumption of circular beams favors simplicity in modeling, uniform power distribution, and practical feasibility especially in rural/remote areas. When beam steering angles are small, the deviation from a circular shape is minimal, and adjustments like beamwidth control help maintain a nearly circular footprint. Interestingly, the number of groups or the number of subsequent beams M is a variable and ranges between $1 \leq M \leq K$, reflecting that there can be at least one beam to serve all ground users or at max K beams when each user is served by a separate beam. In essence, the choice of M is a trade-off between the number of simultaneously served users and SINR. The higher value of M indicates large number of narrow beams to serve the ground users. This results in increased SINR with high-density power but shorter/delayed time slots. On the other hand, the smaller value of M signifies few but wide beams for coverage.

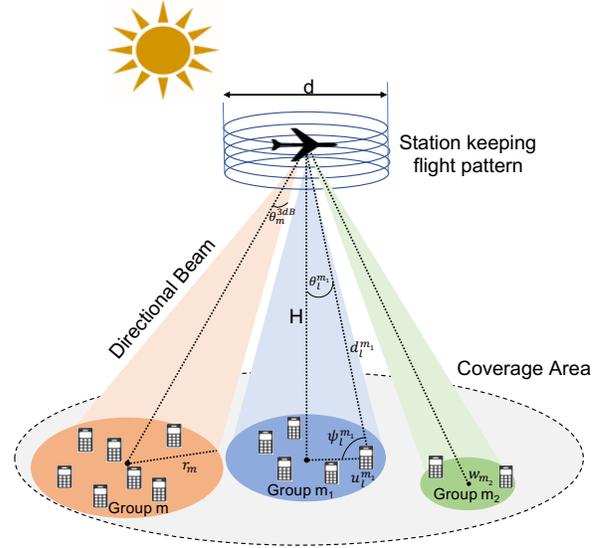


Fig. 1: Aircraft based HAPS Communication System

This renders increased transmission time but reduced SINR. To enhance fairness and reduce interference, each user is served by only one beam. The beam coverage radius can be derived from the half-power beamwidth (HPBW) of m^{th} spot beam as:

$$r_m = H \tan\left(\frac{\theta_m^{3dB}}{2}\right). \quad (1)$$

It is conditioned on $\theta_m^{3dB} \geq 70\pi/D$ where D is the diameter of the antenna arrays. Moreover, the beam centers must be chosen in a way that the entire beam coverage area resides within overall coverage area of the HAPS i.e., $\|\mathbf{w}_m\| + r_m \leq R_H$ for effective resource utilization.

B. User Association

Assuming M beams to serve M user groups with predefined \mathbf{w}_m and r_m , users are associated with groups based on their distances from the center of the beam spots. We define the set of indicators x_k^m to describe users association. The indicator x_k^m is 1 if user k is associated with group m , otherwise 0. It is noteworthy, that each user can only associate with one user group for user fairness and higher efficiency i.e.,

$$\sum_{m=1}^M x_k^m = 1, \forall k. \quad (2)$$

The circular beam coverage area covers radius $r_m \ \forall m$ and all the users in this group are expected to reside in the main lobe of the associated beam to be served simultaneously [27]. Therefore, the associated user must fulfill the following constraint:

$$\sum_{m=1}^M x_k^m \|\mathbf{u}_k - \mathbf{w}_m\| \leq \sum_{m=1}^M x_k^m r_m, \quad \forall k. \quad (3)$$

It is noteworthy that a user l may reside within the radius of user-group j while being associated with the user-group j' , based on the distance inequality $\|\mathbf{u}_l - \mathbf{w}_j'\| \leq \|\mathbf{u}_l - \mathbf{w}_j\|$. Moreover, the apparently overlapping beam radii will not cause any inter-beam interference due to the TDM i.e., these user groups are not served simultaneously but at different time intervals.

C. Non-Orthogonal Multiple Access

Ultra-wide terrestrial coverage of the HAPS can be achieved by splitting the service area into multiple groups, as shown in Fig. 1. Each group is served by a highly directional beam allowing frequency reuse in the neighboring groups for efficient spectrum allocation. We adopt phased antenna array for beamforming and an array controller which are responsible to create the desired beam and steer it in real time, as detailed in [11]. This enables the beam fixation relative to the station-keeping flight pattern for reliable and consistent coverage.

The network consists of K users distributed as per Binomial point process (BPP) which are partitioned into M groups. In each group coverage, the channel gain is expected to vary with the distance from the center as well as on the azimuth angle. The strongest channel gain is available at the center along the boresight direction $\theta = 0$. However, as the distance varies and/or the azimuth direction deviates from the boresight, the performance can be degraded due to the increased path loss and lower antenna beam gain. The striking difference in the channel gains of the users in each group enables us to reap maximum benefits offered by NOMA. Consider the downlink (DL)-NOMA scenario, where m^{th} group is served by a directional beam with superposition coding as ¹

$$v_m = \sum_{k=1}^K \sqrt{P_t \alpha_k^m} s_k^m x_k^m, \quad (4)$$

where P_t is the available power budget for transmission after deducting the aerodynamics, electronics, and night-time operational expenses from the available solar power at a given time as quantified in [28]. Moreover, α_k^m and s_k^m are the fraction of power allocated to and intended information signal for the k^{th} user in m^{th} group, respectively. It is important to highlight that $\sum_{k=1}^K \alpha_k^m \leq 1, \forall m$, in order to limit the power division within given budget. Therefore, using conventional wireless communication model, the received signal at user l in the m^{th} group is given by

$$y_l^m = \underbrace{h_l^m x_l^m \sqrt{P_t \alpha_l^m} s_l^m}_{\text{Desired Signal}} + \underbrace{h_l^m x_l^m \sum_{\substack{k=1 \\ k \neq l}}^K \sqrt{P_t \alpha_k^m} x_k^m s_k^m + w_l^m x_l^m}_{\text{IACI}}, \quad (5)$$

where, h_l^m is the channel gain coefficient between the HAPS array panel and l^{th} user in m^{th} group and w_l^m is the receiver thermal noise modeled as circular symmetric complex Gaussian random variable, i.e., $w_l^m \sim \mathcal{CN}(0, \sigma_l^2)$. The intra-channel interference (IACI) refers to the interference between users within a single user-group sharing the same time and frequency resources. However, inter-channel interference is mitigated by employing TDM between the spot beams of different user groups. In addition, the association parameter x_k^m ensures the incorporation of IACI from the users within the same user-group.

III. PROPAGATION MODEL AND LINK BUDGET

The radio signal propagation from HAPS to the UE undergoes free space path loss (FSPL) and multipath fading due to the significant distance between them and obstacles

¹The transmitted/received signals, channel gains and allocated powers are function of time. However, the time notation is omitted for brevity.

around the UE, respectively. It is noteworthy that the HAPS station-keeping flight does not contribute to the fast fading because there are no moving scatters surrounding the aircraft [29]. Therefore, the propagation loss of the adopted system is modeled as a combination of small-scale and large-scale fading. Hence, the channel coefficient h_l^m can be expressed as follows:

$$h_l^m = \frac{g_l^m \sqrt{G_l^m}}{\sqrt{L(d_l^m)}}, \quad (6)$$

where g_l^m is the small scale fading coefficient between the m^{th} transmitting panel and l^{th} user in m^{th} group, G_l^m is the array gain for the link between m^{th} panel and l^{th} user in its coverage area, and $L(d_l^m)$ is the path loss as a function of d_l^m i.e., the distance between HAPS and l^{th} user in m^{th} group. The computation of these parameters is highlighted in the subsequent sections.

A. Small Scale Multi-path Fading

The received signal comprises of both the LoS and non line-of-sight (NLoS) components pertaining to the HAPS boresight position and independent diffuse multipath reflections from the obstacles. The LoS component is generally deterministic, whereas, the envelope of NLoS component is modeled as a Rayleigh random variable. Hence, the aggregate small-scale multipath fading coefficient g_l^m is modeled as a Rician distributed random variable with power distribution function [30]–[32]

$$f(x | \nu, \sigma_f) = \frac{x}{\sigma_f^2} \exp\left(\frac{-(x^2 + \nu^2)}{2\sigma_f^2}\right) I_0\left(\frac{x\nu}{\sigma_f^2}\right), \quad (7)$$

where I_0 denotes zeroth-order modified Bessel function of the first kind whose shape parameter K_s is defined by the ratio between the average power of LoS component and the average power associated with NLoS multi-path components i.e., $K_s = \nu^2/2\sigma_f^2$.

B. Directivity Gain

The communication panels are equipped with phased array antenna which are responsible for directional beamforming. The transmitter antenna gain G_k^m of user \mathbf{u}_k inside m^{th} group depends on the antenna aperture efficiency η , HPBW of the antenna $\theta_m^{3\text{dB}}$, HAPS altitude H , and the distance of the user \mathbf{u}_k from the center of the beam \mathbf{w}_m [33]

$$[G_k^m]_{\text{dB}} = [G_0^m]_{\text{dB}} - 12 \frac{G_0^m}{\eta} \left(\frac{\theta_k^m}{70\pi}\right)^2, \quad (8)$$

where the peak transmitter antenna gain of the m^{th} beam $G_0^m = \eta (70\pi/\theta_m^{3\text{dB}})^2$ and the beam angle (angle of departure) of the user \mathbf{u}_k is given by

$$\theta_k^m = \tan^{-1}\left(\frac{\|\mathbf{u}_k - \mathbf{w}_m\|}{H}\right). \quad (9)$$

Evidently, the antenna directivity gain reduces while moving away from the boresight position in a horizontal plane.

C. Link Budget

The large scale propagation is characterized as a FSPL model with the distance d_l^m between HAPS and \mathbf{u}_l as $d_l^m = H/\sin \psi_l^m$, where ψ_l^m is the elevation angle of HAPS from \mathbf{u}_l ranging $12\pi/180 \leq \psi_l^m \leq \pi/2$. Any communication link beyond the minimum elevation angle will lose the LoS

path owing to the earth curvature. Evidently, the users in the center group enjoy a larger elevation angle, whereas the edge group users are at relatively smaller elevation angles with $0 \leq \psi_l^e \leq \psi_l^c \leq \pi/2$, where ψ_l^c and ψ_l^e are the elevation angles of users l' and l in the center groups and edge groups, respectively. We employ the close-in path-loss (CIPL) model for the aerial HAPS to compute the received signal path loss $L(d_l^m)$ in dB as [34]

$$[L(H, \psi_l^m)]_{\text{dB}} = 10 \log_{10} \frac{16\pi^2 H^2}{\lambda^2 \sin^2(\psi_l^m)} + \chi_{\sigma_s}^{\text{CIPL}}, \quad (10)$$

where λ is the wavelength corresponding to the carrier frequency and $\chi_{\sigma_s}^{\text{CIPL}}$ is zero-mean shadow fading Gaussian random variable with standard deviation σ_s in dB. The FSPL is computed as the ratio between transmit power and received power. The assumption of isotropic receiver antenna, in the DL-NOMA communication, renders an effective receiver area $\lambda^2/4\pi$ and the received signal intensity is given by the inverse-square law.

IV. PERFORMANCE ANALYSIS

The performance of the DL-NOMA with the given user grouping can be analyzed in the form of user's received SINR, sum rate of all users in the groups and outage performance. Given the superposition coding at the transmitter for each user group, the received signal undergoes SIC to retrieve its own signal based on the known user locations.

A. Signal-to-Interference Noise Ratio

Consider the DL-NOMA where K_m is the set of users in the m^{th} group which are ordered as $\mathbf{u}_1^m, \mathbf{u}_2^m, \dots, \mathbf{u}_{K_m}^m$ depending on their increasing channel strengths. Given this ordered arrangement and SIC at user \mathbf{u}_l^m , it is capable of decoding all users from \mathbf{u}_1^m to \mathbf{u}_{l-1}^m and subtracting these from the received signal. Thus, it can decode its own signal from the resultant by considering the interference from \mathbf{u}_{l+1}^m to $\mathbf{u}_{K_m}^m$ as noise. Therefore, the SINR γ at user \mathbf{u}_l^m is given by

$$\gamma_l^m = \frac{|h_l^m|^2 P_t \alpha_l^m}{|h_l^m|^2 \sum_{k=l+1}^{K_m} P_t \alpha_k^m + \sigma^2}. \quad (11)$$

whereas the SINR of the user with the strongest channel gain is given by

$$\gamma_{K_m}^m = \frac{|h_{K_m}^m|^2 P_t \alpha_{K_m}^m}{\sigma^2}, \quad (12)$$

where the noise power is given as

$$\sigma^2 (\text{dBm}) = -174 + 10 \log_{10}(B) + \text{NF}, \quad (13)$$

with NF denoting the noise figure of the receiver [35] and B depicting the allocated channel bandwidth to serve K_m users simultaneously.

B. Sum Rate Analysis

Assuming perfect receiver channel state information (CSI), we get accurate user-ordering and error-free decoding. Thus, the achievable rate of user \mathbf{u}_l^m is given by

$$R_l^m = B \log_2 [1 + \gamma_l^m]. \quad (14)$$

conditioned on $R_{j \rightarrow l}^m > \tilde{R}_j^m \forall j \leq l$, where \tilde{R}_j^m is the targeted data rate of the j^{th} user in the m^{th} usergroup and $R_{j \rightarrow l}^m$

denoted the rate of the l^{th} user to detect j^{th} user's message, $j \leq l$ i.e.,

$$R_{j \rightarrow l}^m = B \log_2 \left(1 + \frac{|h_l^m|^2 P_t \alpha_j^m}{|h_l^m|^2 \sum_{k=j+1}^{K_m} P_t \alpha_k^m + \sigma^2} \right) \geq \tilde{R}_j^m. \quad (15)$$

Thus, the sum rate R of all users in M groups can be written as $R = \frac{1}{M} \sum_{m=1}^M R_m$, where the sum rate of all users in m^{th} group is given by $R_m = \sum_{l=1}^{K_m} R_l^m$ yielding

$$R = \sum_{m=1}^M R_m = B \sum_{m=1}^M \sum_{l=1}^{K_m} \log_2 [1 + \gamma_l^m], \quad (16)$$

where the received SINR at \mathbf{u}_l^m i.e., γ_l^m in (11) can be expressed using (6) as

$$\gamma_l^m = \frac{\alpha_l^m}{\sum_{k=l+1}^{K_m} \alpha_k^m + \frac{L(d_l^m)}{\varrho_m |g_l^m|^2 G_l^m}}. \quad (17)$$

Likewise $\gamma_{K_m}^m$ in (12) can be manifested using (6) as

$$\gamma_{K_m}^m = \frac{\alpha_{K_m}^m \varrho_m |g_{K_m}^m|^2 G_{K_m}^m}{L(d_{K_m}^m)}, \quad (18)$$

where ϱ_m is the transmit signal-to-noise ratio (SNR) for m^{th} -group i.e., P_t/σ^2 .

C. System Efficiency

The system efficiency of the NOMA based HAPS communication system with the proposed user grouping, user association and beam optimization based on the hybrid multiplexing model can be evaluated in terms of energy efficient, spectrum efficiency and user fairness. The energy efficiency (EE) of the user l in the m^{th} user group can be investigated using

$$\text{EE}_l^m = \frac{R_l^m}{\alpha_l^m P_t + P_c}, \quad (19)$$

where P_c is the circuit power of the system under consideration. The EE_l^m is measured in bits/Joules i.e., a higher value of EE_l^m indicates the higher amount of data in bits that can be sent with minimal energy consumption. The overall energy efficiency of the system can be seen as the AEE of all the users in the coverage area i.e., $\text{AEE} = (KM)^{-1} \sum_{m=1}^M \sum_{l=1}^{K_m} \text{EE}_l^m$.

On the other hand, the spectrum efficiency describes the amount of data transmitted over a given spectrum with minimum transmission errors. Assuming the perfect decoding order and SIC, we can write the spectral efficiency (SE) of the user l in the m^{th} user group as:

$$\text{SE}_l^m = R_l^m / B_l^m. \quad (20)$$

It is a measure of how efficiently a limited frequency spectrum is utilized to transmit the data by the proposed communication system. It is typically measured in bits/s/Hz. The ASE of the NOMA system can be viewed as $\text{ASE} = (KM)^{-1} \sum_{m=1}^M \sum_{l=1}^{K_m} \text{SE}_l^m$ where all the users reap the entire system bandwidth to transmit their data.

D. User Fairness

The user fairness of a communication system is analyzed to determine whether users or applications are receiving a fair share of system resources. For a given user-group, the user fairness can be quantified using the Jain's fairness index as

$$\mathcal{J}_m = \frac{(\sum_{i=1}^{K_m} R_i^m)^2}{K_m \cdot \sum_{j=1}^{K_m} R_j^m}. \quad (21)$$

The proposed NOMA scheme is particularly designed to elevate user fairness by assigning more transmission power

to the users with poor channel conditions and vice versa.

E. Outage Probability Analysis

Considering the scenario of unavailable or erroneous CSI, an outage event may happen in NOMA systems. The outage probability can be described as that the l^{th} user is unable to decode its own message or the message of the weaker user $j < l$ in its user cluster/group [36]. Thus, the outage probability (OP) at the l^{th} user in the m^{th} user group can be written as

$$\text{OP}_l^m = 1 - Pr(\{R_{1 \rightarrow l}^m \geq \tilde{R}_1^m\} \cap \{R_{2 \rightarrow l}^m \geq \tilde{R}_2^m\} \cap \dots \cap \{R_{l-1 \rightarrow l}^m \geq \tilde{R}_{l-1}^m\} \cap \{R_{l \rightarrow l}^m \geq \tilde{R}_l^m\}), \quad (22)$$

Using the notation $E_{j \rightarrow l}^m$ to denote the event of successful detection of user j message at the l^{th} user and (6), we get

$$E_{j \rightarrow l}^m = \{R_{j \rightarrow l}^m \geq \tilde{R}_j^m\}, \\ = \left\{ \frac{|h_l^m|^2 \alpha_j^m}{|h_l^m|^2 \sum_{k=j+1}^{K_m} \alpha_k^m + \varrho_m^{-1}} \geq \varphi_j^m \right\}, \quad (23)$$

where $\varphi_j^m = 2^{\tilde{R}_j^m/B} - 1$. The event $E_{j \rightarrow l}^m$ can be re-written as

$$E_{j \rightarrow l}^m = \left\{ |h_l^m|^2 \geq \frac{\varrho_m^{-1} \varphi_j^m}{\alpha_j^m - \varphi_j^m \sum_{k=j+1}^{K_m} \alpha_k^m} \right\}, \quad (24)$$

Conditioned on $\alpha_j^m \geq \varphi_j^m \sum_{k=j+1}^{K_m} \alpha_k^m$. Further define

$$\Psi_j^m \triangleq \frac{\varrho_m^{-1} \varphi_j^m}{\alpha_j^m - \varphi_j^m \sum_{k=j+1}^{K_m} \alpha_k^m}, \quad \forall j < K_m \quad (25)$$

and

$$\Psi_{K_m}^m \triangleq \frac{\varphi_{K_m}^m}{\alpha_{K_m}^m \varrho_m}. \quad (26)$$

and $\Psi_{\max}^{lm} = \max\{\Psi_1^m, \Psi_2^m, \dots, \Psi_l^m\}$. Then, the outage probability can be written as

$$\text{OP}_l^m = 1 - Pr(|h_l^m|^2 \geq \Psi_{\max}^{lm}) = Pr(|h_l^m|^2 \leq \Psi_{\max}^{lm}). \quad (27)$$

Hence, the outage probability comes out to be the cumulative distribution function (CDF) of rician squared distribution (RSD) $|g_l^m|^2$ as

$$\text{OP}_l^m = Pr\left(|g_l^m|^2 \leq \frac{\Psi_{\max}^{lm} L(d_l^m)}{G_l^m}\right). \quad (28)$$

Given the users location, user ordering and targeted data rates at an instant, the instantaneous OP_l^m can be evaluated, as shown in Appendix A, using the closed form expression as:

$$\text{OP}_l^m = 1 - \mathcal{Q}_1\left[\sqrt{2K_s}, \sqrt{\frac{2L(d_l^m)\Psi_{\max}^{lm}(1+K_s)}{G_l^m\Omega}}\right], \quad (29)$$

where Ω is the total power from both LoS and NLoS paths, and acts as a scaling factor to the Rician distribution i.e., $\Omega = \nu^2 + 2\sigma_f^2$. On the other hand, the outage probability of a user following orthogonal multiple access (OMA) just depends on the decoding of its own message. Such that

$$\text{OMA_OP}_l^m = Pr\{\text{OMA_}R_l^m \leq \tilde{R}_l^m\}, \quad (30)$$

where

$$Pr\{\text{OMA_}R_l^m \leq \tilde{R}_l^m\} = Pr\left\{\frac{|h_l^m|^2 \rho_m^{\text{OMA}}}{K_m} \leq \text{OMA_}\varphi_l^m\right\}, \quad (31)$$

where $\text{OMA_}\varphi_l^m = 2^{\tilde{R}_l^m K_m/B} - 1$. Hence, the outage probability in an OMA scenario is equivalent to cumulative density function.

$$\text{OMA_OP}_l^m = Pr\left(|g_l^m|^2 \leq \frac{\Psi_{\text{OMA}}^{lm} L(d_l^m)}{G_l^m}\right), \quad (32)$$

where $\Psi_{\text{OMA}}^{lm} = \text{OMA_}\varphi_l^m K_m / \rho_m^{\text{OMA}}$ with $\rho_m^{\text{OMA}} = P_t / \sigma_{\text{OMA}}^2$. Note that the orthogonal frequency-division multiple access (OFDMA) based system will allow spectrum segregation among the users and hence the identical noise power/variance σ_{OMA}^2 for users in m^{th} user-group will be given by

$$\sigma_{\text{OMA}}^2 (\text{dBm}) = -174 + 10 \log_{10}(B/K_m) + \text{NF}. \quad (33)$$

Eventually outage probability for OMA system can be written as the following Marcum Q-function

$$\text{OMA_OP}_l^m = 1 - \mathcal{Q}_1\left[\sqrt{2K_s}, \sqrt{\frac{2L(d_l^m)\Psi_{\text{OMA}}^{lm}(1+K_s)}{G_l^m\Omega}}\right]. \quad (34)$$

We can now formulate the optimization problem to design optimal power allocation in order to maximize the sum rate of all users within the allocated power budget.

V. PROBLEM FORMULATION AND PROPOSED SOLUTION

This work aims to jointly optimize numerous design parameters with the objective to maximize sum rate of all users in the coverage area of HAPS while guarantying their QoS, user fairness, and expenses within the available power budget. The optimization problem is targeted at optimizing the following:

- 1) User grouping: M number of user groups to accommodate all user in the coverage area and the central locations for M beam spots i.e., \mathbf{w}_m
- 2) User association: x_l^m decides the association between the users and the defined groups
- 3) Beam optimization: Beam width θ_m or beam radius r_m
- 4) NOMA power allocation: Power allocation coefficients for each user in every user group $\alpha_l^m \forall l, m$

We formulate the design problem for the parameters optimization in a HAPS communication system with the necessary constraints as follows:

$$\mathbf{P1}: \underset{\substack{M, \mathbf{X}, \mathbf{W}, \theta, \\ r, \alpha}}{\text{maximize}} \sum_{m=1}^M \sum_{l=1}^{K_m} x_l^m R_l^m (\alpha_l^m, \theta_m^{3\text{dB}}) \quad (35a)$$

$$\text{s.t.} \quad 1 \leq M \leq K, \quad (35b)$$

$$x_l^m \in \{0, 1\} \& \sum_{m=1}^M x_k^m = 1, \quad \forall k \quad (35c)$$

$$\sum_{m=1}^M x_k^m \|\mathbf{u}_k - \mathbf{w}_m\| \leq \sum_{m=1}^M x_k^m r_m, \quad \forall k \quad (35d)$$

$$\theta_m^{3\text{dB}} \geq \frac{70\pi}{D} \& r_m \geq \frac{0.443\lambda H}{D}, \quad \forall m \quad (35e)$$

$$R_{j \rightarrow l}^m \geq \tilde{R}_j^m, \text{ for } j \leq l \quad \forall m \quad (35f)$$

$$0 \leq \alpha_l^m \leq 1, \& \sum_{l=1}^{K_m} \alpha_l^m \leq 1, \quad \forall l, m \quad (35g)$$

$$\alpha_1^m \geq \alpha_2^m \dots \geq \alpha_{K_m}^m, \quad \forall K_m, m \quad (35h)$$

where $\mathbf{X} \in \mathbb{B}^{K \times M}$ is a Boolean matrix with entries $x_l^m \in \{0, 1\}, \forall l, m$ where $1 \leq l \leq K$ and $1 \leq m \leq M$ and $\mathbf{W} \in \mathbb{R}^{M \times 2}$ contains the 2D coordinates of M beam centers. Moreover, $\theta \in \mathbb{R}^M$ and $\mathbf{r} \in \mathbb{R}^M$ are the vectors comprising of the 3dB-beamwidth and spot beam radii of M beams. Additionally, $\alpha \in \mathbb{R}^{K \times M}$ contains the values of the power coefficients for all users in M groups. It is a sparse matrix with non-zero entries only where $x_l^m = 1$. Importantly, the sum of

Algorithm 1 Geometric Disk Cover Problem

- 1: **Input:** The coordinates of users $\{\mathbf{u}_k\}$ in the horizontal plane of the HAPS coverage area and beam radius $\{r_m\} \forall m$.
 - 2: **Output:** The number $\{M\}$ and the coordinates of the beam centers $\{\mathbf{w}_m\}$.
 - 3: **Initialize** $m \leftarrow 1$
 - 4: **Compute** Distance Matrix $\mathbf{D} \in \mathbb{R}^{K \times K}$ containing distance of a user with every other user and evidently zero diagonal entries.
 - 5: **Define** Boolean matrix $\bar{\mathbf{D}} \in \mathbb{B}^{K \times K}$ with entries $\bar{d}_{ij} = 1$ iff $d_{ij} \leq r_m$ otherwise zero.
 - 6: **while** $(m \leq K \parallel \bar{\mathbf{D}} \neq \mathbf{0})$ **do**
 - 7: **for** $k \leftarrow 1$ to K **do**
 - 8: For each row k , find the non-zero entries $\bar{d}_{kj} \neq 0$ and then from j columns, pick one column l with maximum column sum i.e., $l = \operatorname{argmax}_j \sum_i \bar{d}_{ij}$
 - 9: Mark $\mathbf{w}_m \leftarrow \mathbf{u}_l$
 - 10: Update matrix $\bar{\mathbf{D}}$ by nullifying all j columns and j rows which had non-zero entries $\bar{d}_{kj} \neq 0$.
 - 11: $m \leftarrow m + 1, k \leftarrow k + j$
 - 12: **end for**
 - 13: **end while**
 - 14: **return** $M \leftarrow m$ and \mathbf{w}_m indicates the centers of m user groups.
-

all entries in a row of \mathbf{X} matrix should be equal to 1 since any user can only associate to one user group m whereas the sum of all entries in a column of α should be less than or equal to 1 because the sum of power coefficients of all users within same usergroup cannot exceed the allocated power budget.

The user grouping constraint is given in (35b) whereas the user association constraints are presented by (35c). Any user can only associate to one user group at a given time. Moreover, the constraint (35d) ensures that the associating user resides within the beam coverage area. The beam constraints (35e) are essential lower bounds on the 3dB-beamwidth and spot beam radii, which are adjustable by the beamwidth control. Note that a narrower beam than the given bounds is not achievable with the given antenna array dimensions. The target rate constraint guarantees QoS of all users and warrants the accurate decoding of all users with weaker channel gains which is essential for perfect SIC. The last two constraints on the power allocation coefficients eqs. (35g) and (35h) ensure the transmission power expenses with in power budget and optimal user ordering for user fairness. In any given user group, the maximum power is allocated to the user with weaker channel gains and vice versa.

The objective of this optimization problem is to maximize the sum rate of all users within the given resources while guarantying QoS and user fairness. However, the problem **P1** is a non-convex mixed integer programming problem in the given optimization variables. Therefore, we divide this problem into sub-problems and solve these sub-problems sequentially as presented in Algorithm 2. The subproblems are solved for few optimization parameters assuming that rest all

design parameters are fixed or given. When the beam coverage radius is fixed, problem **P1** can be converted into the following user-grouping problem:

$$\mathbf{P1(a)}: \underset{M, \mathbf{W}}{\text{minimize}} M \quad (36a)$$

$$\text{s.t.} \quad 1 \leq M \leq K, \quad (36b)$$

$$x_l^m \in \{0, 1\} \& \sum_{m=1}^M x_k^m = 1, \quad \forall k \quad (36c)$$

$$\sum_{m=1}^M x_k^m \|\mathbf{u}_k - \mathbf{w}_m\| \leq \sum_{m=1}^M x_k^m r_m, \quad \forall k \quad (36d)$$

This sub-problem finds the optimal locations (beam centers \mathbf{w}_m) of the minimal number of beams required to cover the disk of radius R i.e., the coverage area of HAPS communication system. Problem **P1(a)** is a well-known GDC problem which aims to find minimum number of disks of given radius to cover a set of points in a plane. The famous GDC problem is non-polynomial (NP)-hard highlighting the NP-hardness of **P1**. This problem can be solved using Algorithm 1, where the distance and boolean matrices highlight the nearest neighbors with non-zero entries and the user with maximum number of neighbors is marked as the center of user-group \mathbf{w}_m . Next, we eliminate all the users in the coverage neighborhood of \mathbf{w}_m to find the next beam center. The convergence of the algorithm is guaranteed as it works by eliminating the rows and corresponding columns. The iterations stop when all rows or columns are nullified i.e., all users must reside within the coverage radii of the selected beam centers.

Next subproblem **P1(b)** solves the user association problem and finds out the association parameters $x_l^m, \forall l, m$.

$$\mathbf{P1(b)}: \underset{\mathbf{X}}{\text{maximize}} \sum_{m=1}^M \sum_{l=1}^{K_m} x_l^m R_l^m (\alpha_l^m, \theta_m^{3\text{dB}}) \quad (37a)$$

$$\text{s.t.} \quad x_l^m \in \{0, 1\}, \quad \forall l, m \quad (37b)$$

$$\sum_{m=1}^M x_k^m = 1, \quad \forall k \quad (37c)$$

$$\sum_{m=1}^M x_k^m \|\mathbf{u}_k - \mathbf{w}_m\| \leq \sum_{m=1}^M x_k^m r_m, \quad \forall k \quad (37d)$$

Clearly, the users would like to associate with the beams of closest beam centers to receive maximum SINR which will result in the maximum user rate. As a result, by greedy algorithm, the indicator variables can be obtained as:

$$x_l^m = \begin{cases} 1, & m = \operatorname{argmin}_m \|\mathbf{u}_l - \mathbf{w}_m\|, \\ 0, & \text{otherwise,} \end{cases} \quad (38)$$

This (38) is evaluated for each user and each user can associate with only one closest beam at a given time. Based on the user grouping and user association from **P1(a)** and **P1(b)**, we can carry out beam optimization in the subsequent problem **P1(c)**.

$$\mathbf{P1(c)}: \underset{\mathbf{W}, \mathbf{r}, \theta}{\text{maximize}} \sum_{m=1}^M \sum_{l=1}^{K_m} x_l^m R_l^m (\alpha_l^m, \theta_m^{3\text{dB}}) \quad (39a)$$

$$\text{s.t.} \quad r_m \geq \max\{x_k^m \|\mathbf{u}_k - \mathbf{w}_m\|\}, \forall k, m \quad (39b)$$

$$\theta_m^{3\text{dB}} \geq 70\pi/D, \quad \forall m \quad (39c)$$

$$r_m \geq 0.443\lambda H/D, \quad \forall m \quad (39d)$$

The problem can be solved independently for all user groups.

Interestingly, the antenna beam gain (8) is convex and monotonically decreasing in r_m for a given user group. In addition, G_l^m is directly proportional to γ_l^m and eventually R_l^m . Moreover, for a fixed HAPS altitude H , the beam radius and HPBW are interchangeable as $r_m = H \tan(\theta_m^{3\text{dB}}/2)$. With this background, we can conclude that maximizing the sum rate or SINR is equivalent to maximizing the antenna beam gain. Notably, the maximization of a function that is convex and continuous, and defined on a set that is convex and compact, attains its maximum at some extreme point of that set [37]. Hence, the aforementioned problem can be solved by finding the minimum value of beam radius which satisfies the constraints eqs. (39b) and (39d). This can be achieved in the following two ways:

- 1) The problem is equivalent to solving the MEC problem for a given set of points in the user group thus we can employ the well-known Welzl's algorithm [38] to identify the fine-tuned beam locations W with minimum beam radii r , which can cover the given set of users in a user group.
- 2) Another near-optimal solution is to evaluate $\mathbf{w}_m = K_m^{-1} \sum_{k=1}^K x_k^m \mathbf{u}_k^m$ and $r_m = \max\{\|x_k^m \mathbf{u}_k^m - \mathbf{w}_m\|\}, \forall k$. This simplified closed-form heuristic approach performs close to the optimal solution.

Given r_m , we can evaluate the corresponding HPBW using (1) and the process can be repeated independently for each user group. Once the beam optimization problems are solved, we get the optimal user grouping and user association. This enables us to design the power allocation parameters disjointly for each user group based on their distances from the group/beam center and user ordering as shown in **P1(d)**:

$$\mathbf{P1(d)} \text{ maximize } \sum_{\alpha} \sum_{m=1}^M \sum_{l=1}^{K_m} x_l^m R_l^m (\alpha_l^m, \theta_m^{3\text{dB}}) \quad (40a)$$

$$\text{s.t.} \quad R_{j \rightarrow l} \geq \tilde{R}_j^m, \text{ for } j \leq l \quad \forall m \quad (40b)$$

$$\sum_{l=1}^{K_m} \alpha_l^m \leq 1, \quad \forall m \quad (40c)$$

$$0 \leq \alpha_l^m \leq 1, \quad \forall l, m \quad (40d)$$

$$\alpha_1^m \geq \alpha_2^m \dots \geq \alpha_{K_m}^m, \quad \forall K_m, m \quad (40e)$$

Assuming the same target rate threshold for all users within a user group m i.e., $\tilde{R}_j^m = \Omega_m, \forall j$, the problem **P1(d)** can be solved in a closed-form as presented in [39]:

Lemma 1. *Given the user ordering with decreasing A_l^m i.e., $A_1^m \geq A_2^m \geq \dots \geq A_{K_m}^m$, the sum rate and minimum power coefficients of users in m^{th} group, respectively, are given by*

$$R_m^* = K_m \Omega_m + B \log_2 \left[1 + \frac{1 - \sum_{k=1}^{K_m} \hat{\alpha}_k^m}{A_{K_m}^m} \right], \quad (41)$$

$$\hat{\alpha}_l^m = \left(2^{\Omega'_m} - 1 \right) \left(\sum_{k=l+1}^{K_m} \hat{\alpha}_k^m + A_l^m \right), \quad (42)$$

where

$$A_l^m = \frac{L(d_l^m)}{\varrho_m |g_l^m|^2 G_l^m} \quad \text{and} \quad \Omega'_m = \frac{\Omega_m}{B}. \quad (43)$$

if the following condition holds

$$\left(2^{\Omega'_m} - 1 \right) \left(\sum_{i=1}^{K_m} 2^{(i-1)\Omega'_m} A_i^m \right) \leq 1. \quad (44)$$

Algorithm 2 HAPS Communication Parameters Optimization

- 1: **Input:** $\{R_H\}, \{H\}, \{K\}$, and $\{\mathbf{u}_k\}$.
 - 2: **Output:** $\{M\}, \{\mathbf{w}_m\}, \{r_m\}, \{\theta_m\}, \{x_l^m\}$, and $\{\alpha_l^m\} \forall l, m$.
 - 3: **Initialize** $i \leftarrow 0, R[i-1] \leftarrow R_0$ and $\epsilon \leftarrow \infty$
 - 4: **Select** QoS minimum rate threshold Ω_m and minimum possible beam radius r_{\min}
 - 5: **Set** tolerance $\delta, r[i] = r_{\min}$, and $r_{\text{UB}} = R_H$
 - 6: **Choose** Δr and identical beam radius $r_m[i] = r[i] \forall m$
 - 7: **while** $\epsilon \geq \delta$ & $r_{\min} \leq r_m[i] \leq R_H$ **do**
 - 8: **Let** $i \leftarrow i + 1$
 - 9: **Update** $r_m[i] = r_m[i-1] + \Delta r$ for all user-groups ensuring sequential increment with every iteration.
 - 10: **Determine** $M[i]$ and $\mathbf{w}_m[i] \forall m \in [1, M]$ using Algorithm 1 to solve **P1(a)** given constant $r[i]$.
 - 11: **Associate** users by solving **P1(b)** to evaluate $X[i]$ containing $x_l^m[i]$.
 - 12: **Optimize** individual beams for each user group to evaluate $\tilde{w}_m[i], \tilde{\theta}_m[i]$ and $\tilde{r}_m[i]$ by solving **P1(c)**.
 - 13: **Update** $w_m[i] \leftarrow \tilde{w}_m[i], r_m[i] \leftarrow \tilde{r}_m[i]$ and $\theta_m[i] \leftarrow \tilde{\theta}_m[i]$.
 - 14: **Evaluate** the distance $d_l^m[i]$, elevation angle $\psi_l^m[i]$ and transmit antenna gain $G_l^m[i]$ for each user, in the m^{th} group, from the HAPS station.
 - 15: **Calculate** the Rician channel coefficient using the small scale fading CSI $g_l^m[i]$, pathloss $L(d_l^m[i])$, and beam gain $G_l^m[i]$ for each user l in all M user-groups.
 - 16: **Obtain** the available transmit power $P_t[i]$ of a solar powered HAPS at the chosen location on a given date and time using the power estimation algorithms in [28].
 - 17: **Compute** the power allocation coefficients $\alpha_l^m[i]$ for each user in the m^{th} user group using the closed form solutions of **P1(d)** for the given Ω_m and user ordering.
 - 18: **Evaluate** the users rate $R_l^m[i]$ using $\alpha_l^m[i]$, channel gains $h_l^m[i]$ to find $\tilde{R}[i]$.
 - 19: **Compare** $\tilde{R}[i]$ with $R[i-1]$
 - 20: **if** $\tilde{R}[i] \geq R[i-1]: r_m[i] \leftarrow \tilde{r}_m[i]$ and $R[i] \leftarrow \tilde{R}[i]$
 - 21: **if** $\tilde{R}[i] \leq R[i-1]: r_m[i] \leftarrow r_m[i-1]$ and $R[i] \leftarrow R[i-1]$
 - 22: Update $\epsilon \leftarrow \tilde{R}[i] - R[i-1]$
 - 23: **end while**
 - 24: User Grouping Parameters: $M^* \leftarrow M[i], \mathbf{w}_m^* \leftarrow \mathbf{w}_m[i]$
 - 25: User Association Parameters: $x_l^{m^*} \leftarrow x_l^m[i] \forall l, m$
 - 26: Beam radii: $r_m^* \leftarrow r_m[i]$
 - 27: Half-power beam widths: $\theta_m^* \leftarrow \theta_m[i] \forall m$
 - 28: Power Allocation Parameters: $\alpha_l^{m^*} \leftarrow \alpha_l^m[i] \forall l, m$
 - 29: Sum Rate of Users: $R^* \leftarrow R[i]$
-

The first term of R_m^* is the QoS thresholds of all users in m^{th} user group whereas the second term is the additional rate of K_m user after allocating the remaining power $1 - \sum_{k=1}^{K_m} \hat{\alpha}_k^m$ to it, in order to maximize the sum rate. It is important to highlight that the users are ordered

Lemma 2. *For $\sum_{k=1}^{K_m} \hat{\alpha}_k^m \geq 1$, there exists a user u in $1 \leq u \leq$*

TABLE II: The Adopted System Parameters

ξ	$2^0 14' 2.04''\text{E}$	R	60km
χ	$53^0 28' 0.48''\text{N}$	K (density)	1000 – 15000/km ²
SS: j_d	2460848	D	1.5m
WS: j_d	2461031	S	143m ²
SS: α_{ext}	0.465	W	165kg
WS: α_{ext}	0.29	H	21km
b, AR_w	35m, 30	r_{min}	5.4640km
B	200MHz	ψ_{min}	12 π /180
T_n	870	δ, NF	1e – 4, 5dB
k_B	1.3800e – 23	N_0	–174dBm
Υ	10%	σ_f^2	1
f_c	27.5GHz	Υ	0.1

K_m which satisfies the following condition:

$$\begin{cases} \left(2^{\Omega'_m} - 1\right) \left(\sum_{i=u+1}^{K_m} 2^{i-1} A_i^m\right) \leq 1, \\ \left(2^{\Omega'_m} - 1\right) \left(\sum_{i=u}^{K_m} 2^{i-1} A_i^m\right) \geq 1. \end{cases} \quad (45)$$

Hence, the maximum achievable sum rate is given by

$$R_m^* = (K_m - u) \Omega_m + B \log_2 \left[1 + \frac{\Delta\alpha}{1 - \Delta\alpha + A_u^m} \right], \quad (46)$$

where

$$\Delta\alpha = 1 - \left(2^{\Omega'_m} - 1\right) \left(\sum_{i=u+1}^{K_m} 2^{i-1} A_i^m\right). \quad (47)$$

The first term of R_m^* is the QoS thresholds of users from $u + 1$ to K_m and the second term is the rate of user u with power $\Delta\alpha$. It signifies that only $u + 1$ to K_m can attain QoS threshold with powers $\hat{\alpha}_{u+1}^m, \hat{\alpha}_{u+2}^m, \dots, \hat{\alpha}_{K_m}^m$, respectively, using (42). However, the users before (and including) u^{th} user cannot achieve their target rates. So, the remaining power $\Delta\alpha$ is allocated to the u^{th} user.

The HAPS communication parameters are jointly optimized using the iterative Algorithm 2. This algorithm incorporates the HAPS stratospheric location at a given time and date to evaluate the available transmit power [28] and the QoS user rate minimum threshold Ω_m . We use the coordinates of users in the horizontal plane to determine their distances d_l and elevation angles ψ_l from HAPS. The sum rate is initialized as $R[i-1] = R_0$ with uniform power allocation in the absence of user grouping and beam optimization i.e., $M = 1, r_m = R$, and \mathbf{w}_m is the center of the HAPS circular coverage area. The algorithm solves sub-problems **P1(a)**–**P1(d)** for a given value of beam radius. The lower and upper bounds on the beam radius are updated iteratively, based on the branch and bound algorithm to search the optimal beam radius which maximizes the user's sumrate. This iterative algorithm repeats till it meets the stopping criteria i.e., until the increase in sumrate is insignificant $\epsilon \leq \delta$. Eventually, the algorithm furnishes the optimum values of the $M^*, \mathbf{w}_m^*, r_m^*, \theta_m^*, x_l^{m*}, \alpha_l^{m*}$, and R^* .

A. Complexity and Convergence

The complexity of Algorithm 2 depends on the complexities of solving sub-problems **P1(a)**–**P1(d)** and the number of branch-and-bound iterations. The algorithm solves **P1(a)** to find the number $M[i]$ and centers of user groups $\mathbf{w}_m[i]$ each of radius $r[i]$ to cover all the users in the coverage area. In computational geometry, this classic GDC problem

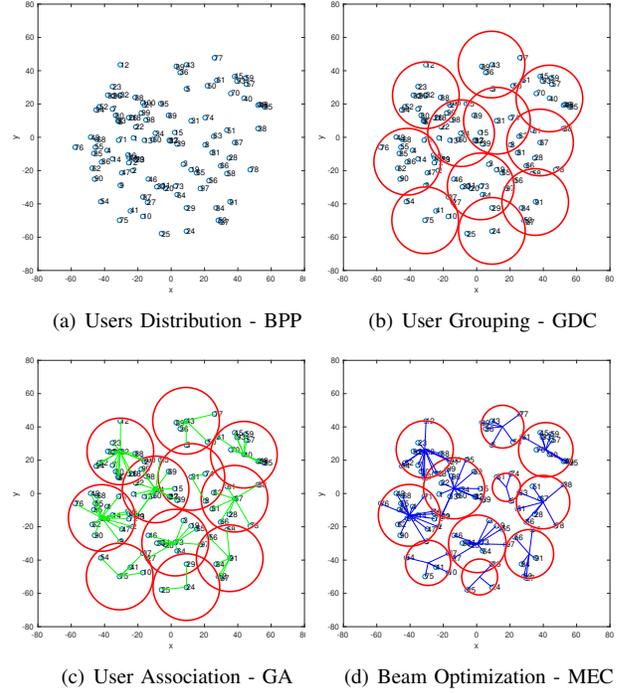


Fig. 2: User grouping, user association, and beam optimization

is a NP-hard problem with complexity $\mathcal{O}(2^K)$. However, the proposed Algorithm 1 solves it by finding \mathbf{D} and then sorting it to find minimum possible M circles with the given radius to cover K users, rendering complexity $\mathcal{O}(K \log K)$. Next, the problem **P1(b)** associates K users with M serving spot beams. The greedy algorithm evaluates M SNR values for each user resulting in complexity $\mathcal{O}(KM)$. Later, each spot beam is individually optimized using either Welzl's algorithm or proposed heuristic solution. This yields the optimal beam parameters $\tilde{w}_m[i], \tilde{r}_m[i]$, and $\tilde{\theta}_m[i]$ for each spot beam. The worst-case complexity of the Welzl's algorithm is quadratic i.e., $\sum_{m=1}^M \mathcal{O}(K_m)^2$, where K_m is the number of users in the m^{th} user group. On the other hand, the heuristic method adds linear complexity $\sum_{m=1}^M \mathcal{O}(2K_m)$. Eventually the proposed NOMA power allocation sub-problem can be solved with complexity $\mathcal{O}(KM \log K)$. Assuming I_{BB} iterations for the branch-and-bound convergence, the overall complexity of the proposed Algorithm 2 can be seen as:

$$\mathcal{C} = \mathcal{O}(I_{\text{BB}}(KM + K(M+1) \log K + \sum_{m=1}^M \mathcal{O}(2K_m))), \quad (48)$$

Considering $K_m \leq K$ and the presence of a dominant term, the computational complexity can be simplified to $\mathcal{C} = \mathcal{O}(I_{\text{BB}}K(M+1) \log K)$.

The convergence of the HAPS parameter optimization algorithm involves multiple interdependent steps. The Branch and Bound method ensures optimal beam radius selection through systematic exploration and pruning of all feasible branches. User grouping using geometric disk cover converges to an efficient clustering solution. The greedy algorithm for user association quickly achieves a local optimum, though global optimality isn't guaranteed. Beam optimization efficiently determines the minimum enclosing circle, converging

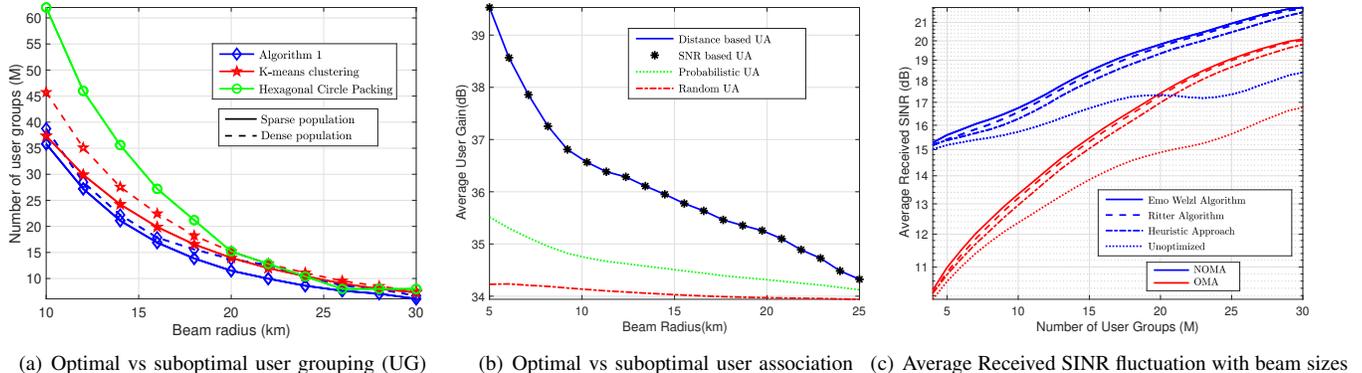


Fig. 3: Impact of user grouping, association, and beam optimization

in expected linear time. Lastly, resource allocation via NOMA closed-form power allocation converges effectively. Overall, the algorithm's success hinges on proper initialization, sequential dependency among steps, and well-defined termination conditions, ensuring a feasible and effective communication strategy across all users.

VI. NUMERICAL RESULTS

The numerical results are evaluated for a HAPS aircraft flying at an altitude 21km and serving the area with coverage radius of 60km over coordinates $2^{\circ} 14' 2.04''\text{E}$ and $53^{\circ} 28' 0.48''\text{N}$. We have adopted the PHASA-35 aircraft model with wingspan b of 35m, total weight W (platform and payload mass) 165kg, wing area S 143m^2 , and maximum achievable altitude H_{\max} 21.336km. We assume K users with BPP distribution in the coverage area. We consider $f_c = 27.5\text{GHz}$ carrier frequency with 200MHz channel bandwidth. Moreover, the phased array antenna is assumed to be 90% efficient with diameter 1.5m. The available transmit power is computed for different hours of the day on the winter solstice (WS) and summer solstice (SS) of 2025 using the solar algorithms [28]. The carrier frequency can be chosen from the microwave band for long-range communication with minimal line losses or from the millimeter-wave band for higher channel bandwidth, depending on the specific requirements. Nonetheless, the choice must be backed by the international telecommunication union (ITU) allocated frequency bands for aerial communications². The adopted values of numerous simulation parameters are presented in Table II unless specified otherwise.

The proposed user grouping, user association, and beam optimization are illustrated in Fig. 2. The distribution of 100 users is depicted in the circular HAPS coverage area of 60km centered at the origin (0,0) on a horizontal plane in Fig. 2(a). Assuming 20km initial beam radius, the Algorithm 1 produces the user grouping depicted in Fig. 2(b). The algorithm renders the minimum number of $M = 11$ beams

²The carrier frequencies of the 2.1GHz is preferred for seamless merger with the existing terrestrial network [11]. It is approved for HAPS base stations offering mobile services according to RR5.388A and ITU Resolution 221(Rec. WRC-07). However, the 27.5GHz band of FR2 millimeter waves in 5G NR offers a much larger bandwidth with shorter range. Interestingly, the presented system models, performance analysis and optimization framework are valid for any frequency range after incorporating the corresponding path losses in the propagation model.

along with their optimal locations i.e., beam centers in order to accommodate all users. Next, the user association is carried out based on the greedy algorithm in Fig. 2(c). The green lines are drawn between the users and the beam centers in a user group to demonstrate their association. Moreover, beam optimization is exhibited in Fig. 2(d) where each spot beam is individually optimized to minimize the beam radius and readjust their centers while serving the same users within a user group. Evidently, this reduces the overlapping regions and concentrates the power density, which maximizes the SINR and consequently the sum rate of the users.

We evaluate the effect of optimal vs sub optimal user grouping for a range of beam radii considering two population densities in Fig. 3(a). The performance of the Algorithm 1 is compared with the popular k-means clustering and hexagonal circle packing. Intuitively, as the beam radius increases, the number of designed user groups decreases, while an increase in population density leads to a rise in the number of user groups. Algorithm 1 and k-means clustering can adapt to varying population densities, while hexagonal circle packing operates independently of these variations. Evidently, the proposed Algorithm 1 renders the minimum number of user groups for any given beam radii and population density. This advantage is particularly significant for sparse populations and small beam radii. For instance, Algorithm 1 requires $M = 28$, k-means clustering requires $M = 35$, and circle packing requires $M = 46$ user groups to serve the same users in a sparsely populated area with spot beams of radius 12km.

Next, we study the impact of distance-based UA versus the SNR-based UA, probabilistic UA (PUA), and random UA (RUA) on the average user directivity gain in Fig. 3(b). The proposed greedy algorithm performs equally good as the generalized SNR-based UA given negligible shadowing and small-scale fading effects at high-altitudes. The direct LoS path to ground users in rural/remote areas with minimal interference from physical obstructions results in trivial shadowing effects. Likewise, the lack of multipath reflections significantly reduces the impact of small-scale fading. Thus, the path loss dominance endorses distance-based UA as a reliable approach without heavily relying on fluctuating SNR metrics, simplifying network management and UA in HAPS

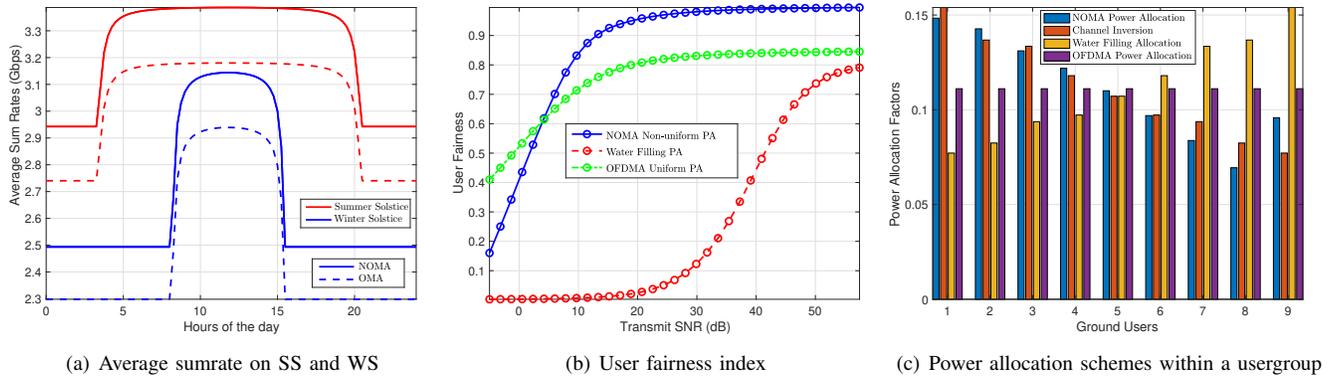


Fig. 4: Impact of Power Allocation Strategy

networks. The proposed scheme is also compared with the PUA and RUA. PUA assigns users to the beams based on the probability of their estimated distances whereas random association blindly associates users to the serving beams. Evidently, the proposed approach outperforms the rest and yields upto 4dB and 5dB average user gain over PUA and RUA, respectively. Clearly, the user antenna/directivity gain reduces with the increasing beam radii and resultant reduced power density.

Subsequently, we evaluate the impact of the proposed beam optimization based on MEC approach with state-of-the-art Ritter's Algorithm, a least-complex heuristic approach, and unoptimized beams on the average received SINR. Fig. 3(c) illustrates the SINR performance with the increasing number of user groups adopting NOMA and OMA power allocation schemes. Interestingly, increasing M to serve a specific set of users within a coverage area raises the power density per beam, resulting in higher SINR for each user group. However, this improvement is accompanied by time delays associated with the addition of spot beams and the use of TDM limits M , to improve overall spectral efficiency. Evidently, the employment of Emo Welzl's algorithm finding the minimum enclosing circles to solve the equivalent problem **P1(c)** outperforms the well-known Ritter's algorithm. Moreover, the proposed heuristic approach closely follows the two popular algorithms rendering immense SINR gains as compared to the unoptimized beams. The data analysis reveals the SINR gains upto 3.44dB, 3.36dB, and 3.16dB by employing Optimal Welzl's algorithm, Ritter's algorithm and heuristic approach, respectively, over unoptimized spot beams in case of NOMA. Likewise, we observe the respective gains of 3.33dB, 3.27dB, and 3.05dB in case of OMA. The gains are particularly significant for higher M . In addition, the average SINR improvement of 5dB and 3dB can be attained by employing NOMA over OMA for $M = 4$ and $M = 15$, respectively. Thus, NOMA scheme is particularly preferred for small M encompassing large number of users within each user group.

Fig. 4(a) presents the average sum rate of users within the HAPS coverage area for different times of the day on both WS and SS in 2025. At the chosen location, we experience approximately 16.75 and 7 hours of daylight on the best (SS)

and worst (WS) cases, respectively. This solar-powered HAPS aircraft harvests solar energy and converts it into electrical energy, which powers its propulsion, transmission, and accessory systems while storing sufficient energy for nighttime operations. We propose using all surplus power generated during the day for wireless transmission, while maintaining a constant transmission power during night due to limited available energy. Transmission power values are based on the solar model from [28], adjusted for inevitable feed line losses. The day is divided into 15-minute intervals, during which the solar elevation angle remains nearly constant. The ground users can achieve a higher sum rate during the day, particularly around noon, pertaining to the higher available transmission power. Additionally, the higher sum rates during the SS compared to the WS are due to the greater availability of solar power, while the extended period of elevated sum rates is attributed to the longer daylight hours during SS. The average sum rate with the NOMA power allocation clearly outperforms the OMA counterpart for both WS and SS. The percentage increase in the average sum rates at WS and SS is upto 6.97% and 6.5% during day and 8.7% and 7.08% during night, respectively, with the optimal power allocation.

The fairness analysis of the presented power allocation strategies with increasing transmit SNR is illustrated in Fig. 4(b). NOMA improves user fairness because it allocates more power to users with weaker channels, ensuring they achieve acceptable data rates. As SNR increases, the fairness gap between strong and weak users tends to narrow. On the other hand, uniform power allocation allocates equal resources irrespective of their channel conditions. Hence, as the SNR increases, the strong users' performance improves significantly, whereas weaker users see less proportional improvement. WFA allocates power across multiple users based on their channel conditions. In the given scenario, it demonstrates the worst fairness performance as it allocates more power to the stronger users (with better channels) and less power to weaker users (with weaker channels) leading to the significant gap between strong and weak users. We notice 22% and 147.5% improvement in average user fairness with the proposed NOMA power allocation as opposed to OFDMA and WFA, respectively, at 40dB transmit SNR.

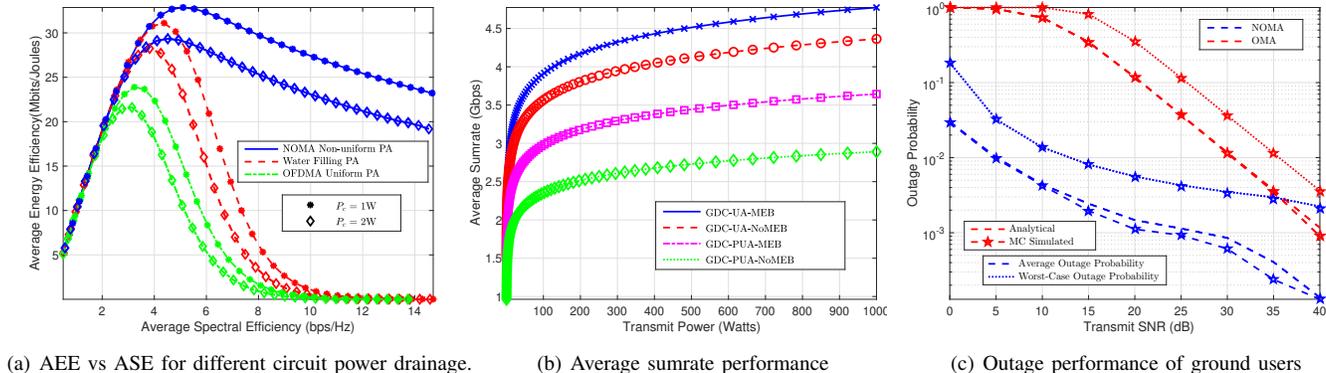


Fig. 5: System Efficiency, Average Sumrate, and Outage Performance

The bar chart in Fig. 4(c) demonstrates the power distribution amongst ordered users (weaker to stronger) within one usergroup. NOMA allocates higher fraction of available power budget to weaker users and vice versa for a fair distribution, enabling every user in the group to meet QoS threshold. Once all users are able to meet the target rate, the excessive power is allocated to the strongest user to maximize the sum data rate. On contrary, channel inversion technique allocates more power to weaker users and vice versa without considering any QoS constraint. This may lead to unfavorable power allocation and reduced data rates for some strong users. OFDMA power allocation assigns equal resources to all users irrespective of their channel strengths and may deprive weak users as the allotted power may be insignificant for them to achieve their target data rates. WFA allocates more power to strong users and vice versa. This favors strong users while the weak users experience outage, leading to higher sumrates with no QoS assurance and fairness guarantee.

We further investigate the AEE versus ASE with different circuit power requirements as demonstrated in Fig. 5(a). We compare the system efficiency of proposed NOMA power allocation with the famous water-filling approach (WFA) and OFDMA uniform power allocation. Interestingly, the AEE increases with the increasing ASE until it reaches its peak value and then decreases with further increase in ASE until it saturates. The NOMA scheme depicts highest AEE values in the entire range of transmit ASE for different circuit power drainage. The NOMA, WFA, and OFDMA show peak AEE of 32.85Mb/J, 31.1Mb/J, and 23.89Mb/J at ASE of 5.1bps/Hz, 4.37bps/Hz, and 3.23bps/Hz, respectively, for $P_c = 1W$. Clearly, NOMA surpasses the other schemes in both AEE and ASE, and it even manages higher AEE saturation levels.

Fig. 5(b) demonstrates the system-level performance to evaluate key performance metric i.e., average sum rate to assess the mutual benefits of the suggested user-grouping algorithm, beam optimization technique, and resource allocation strategy. We assume a dense population of 13,000users/km² distributed in a circular area of 60km radius and grouped into 64 user groups using GDC Algorithm 1. Then, we study the impact of the recommended user association based on greedy algorithm or probabilistic user association. Next, we analyze

the performance with and without advocated beam optimization i.e., MEB or NoMEB, respectively. The overall system performance is assessed as the average of sumrates of each user-group. Clearly, the recommended UA yields higher average sumrates as compared to PUA. Likewise, the advocated MEB outperforms NoMEB scenarios. It is noteworthy that the performance degradation with GDC-PUA-MEB is higher than that with GDC-UA-NoMEB. Eventually, we achieve the best system performance when all the recommended strategies i.e., GDC-UA-MEB. We observe the percentage improvement of upto 9.4%, 30.9%, and 65.05% with GDC-UA-MEB over GDC-UA-NoMEB, GDC-PUA-MEB, and GDC-PUA-NoMEB, respectively.

Eventually, we investigate the outage performance of the proposed HAPS communication system under NOMA and OFDMA for a range of transmit SNR in Fig. 5(c). In this example, the individual users with achievable data rate below the threshold rate 100Mbps are categorized as users in outage. Expectedly, the number of users in outage decreases with the increase in the available transmit power. Therefore, one can deduce less outage as well as high data rates during the day light hours and summer season. The analysis reveals the lower outage probability of NOMA as opposed to OFDMA for different scenarios. We have presented the average outage probability of all users and worst case outage probability of the edge user for both NOMA and OFDMA. Impressively, the average outage probability for NOMA falls upto $1e - 4$ for 40dB SNR despite the long-distance communication and excessive path losses whereas OFDMA can barely make it to $1e - 3$ at the same SNR level rendering a ten-folds gain with the proposed scheme. On the other hand, the worst-case outage probability presents an error floor meaning that the outage probability cannot be improved with further increase in transmit power. The results in Fig. 5(c) show strong agreement between the presented closed-form analytical expression of outage probability (involving Marcum-Q function) and the Monte-Carlo simulations.

VII. CONCLUSION

A self-sustaining, solar-powered stratospheric HAPS is proposed to provide aerial communication services, connecting unconnected ground users in a wide coverage area. This study

explores the impact of user grouping, user association, and beam optimization on system performance, and proposes a downlink NOMA strategy to superpose signals for each user group served by steerable beams from phased array antennas. Optimization algorithms are presented for efficient resource management, including geometric disk cover for user grouping, greedy algorithm for user association, and two minimum enclosing circle methods for beam optimization (i.e., Welzl's algorithm and heuristic approach). Additionally, a closed-form solution for optimal power allocation is provided. Algorithm 2 integrates these methods to maximize overall sum rate while ensuring QoS, fairness, and power constraints. System performance is analyzed through SINR, sum rate, spectral efficiency, energy efficiency, user fairness, and outage metrics. This work highlights the system design and parameter optimization of a HAPS communication system to serve an ultra-wide coverage area, signifying stratospheric aerial communication platforms as the promising candidates for global coverage.

VIII. ACKNOWLEDGMENT

This work is supported by the King Abdullah University of Science and Technology (KAUST) Office of Sponsored Research

APPENDIX A

OUTAGE PROBABILITY DERIVATION

The probability density function of Rician distributed channel coefficient $|g|$ in (7) can also be written as a function of shape parameter K_s and accumulative power coefficient Ω :

$$f(x | K_s, \Omega) = 2x\xi \exp(-K_s + \xi x^2) I_0(2x\sqrt{K_s\xi}), \quad (49)$$

where $\xi = (K_s + 1) / \Omega$. Using the transformation of variables $|g|^2$, we can get

$$f(y | K_s, \Omega) = \xi \exp(-K_s + \xi y) I_0(2\sqrt{K_s\xi y}), \quad (50)$$

Moreover, the cumulative distribution function of $|g|^2$ can be written in variable y , $F(y | K_s, \Omega) = Pr(|g|^2 \leq y)$, as

$$F(y | K_s, \Omega) = 1 - Q_1\left(\sqrt{2K_s}, \sqrt{2\xi y}\right) \quad (51)$$

Given the scaling factors and conditioned on d_l^m , the CDF of $|h|^2 = |g|^2 G_l^m / L(d_l^m)$ can be expressed as

$$F(y | K_s, \Omega) = 1 - Q_1\left(\sqrt{2K_s}, \sqrt{2\xi \frac{yL(d_l^m)}{G_l^m}}\right) \quad (52)$$

Thus, the outage probability can be evaluated using the CDF in (52) as

$$\text{OP}_l^m = Pr(|h_l^m|^2 \leq \Psi_{\max}^m) = F(\Psi_{\max}^m | K_s, \Omega) \quad (53)$$

This yields the outage probability expression in (29).

REFERENCES

- [1] S. Dang, O. Amin, B. Shihada, and M.-S. Alouini, "What should 6G be?" *Nature Electronics*, vol. 3, no. 1, pp. 20–29, 2020.
- [2] A. Anjum, "Adopting a strategy of urgency to achieve cyber resilience," *NDU Journal*, vol. 36, pp. 26–37, 2022.
- [3] E. Yaacoub and M.-S. Alouini, "A key 6G challenge and opportunity—Connecting the base of the pyramid: A survey on rural connectivity," *Proceedings of the IEEE*, vol. 108, no. 4, pp. 533–582, 2020.
- [4] F. Tariq, M. R. Khandaker, K.-K. Wong, M. A. Imran, M. Bennis, and M. Debbah, "A speculative study on 6G," *IEEE Trans. Wireless Commun.*, vol. 27, no. 4, pp. 118–125, 2020.
- [5] X. Cao, P. Yang, M. Alzenad, X. Xi, D. Wu, and H. Yanikomeroglu, "Airborne communication networks: A survey," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 9, pp. 1907–1926, 2018.
- [6] A. Alsharoua and M.-S. Alouini, "Improvement of the global connectivity using integrated satellite-airborne-terrestrial networks with resource optimization," *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5088–5100, 2020.
- [7] S. Karapantazis and F. Pavlidou, "Broadband communications via high-altitude platforms: A survey," *IEEE Commun. Surveys Tuts.*, vol. 7, no. 1, pp. 2–31, 2005.
- [8] Y. Xing, F. Hsieh, A. Ghosh, and T. S. Rappaport, "High altitude platform stations (HAPS): Architecture and system performance," in *93rd Vehicular Technology Conference (VTC2021-Spring)*. Helsinki, Finland: IEEE, 2021, pp. 1–6.
- [9] A. Mohammed, A. Mehmood, F.-N. Pavlidou, and M. Mohorcic, "The role of high-altitude platforms (HAPS) in the global wireless connectivity," *Proceedings of the IEEE*, vol. 99, no. 11, pp. 1939–1953, 2011.
- [10] S. C. Arum, D. Grace, P. D. Mitchell, M. D. Zakaria, and N. Morozs, "Energy management of solar-powered aircraft-based high altitude platform for wireless communications," *Electronics*, vol. 9, no. 1, p. 179, 2020.
- [11] F. Hsieh, F. Jardel, E. Visotsky, F. Vook, A. Ghosh, and B. Picha, "UAV-based multi-cell HAPS communication: System design and performance evaluation," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Taipei, Taiwan, Dec. 2020, pp. 1–6.
- [12] P. Ji, L. Jiang, C. He, Z. Lian, and D. He, "Energy-efficient beamforming for beamspace HAP-NOMA systems," *IEEE Commun. Lett.*, vol. 25, no. 5, pp. 1678–1681, 2020.
- [13] M. A. Azzahra and Iskandar, "NOMA signal transmission over millimeter-wave frequency for backbone network in HAPS with MIMO antenna," in *13th Int. Conf. on Telecommun. Sys., Services, Appl. (TSSA)*. Bali, Indonesia: IEEE, 2019, pp. 186–189.
- [14] A. Anjum, "Securitization of cyber space - from prism of deterrence strategy," *Authorea Preprints*, 2023.
- [15] Z. Lian, L. Jiang, C. He, and D. He, "User grouping and beamforming for HAP massive MIMO systems based on statistical-eigenmode," *IEEE Wireless Commun. Lett.*, vol. 8, no. 3, pp. 961–964, 2019.
- [16] S. Liu, H. Dahrouj, and M.-S. Alouini, "Joint user association and beamforming in integrated satellite-HAPS-ground networks," *IEEE Trans. Veh. Technol.*, vol. 73, no. 4, pp. 5162–5178, 2024.
- [17] W. K. New, C. Y. Leow, K. Navaie, Y. Sun, and Z. Ding, "Interference-aware NOMA for cellular-connected UAVs: Stochastic geometry analysis," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 10, pp. 3067–3080, 2021.
- [18] W. K. New, C. Y. Leow, K. Navaie, and Z. Ding, "Aerial-terrestrial network NOMA for cellular-connected UAVs," *IEEE Trans. Veh. Technol.*, vol. 71, no. 6, pp. 6559–6573, 2022.
- [19] A. Nauman and M. Maqsood, "System design and performance evaluation of high altitude platform: Link budget and power budget," in *19th Int. Conf. on Adv. Commun. Tech. (ICACT)*. PyeongChang, Korea (South): IEEE, 2017, pp. 138–142.
- [20] O. B. Yahia, E. Erdogan, and G. K. Kurt, "HAPS-assisted hybrid RF-FSO multicast communications: Error and outage analysis," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 1, pp. 140–152, 2022.
- [21] M. S. Alam, G. K. Kurt, H. Yanikomeroglu, P. Zhu, and N. D. Dao, "High altitude platform station based super macro base station constellations," *IEEE Commun. Mag.*, vol. 59, no. 1, pp. 103–109, 2021.
- [22] U. S. Atmosphere, *US standard atmosphere*. National Oceanic and Atmospheric Administration, 1976.
- [23] O. Abbasi, A. Yadav, H. Yanikomeroglu, N.-D. Dao, G. Senarath, and P. Zhu, "HAPS for 6G networks: Potential use cases, open challenges, and possible solutions," *IEEE Trans. Wireless Commun.*, vol. 31, no. 3, pp. 324–331, 2024.
- [24] Y. Furuse and G. K. Tran, "A study on a radio source location estimation system using high altitude platform stations (HAPS)," *Sensors*, vol. 24, no. 17, p. 5803, 2024.
- [25] D. Denkovski, M. Angelichinoski, V. Atanasovski, and L. Gavrilovska, "Rss-based self-localization framework for future wireless networks," *Wireless personal commun.*, vol. 78, pp. 1755–1776, 2014.
- [26] M. M. Azari, F. Rosas, and S. Pollin, "Cellular connectivity for UAVs: Network modeling, performance analysis, and design guidelines," *IEEE Trans. Wireless Commun.*, vol. 18, no. 7, pp. 3366–3381, 2019.
- [27] B. Liu, C. Jiang, L. Kuang, and J. Lu, "Joint user grouping and beamwidth optimization for satellite multicast with phased array antennas," in *IEEE Global Commun. Conf. (GLOBECOM)*, Taipei, Taiwan, Dec. 2020, pp. 1–6.
- [28] S. Javed, M.-S. Alouini, and Z. Ding, "An interdisciplinary approach to optimal communication and flight operation of high-altitude long-endurance platforms," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 6, pp. 8327–8341, 2023.
- [29] F. Hsieh and M. Rybakowski, "Propagation model for high altitude platform systems based on ray tracing simulation," in *13th IEEE Eur. Conf. Antennas Propag. (EuCAP)*, Krakow, Poland, 2019, pp. 1–5.

- [30] J. L. Cuevas-Ruiz and J. A. Delgado-Penín, "Channel modeling and simulation in HAPS systems," *Eur. Wirel.*, pp. 24–27, 2004.
- [31] A. G. Kanatas and A. D. Panagopoulos, *Radio wave propagation and channel modeling for earth-space systems*. CRC Press, 2017.
- [32] A. A. O. OLADIPO, "Stratospheric propagation and HAPs channel modeling," *MS Thesis at Blekinge Institute of Tech., School of Eng., Department of Telecommun. Sys.*, 2007.
- [33] M. Takahashi, Y. Kawamoto, N. Kato, A. Miura, and M. Toyoshima, "Adaptive power resource allocation with multi-beam directivity control in high-throughput satellite communication systems," *IEEE Wireless Commun. Lett.*, vol. 8, no. 4, pp. 1248–1251, 2019.
- [34] A. S. Academy, "Space communication calculations," Jan. 2017.
- [35] Y. Shibata, N. Kanazawa, M. Konishi, K. Hoshino, Y. Ohta, and A. Nagate, "System design of Gigabit HAPS mobile communications," *IEEE Access*, vol. 8, pp. 157 995–158 007, 2020.
- [36] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, "On the performance of non-orthogonal multiple access in 5G systems with randomly deployed users," *IEEE Signal Process. Lett.*, vol. 21, no. 12, pp. 1501–1505, 2014.
- [37] H. Bauer, "Minimal places of functions and extremal points," *Archive of Mathematics*, vol. 9, no. 4, pp. 389–393, 1958.
- [38] E. Welzl, "Smallest enclosing disks (balls and ellipsoids)," in *1991 Proc. New Results and New Trends in Computer Science*. Graz, Austria: Springer, 2005, pp. 359–370.
- [39] K. Wang, Y. Liu, Z. Ding, A. Nallanathan, and M. Peng, "User association and power allocation for multi-cell non-orthogonal multiple access networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5284–5298, Nov. 2019.



Mohamed-Slim Alouini (S'94, M'98, SM'03, F'09) was born in Tunis, Tunisia. He earned his Ph.D. from the California Institute of Technology (Caltech) in 1998 before serving as a faculty member at the University of Minnesota and later at Texas A&M University at Qatar. In 2009, he became a founding faculty member at King Abdullah University of Science and Technology (KAUST), where he currently is the Al-Khwarizmi Distinguished Professor of Electrical and Computer Engineering and the holder of the UNESCO Chair on Education to Connect the

Disconnected. Dr. Alouini is a Fellow of the IEEE and OPTICA and his research interests encompass a wide array of research topics in wireless and satellite communications. He is currently particularly focusing on addressing the technical challenges associated with information and communication technologies (ICT) in underserved regions and is committed to bridging the digital divide by tackling issues related to the uneven distribution, access to, and utilization of ICT in rural, low-income, disaster-prone, and hard-to-reach areas.



Sidrah Javed (S'16, M'21) received her Bachelor of Engineering (B.E.) degree in Electrical (Telecommunication) Engineering from the National University of Science and Technology (NUST), Pakistan, in 2012. Following her graduation, she served as a Research Engineer in the Research and Development (R&D) Department at the National Radio and Telecommunication Corporation (NRTC), Pakistan, from 2012 to 2015. Dr. Javed pursued MS/PhD from 2015 to 2021 at King Abdullah University of

Science and Technology (KAUST), Saudi Arabia. In 2023, she contributed her expertise as a Research Consultant at KAUST. Currently, Dr. Javed is a Postdoctoral Research Associate in the Department of Engineering at Durham University, U.K. Her current research interests include modeling, design, and performance analysis of wireless communication systems especially satellite-aerial-terrestrial hybrid communications to eliminate digital inequality.



Citation on deposit: Javed, S., & Alouini, M.-S. (online). System Design and Parameter Optimization for Remote Coverage from NOMA-based High-Altitude Platform Stations (HAPS). IEEE Transactions on Wireless Communications, <https://doi.org/10.1109/twc.2024.3508872>

For final citation and metadata, visit Durham Research Online URL:

<https://durham-repository.worktribe.com/output/3214019>

Copyright statement: This accepted manuscript is licensed under the Creative Commons Attribution 4.0 licence.

<https://creativecommons.org/licenses/by/4.0/>