
Article:
Zakaria, Muhammad Danial Bin, Grace, David orcid.org/0000-0003-4493-7498, Mitchell, Paul Daniel orcid.org/0000-0003-0714-2581 et al. (2 more authors) (2019) Exploiting user-centric joint transmission coordinated multipoint with a high altitude platform system architecture. IEEE Access. pp. 1-16. ISSN 2169-3536

https://doi.org/10.1109/ACCESS.2019.2905844

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received February 20, 2019, accepted March 5, 2019, date of publication March 20, 2019, date of current version April 5, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2905844

Exploiting User-Centric Joint Transmission – Coordinated Multipoint With a High Altitude Platform System Architecture

MUHAMMAD D. ZAKARIA[1,2], DAVID GRACE[1], (Senior Member, IEEE), PAUL DANIEL MITCHELL[1], (Senior Member, IEEE), TAREQ M. AL-SHAMI[1], AND NILS MOROZS[1], (Member, IEEE)

[1]Department of Electronic Engineering, University of York, York YO10 5DD, U.K.
[2]Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Tembila Campus, Malaysia

Corresponding author: Muhammad D. Zakaria (mdz501@york.ac.uk)

This work was supported in part by the European Union's Horizon 2020 Research Programme through the 5GAURA Project under Grant 675806.

ABSTRACT User-centric joint transmission coordinated multipoint (JT-CoMP) has been shown to enhance the capacity of terrestrial cellular systems, by overcoming cell-edge interference. This paper investigates how JT-CoMP can be extended to a new high altitude platform (HAP) system architecture by exploiting a phased array antenna, which generates multiple beams that form cells, each of which can map on to pooled virtual base-station equipment, thereby replacing multiple terrestrial cell sites. The strategy to implement CoMP is designed to best enhance the user experience at the edge of the HAP cells, including the overall performance of the system. Methods to overcome the known tradeoff for JT-CoMP between carrier-to-interference plus noise ratio (CINR) gain and loss of capacity accessible to the users are considered. Two different methods of identifying non-CoMP and CoMP users are introduced based upon the centralized CINR threshold and flexible CINR threshold approaches. For the bandwidth allocation technique, two approaches are used: full bandwidth (FBW) and half bandwidth (HBW). These four approaches are combined, delivering the FBW, HBW, Flex FBW, and Flex HBW schemes that are used to control the JT-CoMP. It is shown that 57% and 45% of users gain benefit from the use of HBW and FBW, respectively. Overall, the schemes based on the flexible CINR threshold approach provide the best balance between loss and gain of the user capacity, while the centralized CINR threshold-based schemes performed well, beneficiary up to 57% of the users, but with the drawbacks of a higher percentage of losing users.

INDEX TERMS Cell mapping, user-centric, CoMP, interference mitigation, HAP, cellular networks, resource allocation.

I. INTRODUCTION

High Altitude Platforms (HAPs) are widely regarded as a flexible, mobile, cost effective and alternative way to provide wireless communication services (e.g. broadband and cellular services) [1], [2]. HAPs are airships or aircraft, operating in the stratosphere approximately 17-22 km above ground [3], [4]. This height is well above commercial airplanes and suffers from less atmospheric turbulence than lower altitudes. HAPs not currently in service can be repositioned to replace failed communication infrastructure and provide extra coverage and capacity when needed for a temporary large crowd event [5]. In terms of permanent service, HAPs are suited to fill the gap in coverage in areas lacking in terrestrial infrastructure. HAPs have the potential to provide a useful alternative to the traditional terrestrial provision because their higher altitude operation provides a better chance of achieving Line of Sight (LoS) connectivity. With a multi-beam deployment capability (each beam can form a cell), a HAP will need less infrastructure to serve more users over a larger service area compared to a corresponding terrestrial system [6]. Many big business entities

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

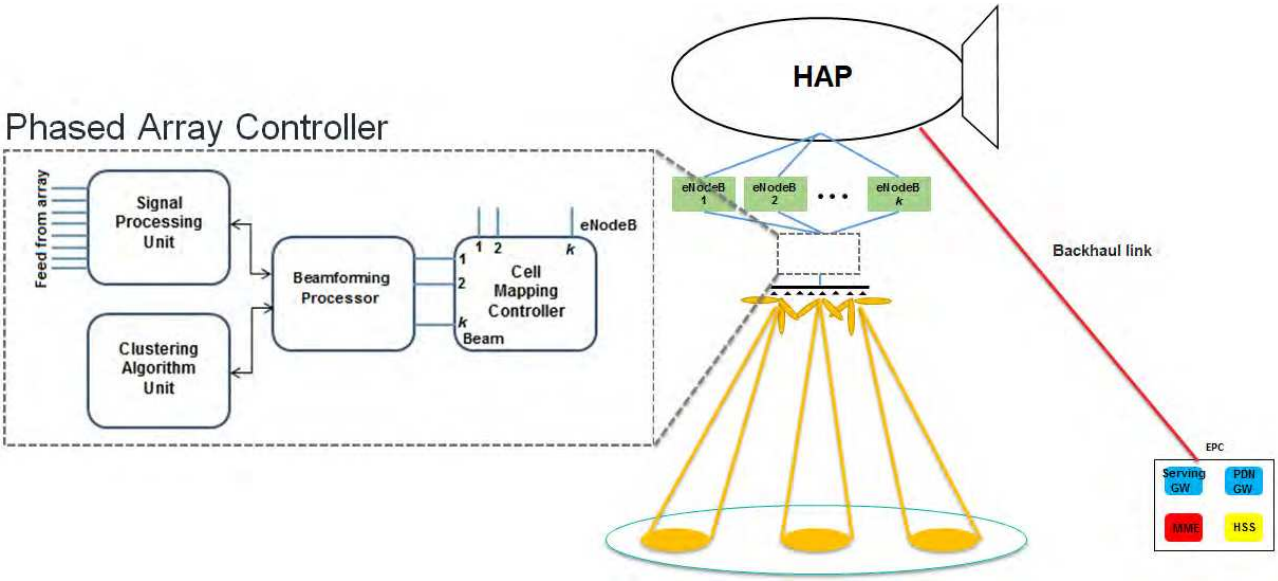

FIGURE 1. HAP system architecture with phased array antenna and pooled virtual eNodeBs mapped onto directional beams generated by the phased array controller.

are currently involved with HAP technologies in both the aeronautic and service sectors. Facebook has partnered with Airbus to perform telecom tests towards end of 2018, using its Zephyr S HAP [7]. This follows earlier development of an aircraft as part of its Aquila program, which successfully completed its second flight test on May 22nd 2017 [8]. Airbus recently flew its Zephyr S continuously for over 23 days, far exceeding the flight endurance record for aircraft [9]. Airbus is also scheduled to have a flight test with their Zephyr T this year [10], while AeroVironment Inc., globally known as a leader in unmanned aircraft systems (UAS) recently announced a joint venture project worth $65 Million with Softbank [11].

The advantage of CoMP is not only to reduce the number of interfering signals, but also in some cases to convert them into useful signals. There are three types of CoMP - Coordinated Scheduling (CS), Coordinated Beamforming (CB), and Joint Transmission (JT). CoMP was first introduced by the third generation partnership project (3GPP) release 11 [12], in order to mitigate inter-cell interference particularly at the cell-edge, thereby improving the capacity of the cell-edge users. The focus in this paper is JT-CoMP which enables two or more simultaneous data transmissions to an intended user in the downlink case. JT-CoMP requires synchronization between the cooperative cells. It can be achieved with the centralization of all virtual eNodeBs within the HAP, so JT-CoMP is an appropriate choice compared to CS and CB CoMP.

In achieving tight synchronization between the terrestrial eNodeBs, special measures need to be carried out especially for a distributed system. It can be achieved by centralization with Cloud-Random Access Network (C-RAN) or a very tight clock synchronization for distributed eNodeBs which can be very complex in term of overhead. An obvious advantage of a HAP is that the virtual eNodeBs are collocated, using a common clock, thus achieving tight synchronization to enable JT-CoMP. To implement JT-CoMP in a HAP system, a new architecture is proposed in order to deliver the coordination needed between important entities. Figure 1 shows the proposed HAP system architecture. The HAP is equipped with a phased array antenna. The phased array controller acts as the entity to control and connect the antenna array beams with the virtual E-UTRAN Node Bs (eNodeBs). The controller consists of a signal processing unit which connects directly to the antenna array. This unit is responsible for setting the weights of the individual antenna elements in order to perform beamforming. The second unit is the Beamforming Processor. This processor acts as the central unit where information about the HAP beams and associated user information (e.g. CINR levels) are collected, processed and forwarded to all connected units. The third unit is the Clustering Unit, which in this work adopts the K-means clustering algorithm. It clusters the users in order to optimize the beam pointing location. This information is passed to the Beamforming Processor. The generated beams aremapped onto the virtual eNodeBs using the Cell Mapping Controller, which manages the feed from the beamforming processor. Using virtual eNodeBs to manage the individual beams as cells provides equivalence to the traditional terrestrial system cell approach, enabling easy integration with existing (including hybrid) systems. Overall, from the access protocol perspective, the architecture behaves in the same way as a traditional terrestrial network. The configuration of this new architecture mirrors the traditional configuration, hence many aspects such as mobility will be handled the same way as in a terrestrial system. Co-location of virtual eNodeBs

on the same HAP provides tight clock synchronization and phase alignment which greatly assists with applications like CoMP and also facilitates handover. Alternatively, a HAP system could have completely centralized processing, where all beams are managed by a single eNodeB, as seen in massive MIMO applications. However, given the potential number of beams (delivering fully functional cells) that can be provided by a HAP system and the resulting capacity, this multiple eNodeB approach with separation of beamforming from higher level functions is much more scalable.

The purpose of this paper is to show how JT-CoMP can be integrated into a HAP system, and how JT-CoMP can increase the capacity of HAP cell-edge users by adapting HAP phased array antenna systems to better integrate with existing approaches to delivering cellular infrastructure. The novelty and contributions of this paper are:

- The introduction of the new HAP system architecture which integrates applications like JT-CoMP.
- A method to better balance the CINR gain and capacity loss trade-off via a new bandwidth allocation technique.
- A new flexible CINR threshold that better selects users who will benefit from CoMP.

HAPs can deploy multiple beams simultaneously, with each beam reusing the same spectrum, which causes interference between the cells as shown by Zakaria *et al.* [6], [13]. Typically, the users at the edge of the cell will experience most interference from the neighboring cells due to their closer proximity. This factor makes the user CINR levels vulnerable. Due to the interconnected layout of HAP cells, there is a trade-off where users will receive less bandwidth compared to when a system does not use CoMP. To solve this issue, we present four different schemes in order to find the appropriate group of users to be included in the CoMP region.

This paper is organized as follows. In section II a brief description of related work is presented. In section III the methods used to model the scenario are explained in detail. The performance of JT-CoMP from HAPs is discussed in section IV, finally, the paper is concluded in section V.

II. RELATED WORK

Significant HAP related communications research has been conducted in the last 20 years. Many models and architectures have been considered to address the scenarios that have been studied. In [14], the ITU recommended antenna profiles are compared with an adapted antenna pattern which exploits elliptic beam lens antenna in order to effectively provide a multi-beam, multi-cell communication network. Also using the ITU recommended antenna profile, Iskandar and Abubaker [15] evaluated an interference mitigation technique for a stratospheric platform (SPF) WiMAX downlink system. In [16], conventional aperture antennas with high directivity are investigated. Based on the simulation results, it is shown that power at the cell edge can be maximized when the cell edge roll-off is approximately 4.5 dB below the boresight gain.

Meanwhile in [17] an antenna array adopting Multiple-Input-Multiple-Output (MIMO) transmissions on HAPs is investigated. The MIMO application is well studied for HAPs. For example, Zakia [18] studied how to provide connectivity to a high-speed train using Ka-band. The aim was to identify the best antenna separation distance for 2×2 MIMO to provide appropriate multiplexing gain. Most of the research above does not specify how signaling from the antenna is handled, and whether a single or multiple eNodeBs are involved.

In the literature, user-centric approach is widely used with different adaptations to best fit the scenarios that are considered. Hashmi *et al.* [19] mentioned that quality of experience (QoE) is lacking in 5G requirements for a C-RAN system.). This is due to the quality of service (QoS) that varies significantly from center to the edge of the cell. To solve that, changing the perception of building a cell around the remote radio head (BS) to building a cell around the UE by using the user-centric approach is essential. In [20], the user-centric approach is adapted to a multihomed user which receives LTE and WiFi service simultaneously. Instead of conventionally using the network to determine the scheduling, Dandachi *et al.* [20] decide to let the UE decide how to split the packet based on the information provided by the network.

Significant research has been carried out on JT-CoMP for terrestrial networks. It has been shown in [21]–[23] that JT-CoMP can provide significant SINR gain; however, JT-CoMP consumes additional bandwidth, as a user that is served by JT-CoMP requires all of its cooperative BSs to reserve an identical physical resource block (PRB) to transmit the same data. This means if a PRB is reserved by one of a users' serving BS, none of the other cooperating BSs of this user can reuse it. As a result, resource allocation should be taken into account when the performance of JT-CoMP is investigated. User-centric JT-CoMP clustering is considered in this work as it has proven its superiority in improving cell-center and cell-edge throughput compared with static clustering [22], [24].

There has been a considerable number of research efforts on JT-CoMP with the aim of finding an optimal user-centric cluster size and allocating radio resources in an efficient way. Nevertheless, most research on JT-CoMP deals with developing an optimal user-centric cluster size and allocating the corresponding resources separately.

In [24], optimal and suboptimal user-centric clustering algorithms are proposed to enhance the performance of users located at the edge. The results have shown that the two proposed algorithms outperform static clustering in terms of average user throughput and cell edge throughput. The work in [25] has applied user-centric JT-CoMP clustering to tackle inter-cell interference in multi-tier networks. In the proposed approach, users can operate under two different modes: non-CoMP mode and CoMP mode. A user operates in CoMP mode only if its second strongest received power is comparable with its strongest received power. Nie *et al.* [26] proposed

a user-centric algorithm with the aim of maximizing energy efficiency in multi-tier networks. A user in the user-centric approach chooses the BSs that provide strong received signal strength as its cooperative BSs. Recently, Bassoy *et al.* [27] applied JT-CoMP in a decoupled control/data architecture with the objective of balancing the load and maximizing spectral efficiency. A user selects the n strongest BSs provided that n does not exceed a maximum user-centric clustering size.

Some research has addressed user-centric JT-CoMP clustering and resource allocation jointly. Liu *et al.* [28] proposed two-step joint user-centric clustering and resource scheduling in ultra-dense multi-tier networks. As a first step, game theory is utilized to design a load aware clustering algorithm. Based on the clustering results obtained in the first step, graph coloring is employed to allocate resources. Further recent work that addressed joint user-centric and resource allocation was presented in [29]. The authors investigated the influence of different power level difference values on the performance of CoMP and non-CoMP users. Taking resource allocation into account, user-centric JT-CoMP clustering is applied to address inter-cell interference in a cell-less architecture where a user is jointly served by multiple BSs and control/data planes are decoupled.

III. SYSTEM MODEL ENABLING JT-CoMP

In this section, the system model is discussed. It is used to describe how this HAP wireless communication system can enable JT-CoMP. In subsection A, the beamforming process is described in detail. This is the core process for pointing the beams and deploying them as cells. Then, how we calculate the link quality and link capacity based on the propagation model and how the cells are deployed is explained in subsection B. The following sub-sections focus on how JT-CoMP can be performed more efficiently. In sub-section C, using set theory, users are defined based on the region in which they are located. Then, the process of determining the CoMP region using CINR as a threshold is discussed in sub-section D. Finally in sub-section E, we propose methods on how to manage the bandwidth allocation between Non-CoMP and CoMP regions with the objective to enhance the cell-edge user performance and the overall performance.

A. PHASED ARRAY ANTENNA PROFILE

Phased array antennas are a robust and flexible solution for beamforming because of their capability to steer the beams towards desired target locations electronically, without requiring any physical displacement of the antennas. Beamforming is achieved by transmitting/receiving a copy of the same signal at many separate, closely spaced antenna elements, but with slightly different delays and phases. In this way, the transmit/receive signals from every antenna element, when combined together, add up in phase and are amplified in some directions, but cancel each other out due to phase differences in other directions.

In this paper, the phased array antennas are used to create mobile network cells by projecting highly directional beams from a HAP approximately 20 km above the ground. For example, Figure 2 shows a projection of the antenna pattern on the ground from a basic linear phased array located 20 km above ground. It consists of 25 antenna elements aligned with the horizontal axis, and spaced $\lambda/2$ apart, where λ is the wavelength of the narrowband carrier. A maximum $\lambda/2$ spacing between the antenna elements is common practice in phased array design to achieve a smooth antenna gain pattern without grating lobes.

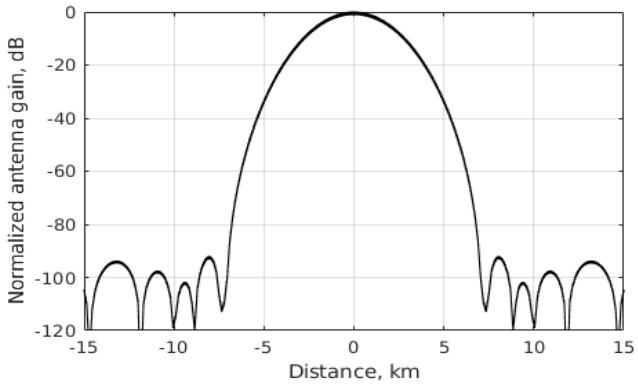

FIGURE 2. Ground projection of the antenna gain pattern from a 25-element linear phased array, located 20km above ground, using half-wavelength element spacing, and Blackman-Harris windowing.

Figure 2 demonstrates that the signals from all 25 elements of the phased array add up coherently at 0 km, i.e. directly underneath the phased array antenna, whereas the amplitude of the combined signal is smaller in other directions due to the phase misalignment of the signals from the individual antenna elements. The crucial features of the phased array antenna pattern are the width of the main lobe and the attenuation of the sidelobes, which will have a direct impact on the size of the cell and the inter-cell interference level respectively. These can be controlled by a range of windowing functions that scale the signals transmitted and received at the separate antenna elements by a particular pattern. For example, the beam pattern in Figure 2 uses the 25-element Blackman-Harris window which dramatically reduces the sidelobe levels to approx. -90 dB, but increases the width of the main lobe as a result.

In order to create and steer narrow beams providing mobile coverage to the users at an arbitrary location on the ground, a two-dimensional phased array is required, as depicted in Figure 3. Here, the direction of the beamforming target is determined in terms of angles in the XZ and YZ planes, i.e. vertical orthogonal planes aligned with the length and width of the phased array mounted on the HAP. In this way, an antenna pattern such as that shown in Figure 2, but in a desired angular direction, can be obtained separately in the XZ and YZ planes, with the resulting 3D pattern being the product of the two orthogonal patterns as shown in Figure 4.

Figure 4 shows an example of a beam being steered to the target location at (3, 5) km, aligned with the X and Y axes of the phased array antenna. A 25×25 element array is used

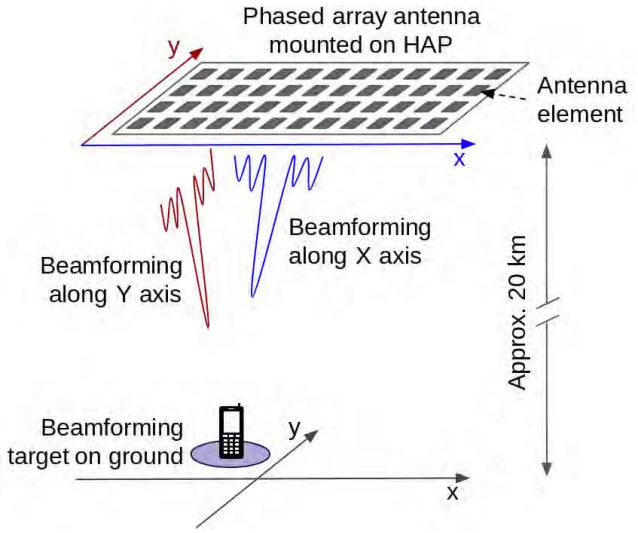

FIGURE 3. Beamforming to a target location on the ground using a horizontally orientated rectangular phased array antenna mounted on a HAP.

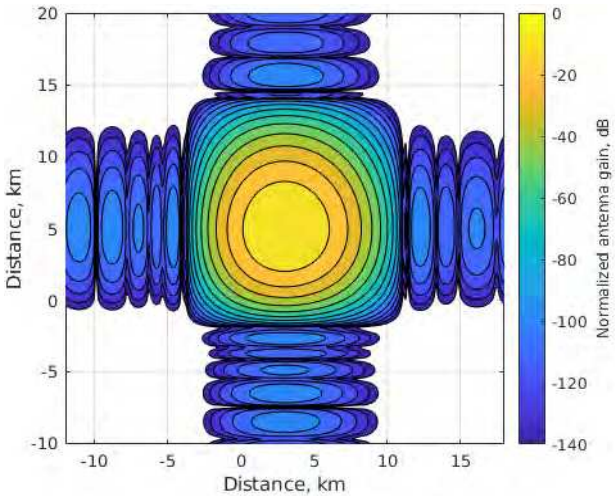

FIGURE 4. Ground projection of the antenna gain pattern from a 25 × 25 square phased array, located 20km above ground with the beamforming target at (3, 5) km, using half-wavelength element spacing, and Blackman-Harris windowing.

and mounted on a HAP 20 km above the ground, using the same Blackman-Harris window function, as in the example in Figure 2. It can be seen that the resulting beam pattern is a product of two orthogonal linear phased array antenna patterns along X and Y axis.

1) DERIVING THE BEAMFORMING COEFFICIENTS

The beamforming coefficients required to steer a beam to an arbitrary XY target location on the ground, as shown in Figure 4, can be derived as follows.

First, let θ_{XZ} and θ_{YZ} be the angles to the desired beamforming target in the XZ and YZ planes relative to the negative direction of the Z axis (vertically down), i.e. aligned with the X and Y axes of the square phased array, as shown in Figure 3. Note, the direction of the target is expressed in θ_{XZ} and θ_{YZ} instead of elevation and azimuth, because the phased array mounted on the HAP is horizontal. Therefore, θ_{XZ} and θ_{YZ}, i.e. an azimuth and elevation angle based polar coordinate system rotated 90°, are the angles that are aligned with the coordinates of the phased array. This allows us to derive the beamforming coefficients separately in two dimensions of the antenna array, thus breaking down the problem into two orthogonal linear phased array beamforming tasks.

Given the desired beamforming direction expressed by θ_{XZ} and θ_{YZ}, the phase shift that is applied at each linear phased array antenna element, i.e. a single linear array along the X axis or Y axis, assuming $\lambda/2$ spacing, is the following:

$$\beta[n] = -\pi(n-1)sin\theta, \tag{1}$$

where θ is the desired beamforming direction, nϵ[1,N] is the index of the antenna element in the linear phased array, and $\boldsymbol{\beta}$ is the vector of phase shifts applied to every element of the linear array, i.e. the signal is progressively delayed by the fixed $-\pi\, sin\theta_{(XZ/YZ)}$ increments along the X/Y dimension of the phased array to achieve beamforming in the desired direction of $\theta_{(XZ/YZ)}$.

In this way the overall complex beamforming weight of a linear phased array, using Blackman-Harris windowing, can be expressed as follows:

$$w[n] = w_B H(n, N) \cdot e^{j\beta[n]} \tag{2}$$

where N is the number of antenna elements in the linear phased array, $w_{BH}[n]$ is the n^{th} element of the N-element vector produced by the Blackman-Harris windowing function, and w[n] is the complex beamforming weight applied to the n^{th} antenna element in the linear phased array, i.e. incorporating both the amplitude scaling and the phase shift applied to the input signal.

Having determined the beamforming coefficients for the equivalent linear phased array along the X and Y axes of the overall antenna array, the overall beamforming coefficients for every antenna element in the square array can be calculated as follows:

$$w[n, m] = w_{XZ}[n] \cdot w_{YZ}[m] \tag{3}$$

where $\mathbf{w_{XZ}}$ and $\mathbf{w_{YZ}}$ are the vectors of linear phased array beamforming coefficients calculated in XZ and YZ planes separately using (2), and w[n,m] is the overall beamforming weight applied to the input signal at the nth element in X axis and m^{th} element in Y axis of the square array.

2) ANTENNA GAIN

The overall phased array antenna gain G in a particular direction expressed by the angles in XZ and YZ planes α_{XZ} and α_{YZ}, consists of two parts:

$$G(\alpha_{XZ}, \alpha_{YZ}) = G_E(\alpha_{XZ}, \alpha_{YZ}).G_{AF}(\alpha_{XZ}, \alpha_{YZ}) \tag{4}$$

where G_E $(\alpha_{XZ}, \alpha_{YZ})$ is the gain of a single antenna element in the direction defined by the α_{XZ} and α_{YZ} angles, and G_{AF}

$(\alpha_{XZ}, \alpha_{YZ})$ is the array factor, i.e. the additional directional gain achieved by phased array beamforming described in this section.

In this paper, we assume isotropic antenna elements, i.e. $G_E(\alpha_{XZ}, \alpha_{YZ}) = 1$, since we consider a large 25×25 phased array, where any small directionality of each antenna element will be negligible compared with the array factor.

Given that a set of beamforming coefficients $\boldsymbol{w}$ was derived for a particular boresight direction defined by the angles θ_{XZ} and θ_{YZ}, as described in the previous subsection, the array factor in a direction defined by α_{XZ} and α_{YZ} can be calculated as follows:

$$G_{AF}(\alpha_{XZ}, \alpha_{YZ}) = \frac{1}{N}\left|\sum_{n}\sum_{m} w[n,m].e^{(j\varphi(n,m,\alpha_X Z,\alpha_Y Z))}\right|^2 \tag{5}$$

which is a sum over the signals to/from every antenna element in the square phased array in the direction of α_{XZ} and α_{YZ}. The square of the amplitude of this overall sum, normalized by the number of antenna elements N, yields the power gain. It takes into account both the amplitude-phase beamforming weight w[n,m] applied at every antenna element, and the phase difference among the signals at every individual antenna element captured by the $\varphi(n, m, \alpha_{XZ}, \alpha_{YZ})$ function, caused by the slight difference in the time of arrival. Alternatively, the array factor $G_{AF}(\alpha_{XZ}, \alpha_{YZ})$ for a uniform rectangular phased array with $\lambda/2$ spacing can be calculated by splitting the 2D phased array problem into two linear phased arrays along the X and Y axis and multiplying the resulting patterns, as follows:

$$G_{AF}(\alpha_{XZ}, \alpha_{YZ}) = \frac{1}{N}\left|\begin{array}{l}(\sum_{n} w_{XZ}[n] \cdot e^{j}(n-1)\pi sin\alpha_{XZ}) \\ \times(\sum_{m} w_{YZ}[m] \cdot e^{j}(m-1)\pi sin\alpha_{YZ})\end{array}\right| \tag{6}$$

where $\boldsymbol{w_{XZ}}$ and $\boldsymbol{w_{YZ}}$ are the complex beamforming coefficients for the linear phased arrays derived in the XZ and YZ planes, as explained in the previous subsection. It reflects the fact that, due to the $\lambda/2$ element spacing, the signal arrives to/from every element in the given linear array with fixed phase lag increments of $\pi sin\alpha_{XZ/YZ}$. The gain value will be at its peak in the desired boresight direction, i.e. when $\alpha_{XZ} = \theta_{XZ}$ and $\alpha_{YZ} = \theta_{YZ}$, because the complex part of the $\boldsymbol{w_{XZ}}$ and $\boldsymbol{w_{YZ}}$ beamforming coefficients derived in (1) cancels out the phase lag only in the $(\alpha_{XZ}, \alpha_{YZ})$ direction, thus resulting in the coherent in-phase addition of the signal across the phased array.

B. PERFORMANCE METRICS

To identify the link quality, we measure the signal and interference experienced by the users in the system for downlink transmission by using parameters that are established and widely used. Based on the scenario in Figure 8, we use Carrier to Noise Ratio (CNR) and Carrier to Interference plus Noise Ratio (CINR). To measure both CNR and CINR, we must first measure the received power level of a user. The received power level, P_R at the Ue can be measured as follows:

$$P_R = \frac{(P_T.G_T.G_R)}{PL} \tag{7}$$

$$CNR = \frac{P_R}{P_N} \tag{8}$$

$$CINR = \frac{P_R}{P_N + \sum P_I} \tag{9}$$

where P_T is the transmit power emitted by the transmitter located at the HAP, G_T is the gain of the transmitter antenna of the HAP, G_R is the receiver antenna gain, and PL is the path loss which will be explained later in (11). While P_N is the noise power and $\sum P_I$ is the summation of the interference power from the neighboring cells. CINR is measured based on the ratio of the signal from the associated cell received by a user, and the total interference from the neighboring cells plus the noise power. The CINR levels of some users will change throughout the simulation as they will be included in CoMP region, so this will be the initial CINR for those users.

The channel capacity experienced by each user is determined by the extended Shannon equation based on [30] and can be described as follows:

$$C = \begin{cases} 0, & CINR_{dB} < 1.8 \\ \alpha B_c log_2\,(1 + CINR)\,, & 1.8 \le CINR_{dB} \le 22 \\ \alpha B_c log_2\,(1 + 158.5)\,, & CINR_{dB} > 22 \end{cases} \tag{10}$$

where α is the implementation loss that is set to be 0.65, B_c is the bandwidth per channel, and $CINR$ is in linear form.

Free Space Path Loss (PL) is considered for the HAP propagation model given the high minimum elevation angle, leading to a higher chance of achieving Line of Sight (LoS) connectivity.

$$PL(\text{dB}) = 20\log10(\text{d}) + 20\log10(f) + 92.45 \tag{11}$$

where d is the distance between HAP (transmitter) and UE (receiver) in km, and f is the carrier frequency in GHz.

C. SET THEORETIC USER DEFINITION

It is essential to correctly define the users within the systems especially when the condition of a user or how a user will operate depends on their specific location. A Venn diagram in Figure 5 represents the system in general (set S), and the users will be defined using the set theory. Sets A, B, and C represent the HAP cells. The sets are described as follows:

$$S = \{Ue : Ue\ is\ user\ demanding\ service\} \tag{12}$$

$$A \cup B \cup C = \{\mathcal{U}e_i | \mathcal{U}e_i \in S, Ue\,(CNR) \ge 9dB \tag{13}$$

$\mathcal{U}e$ is a user equipment placed randomly within the service area S and demanding wireless communication service. A $\mathcal{U}e$ demanding service will have to associate with a cell to be served, given that a certain CNR threshold is met. In the context of set theory, a $\mathcal{U}e$ needs to be included in either set A, B or C meeting the requirement of having at least CNR

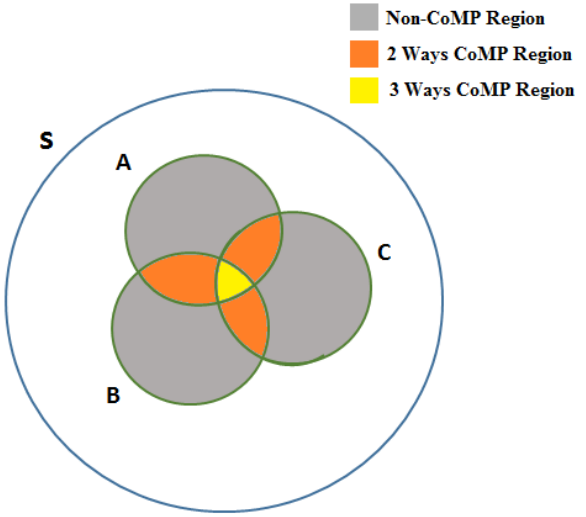

FIGURE 5. Venn diagram representing the service area and overlapping HAP cells.

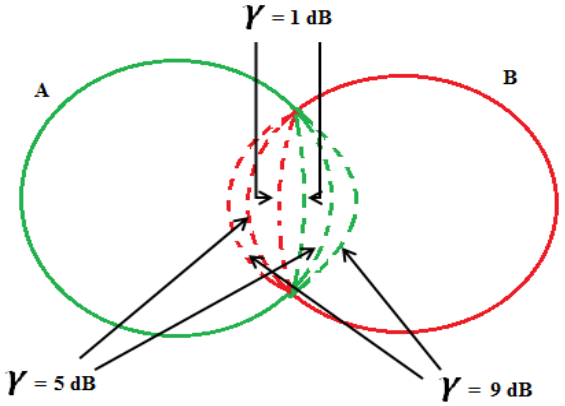

FIGURE 6. Illustration of the impact of different CINR thresholds in determining the overlapping region.

of 9 dB. A $\mathcal{U}e$ that belongs to either set A, B or C will be included in the system as $\mathcal{U}e_i$ as described in (13). In the intersection of sets as seen in Figure 5 a user suffers from a great deal of interference, which is why CoMP is needed to reduce the interference. When CoMP is applied, the sets are as follows:

$$A \cap B \,|C \cap A|\, B \cap C \geq \mathcal{C}_{2w} \tag{14}$$

$$\mathcal{C}_{2w} = \{\mathcal{U}e_c | \mathcal{U}e_c \in A \cup B \cup C, \mathcal{U}e_i\,(CINR) < \gamma \tag{15}$$

$$A \cap B \cap C \geq \mathcal{C}_{3w} \tag{16}$$

$$\mathcal{C}_{3w} = \{\mathcal{U}e_c | \mathcal{U}e_c \in A \cup B \cup C, \mathcal{U}e_c\,(CINR_c) < \gamma \tag{17}$$

$$A \ominus B \ominus C < \mathcal{N} \tag{18}$$

$$\mathcal{N} = \{\mathcal{U}e_n | \mathcal{U}e_n \in A \cup B \cup C, \mathcal{U}e_i(CINR) \geq \gamma \tag{19}$$

The intersections of set A and B, or C and A, or B and C as seen in the Venn diagram in Figure 5 represent the overlapping regions of the cells and known as the CoMP regions ($\mathcal{C}_{2w}$ and $\mathcal{C}_{3w}$). For the CoMP set, user $\mathcal{U}e_i$ with CINR lower than the CINR threshold γ will be included in the $\mathcal{C}_{2w}$. These users are then defined as $\mathcal{U}e_c$. $\mathcal{U}e_c$ with new $CINR_c$ is checked again whether it meet the CINR threshold requirement γ. If the $CINR_c$ is still lower than γ, then the user will be included in the $\mathcal{C}_{3w}$. Meanwhile, the users that belong to set A, B, or C but not their intersections are the $\mathcal{N}$ as shown in Figure 5, where $\ominus$ is the symmetric difference or disjunctive union in (18). These users typically have CINR at least equal to γ and are known as $\mathcal{U}e_n$.

D. CoMP USER CINR THRESHOLD (γ)

1) CENTRALIZED THRESHOLD

To determine appropriate CoMP users ($\mathcal{U}e_c$) and Non-CoMP users ($\mathcal{U}e_n$), we set a range of CoMP user CINR threshold levels (γ) centrally for all cells involved so that we can differentiate the performance of various sizes of CoMP and non-CoMP sets as illustrated in Figure 6. This enables an optimal value of γ to be determined. The set of cooperative cells ($\mathcal{C}c$) that serves the $\mathcal{U}e_c$ in this work is defined as follows:

$$Cc = \begin{cases} \{x_1\} & if\ CINR \geq \gamma\ (\mathcal{N}) \\ \{x_1, x_2\} & if\ CINR < \gamma\ (\mathcal{C}_{2w}) \\ \{x_1, x_2, x_3\} & if\ CINR_c < \gamma\ (\mathcal{C}_{3w}) \end{cases} \tag{20}$$

where x_1, x_2, x_3 are the cells that provide the strongest, second strongest, and third strongest received power level P_R to a particular user respectively, and γ is the CoMP user CINR threshold. The initial CINR is measured based on (9), to determine whether a user is heavily affected by the interference. $CINR_c$, is a re-measurement of CINR after taking into account nullifying the strongest interference from x_2 and possibly x_3 which are then turned into a useful signal. The $CINR_c$ can be defined as follows:

$$CINR_c = \frac{\sum_{j \in Cc} P_{R_j}}{P_N + \sum_{k=|H|; k \notin Cc} P_{I_k}} \tag{21}$$

where P_R in this case is the summation of the useful signals based on (7). Two signals from x_1 and x_2 will be added for a two way CoMP, and the third signal of x_3 will be added if the user is activated in a $\mathcal{C}_{3w}$ CoMP. Whilst the rest of the signals not included in $\sum P_R$, are $\sum P_I$, which are the remaining interference powers.

The steps to define a user are as follows:

Based on (20), a user with CINR less than γ will operate in $\mathcal{C}_{2w}$ mode as $\mathcal{U}e_c$, and will receive signals from both cell x_1 and x_2; otherwise the user will operate in a non-CoMP mode as $\mathcal{U}e_n$. It is because a user with CINR less than γ is regarded as highly affected by interference and located at the cell edge nearest to the strongest interference source. The $\mathcal{U}e_c$ will have their CINR re-calculated using (21) and again checked if the $CINR_c$ is less than γ. Not passing the threshold again will result in an activation of $\mathcal{C}_{3w}$ because of the possible location of the user closer to x_2 and at the same time closer to x_3 which means the interference is still high even after removing the interference and turning it into useful signal of the x_2. The use of CINR of the users to determine whether

Algorithm 1 Centralized Threshold.

For $\boldsymbol{Ue_i} \in \boldsymbol{S}$,

1. Calculate CINR based on (9).
2. **If** CINR $>$ $\boldsymbol{\gamma}$, **then**
3. $\boldsymbol{\mathcal{U}e_i = \mathcal{U}e_n}$
4. **Else**,
5. included in $\boldsymbol{\mathcal{C}_{2w}}$CoMP region; $\boldsymbol{\mathcal{U}e_i = \mathcal{U}e_c}$
6. **End**
7. Calculate $\boldsymbol{CINR_c}$ based on (21).
8. **If** $\boldsymbol{CINR_c > \gamma}$, **then**
9. Stays in $\boldsymbol{\mathcal{C}_{2w}}$ CoMP region.
10. **Else**, included in $\boldsymbol{\mathcal{C}_{3w}}$ region; recalculate $\boldsymbol{CINR_c}$.
11. **End**

the particular user is a $\mathcal{U}e_c$ or $\mathcal{U}e_n$ has never been used previously. This method is more straightforward as CINR is more widely used as the threshold to determine whether a user has a minimum quality of link needed to operate in allocated bandwidths.

2) FLEXIBLE THRESHOLD

Implementing a centralized threshold will affect some users in the system as a trade-off to maximize the capacity of the cell-edge users due to the inclusion of the non-beneficiary users in the CoMP set. A non-beneficiary user means a user that does not benefit from the implementation of CoMP because their CINR improvement only benefits the CoMP users and is achieved at the cost of a reduction in bandwidth needed to deliver CoMP. The boundary between beneficiary and non- beneficiary users varies for each of the cells as it depends on geographical factors of the users associated with each cell, thus the centralized approach cannot be used in solving this matter. So a flexible CoMP user CINR Threshold (γ) is proposed to deal with this unevenness. This flexible threshold means that each of the cells will have their own γ which is derived using the equation as follows:

Capacity per user with implementation of CoMP, C_c should be at least the same with the capacity per user before the implementation of CoMP, C.

$$Capacity, C \leq Capacity_{CoMP}, C_c \tag{22}$$

$$B_c log_2 (1 + CINR) \leq kB_c log_2 (1 + CINR_c) \tag{23}$$

To simplify,

$$\log_2 (1 + CINR) \leq k log_2(1 + CINR_c) \tag{24}$$

where k represents the fractional value of the initial bandwidth. The initial bandwidth is assumed to be 1. The variable k depends on the bandwidth allocation scheme that will be applied together with this flexible threshold. It will be explained more in detail in the later section. Using (24), we can acquire the suitable CINR levels that can be used as the threshold, γ for each cell. Redefining γ using (24) will help to reduce the number of non-beneficiary users included in the CoMP set.

The steps to get the γ for each cell are as follows:

Algorithm 2 Flexible Threshold

1) Determine all $\boldsymbol{\mathcal{U}e_i}$ of the cell whether they meet the requirement of equation (24).
2) **If** (TRUE), **then**
3) $\boldsymbol{\mathcal{U}e_i = \mathcal{U}e_{pass}}$
4) $\boldsymbol{\gamma = \min\left(CINR\ of\ all\ \mathcal{U}e_{pass}\right)}$
5) **Else** $\boldsymbol{\mathcal{U}e_i = \mathcal{U}e_n}$
6) **End**

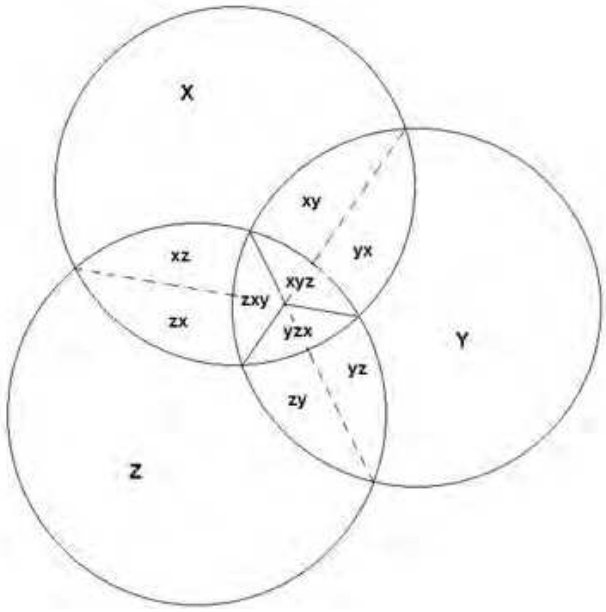

FIGURE 7. Overlapping cells [13].

E. BANDWIDTH ALLOCATION APPROACHES

Two types of overlapping region are considered in this work. According to the illustration in Figure 7, it can be seen that cells X, Y, and Z overlap with each other forming two different types of overlapping regions. An overlapping occurrence between two cells that forms the $\mathcal{C}_{2w}$ region creates two different sub-regions, for example the overlapping region of cells X and Y; sub-region xy where the users are associated to cell X as the primary and Y as the secondary cell, and sub-region yx where the users are associated to cell Y and X as the secondary cell. These sub-regions will have to be defined specifically even though they form one overlapping region because these sub-regions are likely to have a different number of users, which will have an impact on the bandwidth allocation. There are regions where the three cells overlap, and when a user is inside this area, they will experience increased interference, without CoMP. This is because the user is further away from the associated cell's center of coverage, while at the same time it is affected by two strong interfering sources. In this case, only removing one interference source and turning it into a useful signal will be insufficient to improve the CINR, hence the need to turn the second interference source into a useful signal, thereby

creating $\mathcal{C}_{3w}$. Based on Figure 7, sub-regions xyz, yzx, and zxy will form the three way CoMP region. The two types of overlapping region are defined as follows:

1) ***$\mathcal{C}_{2w}$CoMP region***– *An overlapping region involving two cells for example in Figure 7, it is formed by the xy and yx sub-region.*
2) ***$\mathcal{C}_{3w}$CoMP region***– *An overlapping region involving three cells for example in Figure 7, it is formed by the xyz, yzx, and zxy sub-region.*

Bandwidth allocation is non-trivial, especially when implementing CoMP in HAP systems due to the high degree of tessellation and overlap. It is also an important element in providing the balance between improving CINR and losing the capacity. The cooperating cells will have to agree to allocate the same exact bandwidth to the overlapping CoMP region, and the allocated bandwidth cannot be reused by the cooperating cells. A simple way to allocate the bandwidths to the CoMP and non-CoMP regions is to allocate X% of the available bandwidths to CoMP region while the remaining is allocated to the non-CoMP region. This method may be relatively simple, but will result in an inefficient use bandwidth because of the unevenness of the number of the users in the sub-regions. When encountering such a problem, we propose a strategy of using the number of users in both the CoMP and non-CoMP regions to decide what the ratio of bandwidths should be allocated between these regions. For the case of a $\mathcal{C}_{3w}$ region, the number of users in the cooperating two-way CoMP regions and the number of users in the $\mathcal{C}_{3w}$ region will be considered.

From this strategy of allocating bandwidths and based on diagram in Figure 7, we propose two different schemes for allocating the bandwidths as follows:

1) FULL BANDWIDTH (FBW) SCHEME

The FBW scheme allocates the same amount of bandwidth per user as what they would receive if there was no CoMP applied in the system. The FWB scheme can be computed as below:

For $\mathcal{C}_{2w}$;

$$BW_X = \frac{B}{|\mathcal{U}e_i^X|}|\mathcal{U}e_c^{xy}| \tag{25}$$

$$BW_Y = \frac{B}{|\mathcal{U}e_i^Y|}|\mathcal{U}e_c^{yx}| \tag{26}$$

$$BW_{CoMP_{XY}} = \min(BW_X, BW_Y) \tag{27}$$

$$FBW_2 = \frac{BW_{CoMP_{XY}}}{|\mathcal{U}e_c^{xy}| + |\mathcal{U}e_c^{yx}|} \tag{28}$$

This is an example of bandwidth allocation computation between cell X and Y. The total bandwidth for the CoMP region BW_X and BW_Y is calculated based on the total number of users and number of $\mathcal{U}e_c$ (e.g. number of users in xy ($\mathcal{U}e_c^{xy}$) and yx ($\mathcal{U}e_c^{yx}$) region respectively) in both cell X and Y. Both bandwidths are then compared between each other in (27) and the cell with lower bandwidth will be chosen. The other party will have to agree with the chosen bandwidth and allocate the same bandwidth to the CoMP region. The reason for this step is because with FBW the cell already offers the maximum bandwidth for the CoMP region, and going beyond that means that bandwidth for the Non-CoMP region will be sacrificed. Selecting the cell with the lower bandwidth offering will prevent the sacrificing of the Non-CoMP bandwidth. We then acquire the CoMP bandwidth per user for the $\mathcal{C}_{2w}$ region (FBW_2) by dividing the total bandwidth allocated to the CoMP region ($BW_{CoMP_{XY}}$) with the total number of users in the CoMP region in this case xy and yx sub-regions.

For $\mathcal{C}_{3w}$;

$$BW_{XY} = \frac{BW_{CoMP_{XY}}}{|\mathcal{U}e_c^{xy}| + |\mathcal{U}e_c^{yx}|}|\mathcal{U}e_c^{xyz}| \tag{29}$$

$$BW_{YZ} = \frac{BW_{CoMP_{YZ}}}{|\mathcal{U}e_c^{yz}| + |\mathcal{U}e_c^{zy}|}|\mathcal{U}e_c^{yzx}| \tag{30}$$

$$BW_{ZX} = \frac{BW_{CoMP_{ZX}}}{|\mathcal{U}e_c^{zx}| + |\mathcal{U}e_c^{xz}|}|\mathcal{U}e_c^{zxy}| \tag{31}$$

$$BW_{CoMP_{XYZ}} = \min(BW_{XY}, BW_{YZ}, BW_{ZX}) \tag{32}$$

$$FBW_3 = \frac{BW_{CoMP_{XYZ}}}{|\mathcal{U}e_c^{xyz}| + |\mathcal{U}e_c^{yzx}| + |\mathcal{U}e_c^{zxy}|} \tag{33}$$

In the case where a user activates the $\mathcal{C}_{3w}$, based on the illustration in Figure 7, it will involve three 2 way CoMP sub-regions and the bandwidth will be allocated from $BW_{CoMP_{XY}}$, $BW_{CoMP_{YZ}}$, and $BW_{CoMP_{ZX}}$. The bandwidth for the $\mathcal{C}_{3w}$ (e.g. BW_{XY}) will be decided based on the total number of $\mathcal{U}e_c$ (e.g. users in xy ($\mathcal{U}e_c^{xy}$) and yx ($\mathcal{U}e_c^{yx}$) sub-region) and the number of $\mathcal{C}_{3w}$ users (e.g. users in xyz ($\mathcal{U}e_c^{xyz}$) region). Just as in the $\mathcal{C}_{2w}$CoMP case, the lowest bandwidth among the three regions will be selected based on (32) for the same reason. If one of the sub-region results in zero bandwidth assignment, which means that there are zero $\mathcal{C}_{3w}$ users in that region, the $\mathcal{C}_{3w}$ region of cell X, Y, and Z will be shut down and all the other $\mathcal{C}_{3w}$ users from other sub-regions will be revert back to being $\mathcal{C}_{2w}$ users. This is to make sure that the sub-sections that have zero $\mathcal{C}_{3w}$ users do not need to reserve any bandwidth for the $\mathcal{C}_{3w}$ region which will results in degradation in the $\mathcal{C}_{2w}$ users' performance. Finally, the CoMP bandwidth per user for the $\mathcal{C}_{3w}$region (FBW_3) can be calculated by dividing the total bandwidth allocated ($BW_{CoMP_{XYZ}}$) with the total number of $\mathcal{C}_{3w}$ users ($\mathcal{U}e_c^{xyz}$, $\mathcal{U}e_c^{yzx}$, and $\mathcal{U}e_c^{zxy}$).

2) HALF BANDWIDTH (HBW) SCHEME

The HBW scheme allocates half of the amount of bandwidth per user compared with what they receive if no CoMP is applied in the system. The HBW scheme can be computed as follows:

For $\mathcal{C}_{2w}$;

$$BW_X = \frac{B}{|\mathcal{U}e_i^X| - |\mathcal{U}e_c^{xy}|/2}|\mathcal{U}e_n^X| \tag{34}$$

$$BW_Y = \frac{B}{|\mathcal{U}e_i^Y| - |\mathcal{U}e_c^{yx}|/2}|\mathcal{U}e_n^Y| \tag{35}$$

$$BW_{CoMP_{XY}} = \text{Max}(B - BW_X, B - BW_Y) \tag{36}$$

$$HBW_2 = \frac{BW_{CoMP_{XY}}}{|\mathcal{U}e_c^{xy}| + |\mathcal{U}e_c^{yx}|} \quad (37)$$

To calculate the $\mathcal{C}_{2w}$ HBW, we first find out the bandwidth for the non-CoMP region for both cells X and Y (BW_X and BW_Y) by dividing the total bandwidth of the system by the total number of users in the cell minus half of the total $\mathcal{U}e_c$ of that cell and multiply it with the number of $\mathcal{U}e_n$. To decide on the CoMP allocation, the CoMP bandwidth from both cells is the total bandwidth minus the bandwidth for non-CoMP region. The bandwidth in each cell is compared and the highest bandwidth that both cells can offer is assigned to the CoMP region, as in (36). This approach is different from the FBW scheme because for HBW only half of the bandwidth is considered, so if we choose the lowest bandwidth available the other cell that can offer more will have a much reduced bandwidth allocation. Hence, this will result in much lower bandwidth allocation for $\mathcal{U}e_c$, and fail to deliver the capacity improvements in many cases arising from the improved CINR. With the $BW_{CoMP_{XY}}$ decided, the CoMP bandwidth per user (HBW_2) can be calculated by dividing $BW_{CoMP_{XY}}$ with the total number of $\mathcal{U}e_c^{xy}$ and $\mathcal{U}e_c^{yx}$.

For $\mathcal{C}_{3w}$;

$$BW_{XY} = \frac{BW_{CoMP_{XY}}}{|\mathcal{U}e_c^{xy}| + |\mathcal{U}e_c^{yx}| - |\mathcal{U}e_c^{xyz}|/2} \times (|\mathcal{U}e_c^{xy}| + |\mathcal{U}e_c^{yx}| - |\mathcal{U}e_c^{xyz}| \quad (38)$$

$$BW_{YZ} = \frac{BW_{CoMP_{YZ}}}{|\mathcal{U}e_c^{yz}| + |\mathcal{U}e_c^{zy}| - |\mathcal{U}e_c^{yzx}|/2} \times (|\mathcal{U}e_c^{yz}| + |\mathcal{U}e_c^{zy}| - |\mathcal{U}e_c^{yzx}|) \quad (39)$$

$$BW_{ZX} = \frac{BW_{CoMP_{ZX}}}{|\mathcal{U}e_c^{zx}| + |\mathcal{U}e_c^{xz}| - |\mathcal{U}e_c^{zxy}|/2} \times (|\mathcal{U}e_c^{zx}| + |\mathcal{U}e_c^{xz}| - |\mathcal{U}e_c^{zxy}|) \quad (40)$$

$$BW_{CoMP_{XYZ}} = \text{Max}(BW_{CoMP_{XY}} - BW_{XY}, BW_{CoMP_{YZ}} - BW_{YZ}, BW_{CoMP_{ZX}} - BW_{ZX}) \quad (41)$$

$$HBW_3 = \frac{BW_{CoMP_{XYZ}}}{|\mathcal{U}e_c^{xyz}| + |\mathcal{U}e_c^{yzx}| + |\mathcal{U}e_c^{zxy}|} \quad (42)$$

To calculate the $\mathcal{C}_{3w}$ for the HBW case, all the cooperative sub-regions will calculate the bandwidth they can offer for the $\mathcal{C}_{3w}$ region by first determining the bandwidth for their 2 way sub-region (BW_{XY}, BW_{YZ}, and BW_{ZX}) by considering the number of $\mathcal{C}_{2w}$ users and $\mathcal{C}_{3w}$ users. Each of the bandwidth assignments can be offered to the $\mathcal{C}_{3w}$ region which is the total bandwidth allocated to the CoMP region (e.g. $BW_{CoMP_{XY}}$) minus bandwidth for the $\mathcal{C}_{2w}$region (e.g. BW_{XY}) are compared and the highest among the three offers are selected as the $\mathcal{C}_{3w}$ bandwidth ($BW_{CoMP_{XYZ}}$). Lastly, the $\mathcal{C}_{3w}$ region bandwidth per user (HBW_3) is calculated by dividing $BW_{CoMP_{XYZ}}$ with the total number of $\mathcal{C}_{3w}$ users ($\mathcal{U}e_c^{xyz}$, $\mathcal{U}e_c^{yzx}$, and $\mathcal{U}e_c^{zxy}$).

IV. RESULTS AND DISCUSSION

Figure 8 illustrates the system, where the HAP is located at the center of a 30 km radius service area at an altitude of 20 km above ground. The HAP cells are then deployed in the service area with overlapping areas between the cells. Three are shown, but in practice there can be many more. Two types of user are considered in the system - the non-CoMP user equipment ($\mathcal{U}e_n$) and CoMP user equipment ($\mathcal{U}e_c$) seen in Figure 8. Users are randomly distributed across the service area according to a uniform distribution. The HAP is considered to be equipped with 25×25 element planar phased array antenna which uses beamforming, which forms the multiple cells used to deliver wireless communication service. The locations of the cells are determined based on the clustering of users using the K-Means clustering algorithm. The algorithm determines the optimum centroid positions using the mean of clustered user's positions. The process of determining centroid positions will be achieved by integrating until the optimum point is reached. With this clustering algorithm, specific high density user groups can also be identified inside the service area according to the work in [5].

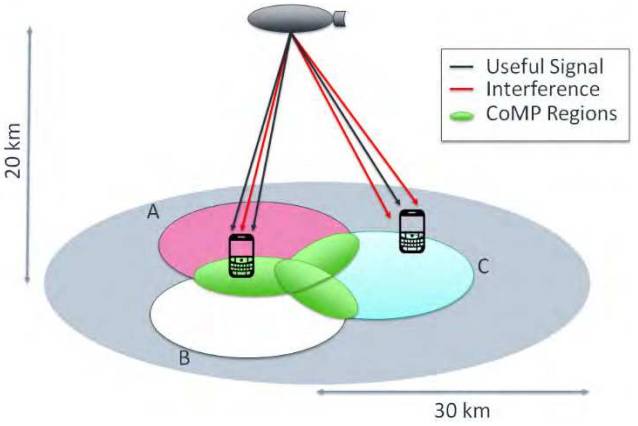

FIGURE 8. HAP cell footprints and the overlapping region as CoMP region.

TABLE 1. Simulation parameters.

Parameter	Value
HAP Transmit Power	40 dBm
Receiver Antenna Gain	0 dBi
HAP Antenna Gain (Boresight)	27.9 dBi
Carrier Frequency	2.6 GHz
Noise Power	-100 dBm
CNR Threshold	9 dB
CINR Threshold	0 dB
Number of Users	2900

In order to evaluate all the proposed methods and schemes, simulations using MATLAB were carried out based on the system layout in Figure 8. Traffic is modeled based on the full-buffer traffic model. The parameters are presented in Table 1 below:

Figure 9 shows the percentage of Ue_n and Ue_c for several threshold γ values from 0 to 19 dB used throughout the simulation. At γ of 0 dB, almost 0% users operate in CoMP while most of the users operate in the non-CoMP region, i.e. it can be assumed that the system operates with no CoMP.

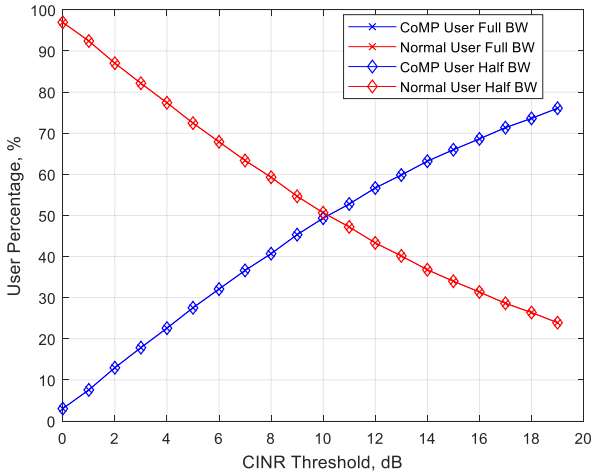

FIGURE 9. Percentage of CoMP and non-CoMP users with variation of the CINR threshold.

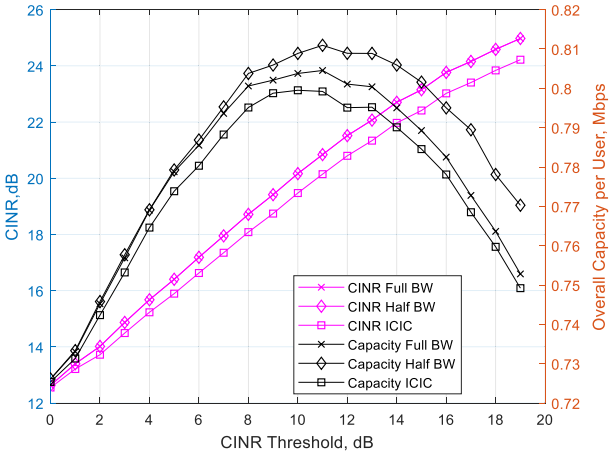

FIGURE 11. Mean CINR vs mean capacity per user for all schemes.

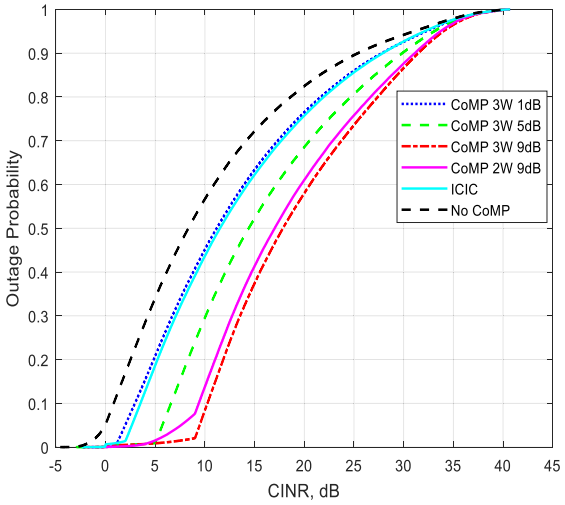

FIGURE 10. The outage probability of different γ with $\mathcal{C}_{3w}$, $\mathcal{C}_{2w}$ and ICIC.

As γ increases, the percentage of Ue_c increases, and contrarily, the percentage of Ue_n decreases. This is because the higher the γ, the more users that are included into the CoMP region, hence the increase of Ue_c and the decrease of Ue_n as a percentage.

Implementing CoMP means that users can improve their CINR levels. In Figure 10, the CDF plot of CINR levels of no CoMP, $\mathcal{C}_{3w}$ (1, 5, 9 dB), $\mathcal{C}_{2w}$ (9 dB), and inter-cell interference coordination (ICIC) are presented. The minimum CINR required for a user to be able to operate on a given channel is 1.8 dB, according to (20). It is shown that approximately 12 % of the included users ($\mathcal{U}e_i$) have a CINR below the operational CINR threshold. Implementing CoMP reduces the user outage which is a positive sign. It is also observed from the graph that with higher $\boldsymbol{\gamma}$, more users are included in the CoMP region, hence a better CINR performance. Besides that, it can be clearly seen that $\mathcal{C}_{3w}$ CINR is better than $\mathcal{C}_{2w}$ CINR. ICIC performs poorly compared with CoMP in terms of CINR performance. $\mathcal{C}_{3w}$ CINR should be higher than CINR because $\mathcal{C}_{3w}$ users will have one extra signal source (the addition of 3 signal sources) and one fewer interference source compared to $\mathcal{C}_{2w}$. For ICIC, one interference source will be removed because the bandwidth used for ICIC will not be reused by the neighboring cell [31], but it will not benefit from the simultaneous data transmission like CoMP.

From the previous graph, the higher the γ, the better the CINR performance improvement. However, there is a trade-off, because by including more users into the CoMP region in order to increase the CINR levels means that the amount of bandwidth that can be allocated per user is decreased. Despite the improvement shown in CINR levels, user capacity will reach its peak and the performance will start to decline. The mean CINR and mean capacity per user performance is presented in Figure 11 to directly compare the performance of CINR and capacity per user. It is shown that while the mean CINR keeps increasing with increasing of γ, the mean capacity for all schemes starts to drop after $\gamma = 10$ dB. The mean capacity starts to drop because at that point the system has started to include the users that have better performance without CoMP. These users receive less bandwidth when included in the CoMP region, and the CINR level increase cannot compensate for the reduction in bandwidth. The cut in bandwidth is also caused by the unevenness of the number of users in cooperative cells.

Figures 12, 13, 14, and 15 present the capacity difference which indicates whether the system benefits (positive difference) or loses (negative difference). Figure 12 shows the mean capacity difference of both FBW and HBW schemes for $\mathcal{U}e_i$, $\mathcal{U}e_c$ and $\mathcal{U}e_n$ users. The $\mathcal{U}e_n$ capacity difference keep increasing as the CINR threshold, γ increases, while the $\mathcal{U}e_c$ will reach a peak before having a degradation in capacity fairness. The $\mathcal{U}e_n$ capacity difference keeps increasing because when more users are being included into CoMP region, the $\mathcal{U}e_n$ will receive more bandwidth resulting in less bandwidth sharing. On the other hand, $\mathcal{U}e_c$ performance degrades at a certain point of the simulation because the users

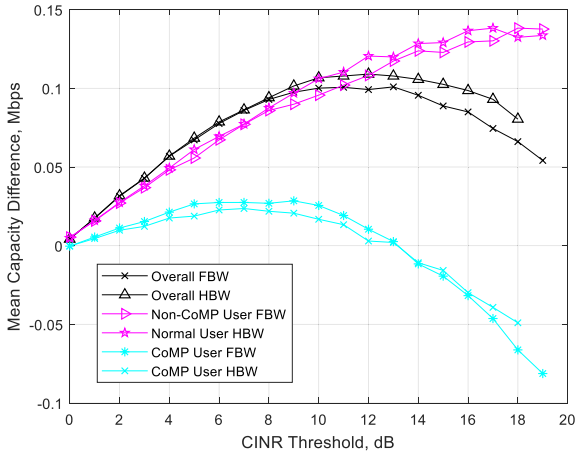

FIGURE 12. The average capacity difference of FBW and HBW for all types of users.

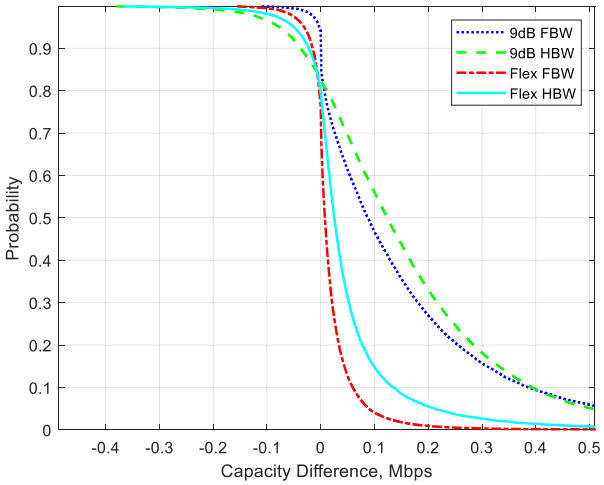

FIGURE 14. Non-CoMP users capacity difference CCDF for all schemes.

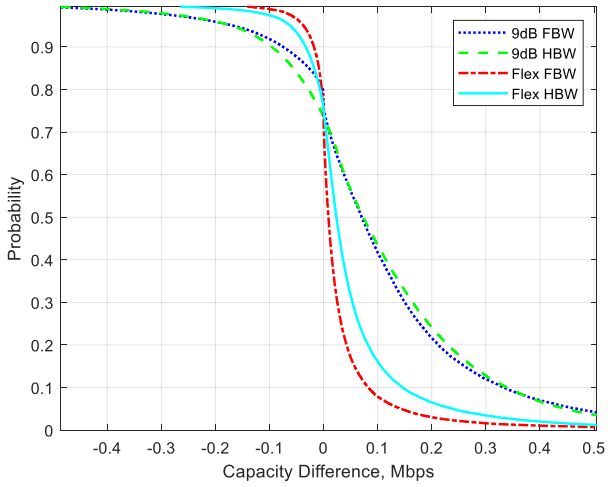

FIGURE 13. Overall users capacity difference CCDF for all schemes.

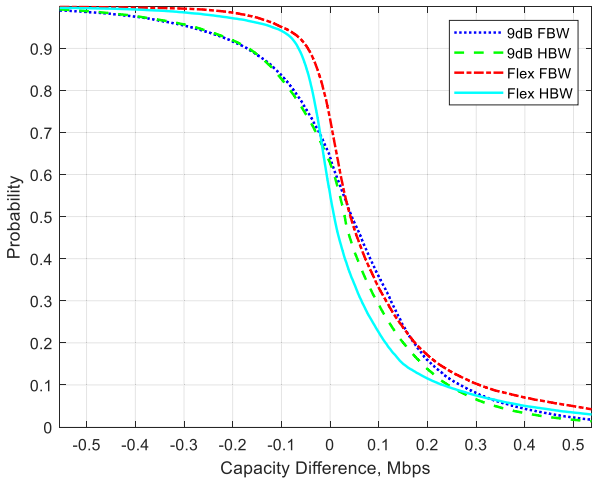

FIGURE 15. CoMP users capacity difference CCDF for all schemes.

that can perform better without CoMP start to be included. This kind of user receives less bandwidth with CoMP, and the CINR improvement is not sufficient to compensate for the bandwidth loss.

To go deeper into the behavior of the capacity difference, Complementary CDF (CCDF) graphs are presented in Figure 13, 14, and 15 for $\mathcal{U}e_i$, $\mathcal{U}e_c$ and $\mathcal{U}e_n$ respectively. The capacity differences of four different schemes are compared to establish the most suitable scheme to use in this scenario. The schemes are FBW with 9 dB γ, HBW with 9 dB γ, Flexible FBW, and Flexible HBW. The 9 dB performance threshold was chosen for both FBW and HBW because it was determined based on Figure 11, which illustrates that it is an optimal value of γ. For the $\mathcal{U}e_i$ capacity difference, it is shown in Figure 13 that 9 dB HBW has both the highest increase and decrease in performance. This is followed by 9 dB FBW, Flex HBW, and Flex FBW respectively. The users in the system, $\mathcal{U}e_i$ form two different user groups when CoMP is applied to the system - the CoMP users ($\mathcal{U}e_c$) and non-CoMP users ($\mathcal{U}e_n$). This graph represents the overall performance.

In Figure 14, the $\mathcal{U}e_n$ capacity difference is presented. All the schemes show that 80% of users benefit from CoMP, with the 9 dB HBW scheme being the best followed by 9 dB FBW, Flexible HBW, and Flexible FBW. Both HBW based schemes are better than the FBW based schemes because with HBW, $\mathcal{U}e_c$ are only allocated half of what they originally get without CoMP which leaves the $\mathcal{U}e_n$ extra bandwidth. While on the negative difference side, both HBW based schemes perform worse than the FBW based schemes because of the maximum value agreement as discussed in section III subsection F based on equation (35). Some $\mathcal{U}e_n$ of HBW will lose more bandwidth compared with $\mathcal{U}e_n$ using FBW.

For the $\mathcal{U}e_c$ capacity difference, a CDF graph is presented in Figure 15. It can be seen that Flex FBW outperforms the other schemes by having 75% of $\mathcal{U}e_c$ beneficiary from CoMP, while having less degradation (negative difference) compared with other schemes. The 9 dB FBW and 9 dB HBW have

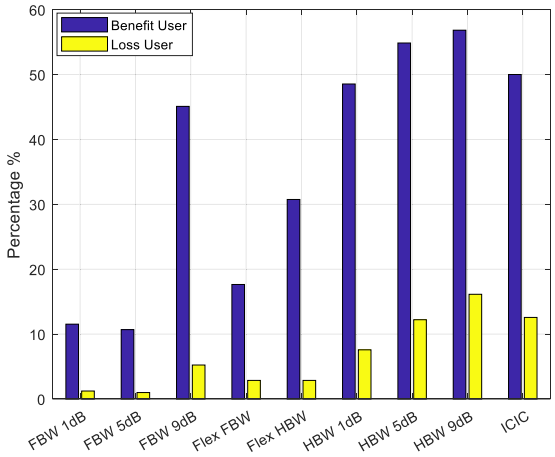

FIGURE 16. Benefit and loss trade-off for all schemes.

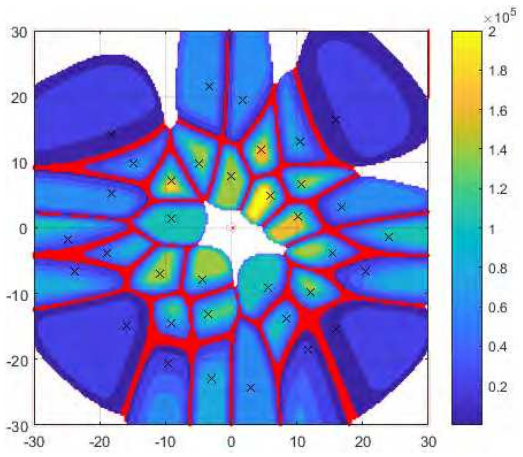

FIGURE 17. Contour plot of HAP cells. The dark blue to yellow regions indicate the lowest to highest capacity per user respectively, 'X' marks are the center of the HAP cells, and red regions are where the users have CINR levels of below 1.8 dB before implementation of CoMP in 30 km service area.

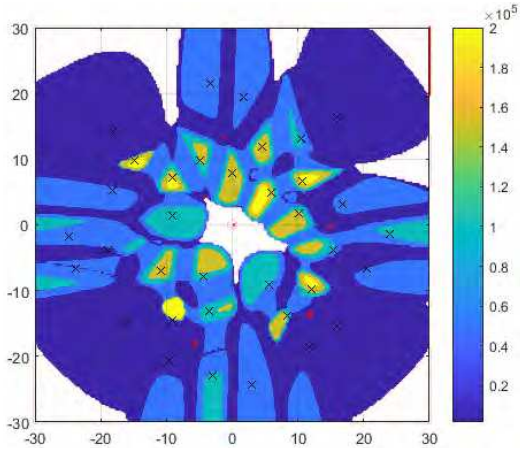

FIGURE 18. Contour plot of HAP cells. The dark blue to yellow regions indicate the lowest to highest capacity per user respectively, 'X' marks the center of the HAP cells, and red regions show where the users have CINR levels below 1.8 dB after implementation of CoMP with FBW (γ 9 dB) in 30 km service area.

almost the same performance with both having a great loss of capacity, while 9 dB FBW has slightly better than 9 dB HBW in terms of the capacity gain. For the $\mathcal{U}e_c$, it is expected that the FBW based scheme has better performance because of the nature of the scheme, which allocates more bandwidth to the CoMP region compared to the HBW.

In Figure 16 the percentage of beneficiary users which represents the users that have at least 20% capacity increase. While the percentage of losing users represents the users that have at least 20% capacity loss. The 20% threshold for both capacity increase and loss is used because under 20%, the increase or loss in capacity is considered insufficient to warrant CoMP. This measure can be used to help determine which scheme works best, because the trade-off between beneficiary and losing users for each scheme can be compared directly. For the case of HBW and FBW the parameters of 1, 5, and 9 dB $\boldsymbol{\gamma}$ is used to show the effects of using different threshold level. The scheme with the fewest drawbacks is the FBW, however the beneficiary user percentage is not that impressive. The highest percentage of beneficiary users occurs with the HBW scheme, however it also has more drawbacks. Obviously the best possible performance is to have maximum beneficiary users and very few losing users but this depends on what is valued for the system. In terms of capacity increase, the Flex HBW is better than the Flex FBW while the capacity decrease is similar for both.

Figure 17 presents a coverage plot showing the HAP cells covering the service area for approximately 96% of the total users, prior to CoMP. The white area is the area that is not covered by the HAP cells. The red color in between the cells illustrates the region where the users have a CINR level lower than the operational threshold (1.8 dB). It also represents the overlapping region of the cells. The color bar in Figure 17 and 18 represents capacity per user in bits per second.

After implementing CoMP, certain areas are improved as seen in Figure 18. An obvious improvement can be seen is that almost all red marks that represent a user having CINR below 1.8 dB are removed. This is an indication that the CINR of the cell edge users has been improved. In terms of the capacity increase, a clear difference can be seen in Figure 19 and 20 below.

The spatial effects of implementing CoMP with FBW, γ of 9 dB in a HAP system can be seen in Figure 19. It is clearly shown how the overlapping region is improved after being significantly affected by the interference as seen in Figure 17 earlier. As previously discussed, this is where users are located which have a degradation in performance when CoMP is applied. The darker region represents the area where the users have a degradation. From the authors' perspective this sacrifice can be made when it is important to have consistent wide area coverage.

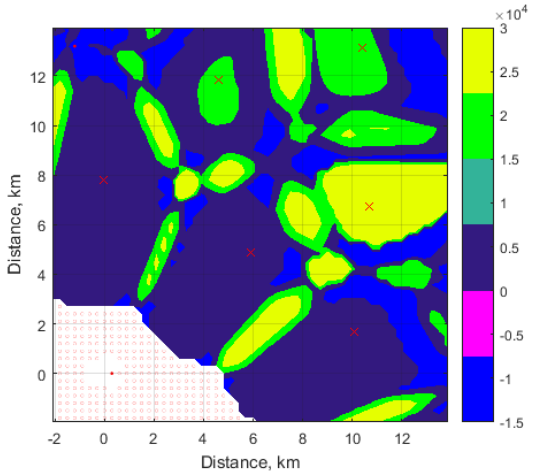

FIGURE 19. Contour plot focusing on overlapping areas (zoom in from the 30 km service area). The yellow areas indicate the areas with most improved users, dark blue areas indicate the areas with almost unaffected users, and light blue areas indicate the areas with highest loss users with 9 dB FBW CoMP (colorbar indicates capacity difference in bits per second).

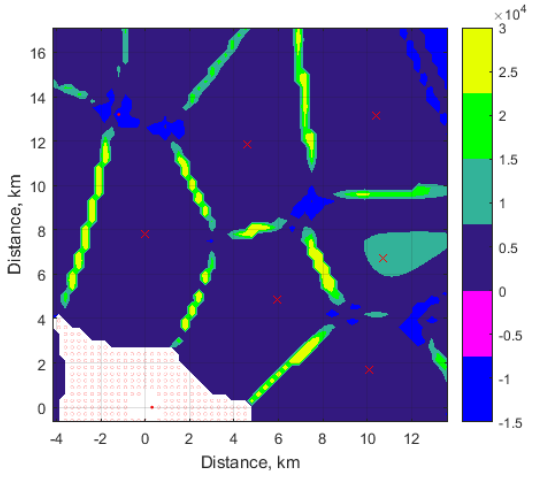

FIGURE 20. Contour plot focusing on overlapping areas (zoom in from the 30 km service area). The yellow areas indicate the areas with most improved users, dark blue areas indicate the areas with almost unaffected users, and light blue areas indicate the areas with highest loss users with Flex FBW CoMP (colorbar indicates capacity difference in bits per second).

Lastly in Figure 20, a contour plot of flexible FBW is presented. It shows how this flexible scheme helps reduce the users included into the CoMP region, restricting membership to those who can really benefit from CoMP.

V. CONCLUSION

JT-CoMP is shown to give significant benefits to the users at the cell edge in a HAP multi-beam system by improving both the CINR levels and capacity per user, whilst the same time improving the overall performance of the system. By identifying the trade-off between CINR and capacity, two types of threshold are proposed: the centralized CINR threshold, and the flexible CINR threshold. To deal with the unevenness of users in each cell, a flexible CINR threshold is implemented with each individual cell having a different threshold. Two different methods of allocating the bandwidth between the non-CoMP and CoMP regions have been proposed; the FBW and HBW schemes both bring benefits to 57% and 46% of users respectively. The FBW scheme works better for the CoMP region improving the user experience at the cell-edge. It is shown how γ can be used to control the overall user capacity and reaches an optimum. A flexible threshold is proposed in order to carefully select the users to be included into the CoMP region. With this approach the number of users that lose capacity can be minimized. Implementing CoMP is possible because of the newly proposed HAP architecture that enables the system to treat individual HAP beams as a serving cell which can be managed by virtual eNodeBs. This provides equivalence to the traditional cell approach used with the terrestrial systems, thus providing the capability to perform such functions in a flexible way.

REFERENCES

[1] D. Grace, M. Mohorcic, M. Oodo, M. Capstick, M. B. Pallavicini, and M. Lalovic, "CAPANINA-communications from aerial platform networks delivering broadband information for all," in *Proc. 14th IST Mobile Wireless Commun. Summit*, 2005, pp. 1–5.

[2] T. C. Tozer and D. Grace, "High-altitude platforms for wireless communications," *Electron. Commun. Eng. J.*, vol. 13, no. 3, pp. 127–137, 2001.

[3] J. Thornton, D. Grace, C. Spillard, T. Konefal, and T. Tozer, "Broadband communications from a high-altitude platform: The european HeliNet programme," *Electron. Commun. Eng. J.*, vol. 13, no. 3, pp. 138–144, 2001.

[4] S. Karapantazis and F.-N. Pavlidou, "The role of high altitude platforms in beyond 3G networks," *IEEE Wireless Commun.*, vol. 12, no. 6, pp. 33–41, Dec. 2005.

[5] K. Gomez *et al.*, "Aerial base stations with opportunistic links for next generation emergency communications," *IEEE Commun. Mag.*, vol. 54, no. 4, pp. 31–39, Apr. 2016.

[6] M. D. Zakaria, D. Grace, and P. D. Mitchell, "Antenna array beamforming strategies for high altitude platform and terrestrial coexistence using k-means clustering," in *Proc. IEEE 13th Malaysia Int. Conf. Commun. (MICC)*, Nov. 2017, pp. 259–264.

[7] *Facebook Cancels Its Internet Drone Program, Will Partner With Airbus Instead*. [Online]. Available: https://www.geekwire.com/2018/facebook-cancels-internet-drone-program-will-partner-airbus-instead/

[8] *Aquila's Successful Second Flight: Another Step Forward in Bringing the World Closer Together*. [Online]. Available: https://code.facebook.com/posts/200887800439084/aquila-s-successful-second-flight-another-step-forward-in-bringing-the-world-closer-together/

[9] *Airbus Zephyr Solar High Altitude Pseudo-Satellite Flies for Longer Than Any Other Aircraft During Its Successful Maiden Flight*. Accessed: Oct. 17, 2018. [Online]. Available: https://www.airbus.com/newsroom/press-releases/en/2018/08/Airbus-Zephyr-Solar-High-Altitude-Pseudo-Satellite-flies-for-longer-than-any-other-aircraft.html

[10] F. A. D'Oliveira, F. C. L. de Melo, and T. C. Devezas, "High-altitude platforms—Present situation and technology trends," *J. Aerosp. Technol. Manage.*, vol. 8, no. 3, pp. 249–262, 2016.

[11] *AeroVironment Announces Joint Venture and Solar High-Altitude Long-Endurance Unmanned Aircraft System Development Program*. [Online]. Available: https://www.businesswire.com/news/home/20180103005647/en/AeroVironment-Announces-Joint-Venture-Solar-High-Altitude-Long-Endurance

[12] *Coordinated Multi-point Operation for LTE Physical Layer Aspects*, document 3rd Generat. Partnership Project (3GPP), 36.819 R11 v11.2.0, ed, 2013.

[13] M. D. Zakaria, D. Grace, P. D. Mitchell, and T. M. Shami, "User-centric JT-CoMP for high altitude platforms," in *Proc. 26th Int. Conf. Softw., Telecommun. Comput. Netw. (SoftCOM)*, Sep. 2018, pp. 1–6.

[14] J. Thornton, D. A. J. Pearce, D. Grace, M. Oodo, K. Katzis, and T. C. Tozer, "Effect of antenna beam pattern and layout on cellular performance in high altitude platform communications," *Wireless Pers. Commun.*, vol. 35, nos. 1–2, pp. 35–51, 2005.

[15] Iskandar and A. Abubaker, "Co-channel interference mitigation technique for mobile WiMAX downlink system deployed via stratospheric platform," in *Proc. 8th Int. Conf. Telecommun. Syst. Services Appl. (TSSA)*, Oct. 2014, pp. 1–5.
[16] D. Grace and M. Mohorčič, *Broadband Communications via High Altitude Platforms*. Hoboken, NJ, USA: Wiley, 2011.
[17] Q. Zhang, Q. Xi, C. He, and L. Jiang, "User clustered opportunistic beamforming for stratospheric communications," *IEEE Commun. Lett.*, vol. 20, no. 9, pp. 1832–1835, Sep. 2016.
[18] I. Zakia, "Capacity of HAP-MIMO channels for high-speed train communications," in *Proc. 3rd Int. Conf. Wireless Telematics (ICWT)*, Jul. 2017, pp. 26–30.
[19] U. S. Hashmi, S. A. R. Zaidi, and A. Imran, "User-centric cloud RAN: An analytical framework for optimizing area spectral and energy efficiency," *IEEE Access*, vol. 6, pp. 19859–19875, 2018.
[20] G. Dandachi, S. E. Elayoubi, T. Chahed, and N. Chendeb, "Network-centric versus user-centric multihoming strategies in LTE/WiFi networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 5, pp. 4188–4199, May 2017.
[21] G. Nigam, P. Minero, and M. Haenggi, "Coordinated multipoint joint transmission in heterogeneous networks," *IEEE Trans. Commun.*, vol. 62, no. 11, pp. 4134–4146, Nov. 2014.
[22] V. Garcia, Y. Zhou, and J. Shi, "Coordinated multipoint transmission in dense cellular networks with user-centric adaptive clustering," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4297–4308, Aug. 2014.
[23] T. M. Shami, D. Grace, A. Burr, and M. D. Zakaria, "Radio resource management for user-centric JT-CoMP," in *Proc. 15th Int. Symp. Wireless Commun. Syst. (ISWCS)*, Lisbon, Portugal, Aug. 2018, pp. 1–5.
[24] H. S. Kang and D. K. Kim, "User-centric overlapped clustering based on anchor-based precoding in cellular networks," *IEEE Commun. Lett.*, vol. 20, no. 3, pp. 542–545, Mar. 2016.
[25] A. H. Sakr and E. Hossain, "Location-aware cross-tier coordinated multipoint transmission in two-tier cellular networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 11, pp. 6311–6325, Nov. 2014.
[26] W. Nie, F.-C. Zheng, X. Wang, W. Zhang, and S. Jin, "User-centric cross-tier base station clustering and cooperation in heterogeneous networks: Rate improvement and energy saving," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 5, pp. 1192–1206, May 2016.
[27] S. Bassoy, M. Jaber, M. A. Imran, and P. Xiao, "Load aware self-organising user-centric dynamic CoMP clustering for 5G networks," *IEEE Access*, vol. 4, pp. 2895–2906, 2016.
[28] L. Liu, Y. Zhou, V. Garcia, L. Tian, and J. Shi, "Load aware joint CoMP clustering and inter-cell resource scheduling in heterogeneous ultra dense cellular networks," *IEEE Trans. Veh. Technol*, vol. 67, no. 3, pp. 2741–2755, Mar. 2017.
[29] T. M. Shami, D. Grace, A. Burr, and M. D. Zakaria, "User-centric JT-CoMP clustering in a 5G cell-less architecture," in *Proc. 29th Annu. Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Bologna, Italy, Sep. 2018, pp. 177–181.
[30] A. Burr, A. Papadogiannis, and T. Jiang, "MIMO truncated Shannon bound for system level capacity evaluation of wireless networks," in *Proc. IEEE Wireless Commun. Netw. Conf. Workshops (WCNCW)*, Apr. 2012, pp. 268–272.
[31] X. Fan, S. Chen, and X. Zhang, "An inter-cell interference coordination technique based on users' ratio and multi-level frequency allocations," in *Proc. Int. Conf. Wireless Commun., Netw. Mobile Comput.*, Sep. 2007, pp. 799–802.

MUHAMMAD D. ZAKARIA received the B.Eng. degree in computer and communication systems engineering from Universiti Putra Malaysia, in 2014. He is currently pursuing the Ph.D. degree with the Communication Technologies Research Group, Department of Electronic Engineering, University of York. He was a Fellow with Universiti Sultan Zainal Abidin, under the Fellowship Scheme. His research focuses on the high altitude platform (HAP)-based communication and coexistence with terrestrial networks.

DAVID GRACE (S'95–A'99–M'00–SM'13) received the Ph.D. degree from the University of York, in 1999, with the subject of his dissertation on "Distributed Dynamic Channel Assignment for the Wireless Environment." In 2000, he jointly founded SkyLARC Technologies Ltd., and was one of its directors. He was one of the lead investigators of FP7 ABSOLUTE and focused on extending LTE-A for emergency/temporary events through the application of cognitive techniques. He was the Technical Lead of the 14-Partner FP6 CAPANINA Project that dealt with broadband communications from high altitude platforms. Since 1994, he has been a member of the Department of Electronic Engineering, University of York, where he is currently a Professor (Research) and the Head of the Communication Technologies Research Group. He is also a Co-Director of the York–Zhejiang Lab on Cognitive Radio and Green Communications, and a Guest Professor with Zhejiang University. He is the Lead Investigator of H2020 MCSA 5G-AURA and H2020 MCSA SPOTLIGHT. He has authored over 220 papers, and has authored/edited two books. His current research interests include aerial platform-based communications; cognitive green radio, particularly applying distributed artificial intelligence to resource and topology management to improve overall energy efficiency; 5G system architectures; dynamic spectrum access; and interference management. He was the Chair of the IEEE Technical Committee on Cognitive Networks for the period 2013/2014. He is a Founding Member of the IEEE Technical Committee on Green Communications and Computing.

PAUL DANIEL MITCHELL (M'00–SM'09) received the M.Eng. and Ph.D. degrees from the University of York, York, U.K., in 1999 and 2003, respectively. His Ph.D. research was on medium access control for satellite systems, which was supported by British Telecom. He has gained industrial experience with BT and QinetiQ. Since 2002, he has been a member of the Department of Electronic Engineering, University of York, where he is currently a Reader. He has authored more than 110 refereed journal and conference papers. His research interests include medium access control and routing, underwater acoustic networks, wireless sensor networks, cognitive radio, traffic modeling, queuing theory, and satellite and mobile communication systems. He is a member of the IET and a Fellow of the Higher Education Academy. He has served on numerous international conference program committees. He was the General Chair of the International Symposium on Wireless Communications Systems which was held in York, in 2010, the Track Chair of the IEEE VTC, in 2014, and the TPC Co-Chair of the ISWCS 2019. He is an Associate Editor of the *IET Wireless Sensor Systems Journal* and the *Sage International Journal of Distributed Sensor Networks*.

TAREQ M. AL-SHAMI received the B.Eng. degree (Hons.) and the M.Sc.Eng. degree in electronics, majoring in telecommunications, from Multimedia University, Cyberjaya, Malaysia, in 2012 and 2017, respectively. He is currently pursuing the Ph.D. degree with the Department of Electronic Engineering, University of York. Since 2016, he has been an Early Stage Researcher of the European-funded Marie Curie ITN-5GAuRA Project. His current research interests include 5G small-cell networks, radio resource management, and particle swarm optimization.

NILS MOROZS (S'13–M'17) received the M.Eng. and Ph.D. degrees in electronic engineering from the University of York, in 2012 and 2015, respectively. His Ph.D. research was part of the EU FP7 ABSOLUTE Project, where he developed LTE-compliant dynamic spectrum access methods for disaster relief and temporary event networks. He was a Researcher in Wi-Fi and wireless convergence with BT, Martlesham, U.K., and a Research Associate in the area of high altitude platform (HAP)-based communications with the Department of Electronic Engineering (EE), University of York. He is currently a Research Associate with the Department of EE, University of York, focusing on channel modeling and medium access control for underwater acoustic sensor networks as part of the EPSRC USMART Project (EP/P017975/1). His research interests include the development of protocols, and architectures for wireless radio and acoustic networks.