Coverage Extension and Capacity Enhancement Using High Altitude Platforms

Steve Chukwuebuka Arum

Ph.D.

University of York

Electronic Engineering

January 2022

Abstract

This thesis develops and investigates approaches to extending coverage and enhancing capacity of wireless communications using a high altitude platform (HAP) with a multi-element planar phased array antenna (PAA). The purpose of the research is to demonstrate how a HAP coverage area can be significantly extended beyond the 30 km radius area state-of-the-art and user capacity enhanced to further maximise the utility of HAPs. The feasibility of using solar-powered HAPs to provide such extended coverage over a long duration, given the payload and energy constraints, is evaluated based on some example platforms at different locations. It is shown, using some derived energy models, that HAP operation with adequate energy management is feasible at equatorial regions when using an aircraft with wingspan greater than 35 m. However, airships, or aircraft with larger wingspan and alternative energy sources are required at locations further away from the equator. With feasibility established, it is shown that contiguous coverage can be achieved over a service area of at least 60 km radius using a proposed beam-pointing algorithm, which is based on PAA beamforming. The algorithm, which directly considers beam broadening and overlap at lower elevation angles, significantly extends the coverage area and achieves an improvement of between 5–15 dB carrier-to-interference-plus-noise ratio (CINR) compared with alternative schemes including the state-of-the-art. Relevant models for estimating the theoretical limits of the achievable coverage extension are also derived. Furthermore, a two-tier architecture, which forms two tiers of contiguous tessellated cells over an extended area with users having the flexibility of connecting to the best tier based on CINR, is proposed. The architecture, formed using multi-element planar PAA, considerably improves coverage and user throughput, and reduces edge-of-cell connectivity by up to 25% compared with the typical one-tier architecture. Insights on the possibility of using tier-based HAP architecture to facilitate 6G communication, which includes possible deployment scenarios, architectures, and use cases, are presented. A modified two-tier architecture involving the creation of an additional ad-hoc tier that is optimised for hotspot coverage and changing network user and traffic demands is also developed. The tier-based architectures, which include the adaptable and reconfigurable three-tier architecture, achieve over 99% coverage and can dynamically optimise the beam-pointing for enhanced capacity, which are useful for 6G.

Contents

List of Figures

List of Tables

Acknowledgements

First and foremost, I would like to thank the Almighty God for His sufficient grace on me. I am eternally grateful for His guidance and protection throughout this study.

My deepest gratitude goes to my supervisors Prof. David Grace and Prof. Paul Mitchell for their dedication and insightful guidance, without which this work would not have come this far. Thanks to my thesis advisor, Prof. Alister Burr, for the valuable discussions and feedback on my work throughout the Ph.D. Thanks also to my external examiner, Dr. Oluwakayode Onireti for the in-depth discussions during the viva and detailed feedback. My appreciation also goes to Dr. Nils Morozs and Dr. Mario Maurelli for contributing to this work.

My eternal gratitude goes to my family. Especially to my beloved wife, Chioma, who has been incredibly supportive and making countless sacrifices to enable me get to this point. And to my son, Alinachukwu, for the extra motivation to finish this work with expediency. Also, to my parents, HRH. Okeke and Lolo Amoge, and siblings, Arinze, Adaeze, Chiemerie, and Chinagorom, for their continued support and encouragement. Daalụ nụ o ndị nke m. Thanks also to my 'adopted' family in York - Phil and Gill Wilson, Elisa, Isaac, and Martha who have made my family a part of theirs and taken such great care of us.

Thanks to my colleagues at the Department of Electronic Engineering for creating a friendly, supportive and intellectually stimulating research environment. In particular, thanks to Kayode, Qiao (Joe), and Muheeb for the countless activities and discussions that made this work and life in York enjoyable. Thank you to many others too numerous to mention, who in one way or the other also helped me in the course of this study.

Finally, I would like to acknowledge that this work has been partially funded by Orange under research agreement H09121.

Declaration

All work presented in this thesis is original to the best knowledge of the author. References to other researchers have been given as appropriate. This work has not previously been presented for an award at this or any other institution. The research presented in this thesis features in a number of the author's publications listed below.

Journal Articles

S. C. Arum, D. Grace, and P. D. Mitchell, "A review of wireless communication using high-altitude platforms for extended coverage and capacity," *in Elsevier Computer Communications*, vol. 157, pp. 232-256, 2020.

S. C. Arum, D. Grace, P. D. Mitchell, M. D. Zakaria, N. Morozs, "Energy management of solar-powered aircraft-based high-altitude platform for wireless communications," *in MDPI Electronics*, vol. 9, pp. 179 – 204, 2020.

S. C. Arum, D. Grace, and P. D. Mitchell, "Cellular service provisioning from high altitude platforms using a two-tier Architecture," *submitted to IEEE Transactions on Mobile Computing*, under revision, 2021.

S. C. Arum, D. Grace, P. D. Mitchell, and M. D. Zakaria, "Delivering extended cellular coverage and capacity using high-altitude platforms," *to be submitted to MDPI Electronics*, in preparation, 2022.

Conference Paper

S. C. Arum, D. Grace, P. D. Mitchell, and M. D. Zakaria, "Beam-pointing algorithm for contiguous high-altitude platform cell formation for extended coverage," *in Proc. IEEE 90th Vehicular Technology Conference (VTC2019-Fall)*, Sept. 2019, pp. 1-5.

Patent

S. C. Arum, D. Grace, P. D. Mitchell, and L. Reynaud, "Method for allocating a frequency resource to at least one terminal and associated device," France Patent Application 21 06709, 23 Jun. 2021.

Chapter 1. Introduction

Contents

1.1. Overview

For many years, stakeholders have progressively looked at the possibility of using high altitude platforms (HAPs) to provide cost-effective communications, but mostly within a relatively small coverage area of 30 km radius. Given a HAP's favourable propagation environment and its elevated look angle, a coverage area of 30 km radius is thought to be underwhelming. Increasing this coverage area can further maximise the cost-effectiveness and achievable utility. One such utility is the capability of HAPs to communicate directly with user equipment (UE) in the same way as a terrestrial system. Considering that technology has matured so much that HAPs are now beginning to come to fruition, it is worth looking at how the limited coverage area of a HAP can be increased to enhance cost-effectiveness and other benefits, which is desirable especially in rural or remote areas with minimal or no coverage from terrestrial or satellite systems. Interestingly, this fits with the vision of future communication networks providing cost-effective ubiquitous coverage and capacity enhancement.

Providing extended coverage from HAPs in rural areas will deliver benefits such as increased connectivity, improved education and health service delivery, rural-urban migration mitigation, which will generally improve the socio-economic status of the inhabitants [13] in line with the United Nation's sustainable development goals. HAPs present a compelling use case because of the characteristically low user densities in rural areas and prohibitive cost of the alternative terrestrial or satellite communication systems, which can strain economic delivery of services and render their deployment commercially unattractive to the operators. Furthermore, the design and deployment of wireless communication systems in rural areas must address the need for wide area coverage, operation in varying geographic terrain, robustness, and ease of installation [13]. In order to satisfy the large coverage requirement using terrestrial systems, extremely tall base station (BS) masts are needed with signals transmitted at significantly high power [3], which pose some regulatory challenges among other issues. It is also important

to consider that some rural areas are very remote from the nearest backhaul infrastructure or do not have grid-based power for the terrestrial systems or both. Therefore, HAPs are desirable for providing increased coverage over a wide area and enhanced capacity at a reduced cost with ease of deployment, and the possibility of incremental deployment [8, 14, 15], which encapsulates the requirements of a rural wireless communication system [13].

However, notwithstanding the benefits and desirability of HAPs, there is still the question of the feasibility of solar-powered HAPs, which is the most popular and technologically matured HAP currently, for long duration wireless communication missions especially at latitudes further away from the equator given the state-of-the-art. Central to the feasibility is energy management, which is under-explored despite its importance. With the deployment of HAPs for wireless coverage close to being a reality, it is fundamental that energy management of a HAP communication system is given careful consideration to ensure operability. This governs how much solar energy is needed to maintain flight and support the payload during the day, and how much energy needs to be harvested and stored for night time usage for continuous operation throughout the desired mission duration. With only a few existing studies [16, 17] considering solar cell energy harvesting in relation to solar-powered high altitude aircraft, more work is clearly needed in this area. This involves developing realistic models for energy harvesting as a function of latitude and platform size, and consumption by the platform and payload.

Consequently, this thesis investigates the provisioning of wireless communications coverage and capacity especially in rural and remote areas using a HAP, with specific focus on extending the coverage and enhancing capacity achievable from a HAP. It also explores the feasibility of such missions within the context of energy management, by looking at bringing together the necessary calculations needed to evaluate the feasibility and operability of a solar-powered HAP for providing continuous wireless coverage over a given area at any given latitude. Until now, the majority of HAP-related research and projects mostly consider a service area of around 30 km radius. However, since some terrestrial macro cells can provide coverage within an area of up to 30 km radius [18], and considering the better propagation environment, a HAP should be able to provide coverage significantly beyond the common 30 km radius and enhance the achievable capacity, subject to the limitations of the aerial platform. It is thought that the benefits of HAPs already highlighted earlier have the potential to be extended more widely through an increased service area, and this thesis will focus on this aspect. Although the thesis highlights rural areas, which is mainly because they have a ready-made market for HAPs, the work presented can be directly applied to other environments with minimal modifications.

1.2. Hypothesis

The following hypothesis guides the work presented in this thesis.

"*Significant contiguous and non-contiguous wireless communications coverage and capacity can be delivered over an extended service area at low elevation angle using high altitude platforms with appropriate communication strategies.*"

The capability of a HAP to deliver wireless communications at low elevation angles is assessed by understanding the feasibility of using the current state-of-the-art given the characteristics, latitude of operation, radius of coverage of a platform, as well as developing and evaluating strategies that mitigate the effect of beamforming at low elevation angles. The feasibility of using the current platforms is established by using models of energy collection as a function of operating latitude and consumption developed and evaluated for the common and extended radius of coverage scenarios. It is essential for the platforms to have sufficient payload carrying capability and harvest sufficient energy to power both the avionics and wireless communications payload. Additionally, it is also critical to overcome the effects of propagation and beamforming at low elevation angles such as beam broadening resulting in severe beam overlap. Furthermore, strategies to effectively mitigate interference and fairly distribute the system resources should be devised considering the disproportionality of beam and cell sizes in the network are necessary. These aspects are the focus of this thesis and the solutions proposed are evaluated with simulations and theoretical analysis.

Therefore, to prove the above hypothesis, the following objectives provide the needed framework. They also briefly introduce the focus of the technical work chapters in this thesis.

1. Develop energy management models to estimate the HAP energy collection and consumption while considering size, weight, and power (SWaP) requirements of both the platform and wireless communication payload. The models will be used to determine the feasibility of current state-of-the-art HAPs for long duration wireless communication missions at any given location.

2. Develop beam pointing and cell formation schemes that can overcome the limitations of forming beams at low elevation angles, which results in significant beam broadening and overlap. The schemes will enable the deployment of adequate contiguous cellular network over an extended service area.

3. Develop robust, adaptable, and reconfigurable wireless communication architectures and

techniques that can mitigate interference, guarantee a minimum level of quality of service (QoS), and ensure fairness to users.

4. Evaluate the flexibility of the developed architectures and techniques under varying usage conditions.

1.3. Thesis Outline

The rest of the thesis is organised as follows.

- Chapter 2 reviews the existing literature on using HAPs for wireless communications. It presents an overview of HAPs including their classifications, regulations, and some relevant projects. It then discusses the background of wireless communication, introducing the network topologies, cellular architectures, antenna beam-pointing and cell formation strategies. Furthermore, detailed discussions on providing extended coverage using HAPs including factors affecting coverage extension are presented. In addition, some challenges facing the use of HAPs for wireless communications are highlighted as well as discussions on the HAPs feasibility considering the fundamental energy and payload constraints. This work has resulted in a review article publication in Elsevier Computer Communications Journal.

- Chapter 3 discusses the energy management modelling and feasibility of using a HAP for wireless communications over a long period in the context of energy availability and payload capabilities of current platforms. It presents models of solar energy collection and consumption by a HAP, which are used to evaluate long-endurance HAP feasibility at different locations. Three representative example platforms are assessed using the proposed models. In the context of coverage area, the last section of the chapter analyses the feasibility of HAPs for wireless communication service delivery both over the common 30 km and an extended 60 km radius coverage area proposed in this thesis. This work has resulted in a publication in MDPI Electronics Journal.

- Chapter 4 proposes a novel algorithm that points HAP beams over an extended coverage area. The proposed algorithm directly considers the broadening of beams, especially when pointed at the extended locations, to minimise overlap and significantly mitigate interference. The analytical framework of the proposed algorithm is presented. Some theoretical models for estimating the area of a HAP beam footprint, achievable average capacity and area spectrum efficiency which can be used to estimate the possible

limits of coverage extension, are derived. Combining the models, one can estimate the possible HAP extended coverage bounds given the operational parameters. The chapter also presents system performance results of the proposed algorithm evaluated using simulations as well as numerical results from the derived models. Part of this work was presented in an IEEE Vehicular Technology Conference in 2019. An extended version of the conference paper consisting of the work in this chapter has been submitted to IEEE Transactions on Aerospace and Electronic Systems.

- Chapter 5 proposes a novel tier-based HAP cellular network architecture that significantly mitigates the effects of inter-cell interference (ICI) and disproportionately sized cells that characterise the extended HAP coverage. The architecture comprising two homogeneous tiers of the contiguous cellular structure obtained in Chapter 4, with one tier rotated in angle, significantly improves the downlink signal quality. Furthermore, two relevant algorithms that deploy the two tiers and perform user association are proposed in this chapter. The performance of the proposed architecture is analysed and compared with the benchmark one-tier architecture using simulations. This work is currently under revision for resubmission in IEEE Transactions on Mobile Computing.

- Chapter 6 presents a futuristic view of the use of HAPs in future networks such as 6G. It discusses the different scenarios for delivering 6G communications from a HAP within a rural area context by highlighting the possible use cases and architectures. The chapter presents insights on how tier-based architectures can enable 6G from HAPs and proposes a modification of the two-tier cellular architecture proposed in Chapter 5 by introducing an additional ad-hoc tier to form a three-tier HAP cellular network. The ad-hoc tier can provide service to hotspot users within the HAP service area, taking the size and distributions of the hotspots into consideration. Additionally, the algorithms for the proposed architecture and beam pointing optimisation for improved coverage and capacity suitable for user hotspots are presented. As in the previous chapters, the proposed architecture is assessed using simulations but with non-uniform user densities unlike in the previous chapters. This work will later form a publication to be submitted to IEEE Transactions on Mobile Computing.

- Chapter 7 presents the conclusions of this thesis, summarises its original contributions, and discusses a number of recommendations for further work.

Chapter 2. Literature Review

Contents

2.1. Introduction

The development of aerial platforms for different applications dates back in history but was stalled by the Hindenburg disaster [3]. However, there has been an increasing interest and significant improvements in aerial platforms and their applications in the past 20–30 years, with the time-to-market of the technology for various applications now projected to be within a few years. This is due to the rapid development in the key enabling fields such as advanced material science, which allows for the manufacturing of lighter but durable materials. Additionally, solar cells have evolved so much that there are currently ultra-thin, flexible and lightweight solar panels with improved efficiency [17]. This is driving the development of solar-powered

platforms where weight, flexibility, and efficiency is an issue. Advancement in energy storage systems and advanced materials for the realisation of aerial platforms are some other factors fast-tracking developments. This chapter focuses mainly on the application of HAPs, a special type of aerial platform, for wireless communications service delivery, providing a brief background on the platform and communication concepts with some challenges of using HAPs to deliver wireless communications.

HAPs are aerial platforms, which are either manned or unmanned, located conventionally between 17–22 km altitude and used for a variety of applications including wireless communications [8]. The advantages of HAPs over terrestrial wireless and satellite systems in addressing some wireless communication issues are clearly highlighted in the literature [19, 20]. HAPs offer some advantages over terrestrial systems due to their elevated look-angle resulting in better coverage and propagation performance. Furthermore, compared with both terrestrial and satellite systems, HAPs are cheaper and quicker to develop, deploy, operate, and maintain. The increasing optimism in HAPs is partly due to the possibility of using one platform for multiple applications and also due to their potential for low cost, high availability wireless communications service provision over an extended area. There are various previous and ongoing HAP projects both on platform development and their use for wireless communications.

2.2. HAP Overview

Although this chapter aims to introduce the background and concepts of HAPs for wireless communications, this section presents some relevant information facilitating that. It covers areas such as the different classifications of HAPs that can be deployed and their distinguishing features, relevant aviation and radio regulations, as well as some past and current HAP projects. Note that other names that have been used for HAPs include high altitude platform station (HAPS) [21], high altitude pseudo satellites [14], stratospheric platforms [22, 23], etc.

2.2.1. Classification

Generally, a HAP can be manned or unmanned [8, 24]. A typical manned HAP is the Proteus [25] specifically designed for long endurance operation at high altitudes. Since HAP missions, especially for wireless communications, are of very-long duration at altitudes not suitable for human pilots due the harsh stratospheric environment, unmanned platforms enjoy more attention from HAP stakeholders [8]. Hence, this thesis focuses on unmanned HAPs.

HAPs can be classified, based on the physical principle providing the lifting force, as aerostatic

or aerodynamic. They can also be classified based on the platform design. Commonly, HAPs are classified as balloons, airships, or airplanes, and their main characteristics are shown in Table 2.1. The HAP class suitable for any given communication mission depends on a number of factors such as the payload weight, the energy available to the platform, and mission duration. Additionally, an important factor in determining the appropriate type of HAP is the regulations guiding the deployment and operation, which are often country specific.

HAP Category	Characteristics
Aircraft	• Can potentially support up to 250 kg payload • Can be manned (e.g Global Hawk) or unmanned (e.g Zephyr-S) • Up to 20 kW of power available for payload • Solar-powered operation capability with solar cells on the wings
Airships	• Significant payload support up to 1000 kg • Fly up to 30 km altitude • Active station keeping capability • Between 5–20 kW of power available to payload • Solar-powered operation capability with large surface area for solar cells
Balloons	• Can support less than 100 kg • Typical operating altitude 17 km • No active station keeping • Less than 1 kW of power available to payload • Solar-powered operation capability with limited space for solar cells

Table 2.1: Classification and features of High Altitude Platforms [7–10].

2.2.2. Regulations

In order to facilitate the integration of HAPS into the wider global communication system, internationally accepted regulations are needed [26]. There are different types of regulations that will govern the use of HAPs for wireless communications. On the one hand, there are aviation regulations, which govern the licensing and operation of all the platforms typically within the jurisdiction of international and national civil aviation authorities. Unfortunately, their is no internationally recognised pathway to achieve this yet. On the other hand, there are radio regulations governing the spectrum usage strictly for HAPS operating at altitudes between 20–50 km as regulated by International Telecommunication Union - Radiocommunication sector (ITU-R). Conventional, HAPs operating at altitudes between 17–22 km follow ITU-R regulations.

Aviation Regulation

The deployment of any given class of HAP is subject to a set of aviation regulations, which differ in different countries. These regulations more often than not specify how the platforms are allowed to operate and the maximum allowable altitude for each class. Generally, it is typical for unmanned aircraft operators, irrespective of size, to obtain a license from the civil aviation authority of the country where they intend to operate. The major aim of the civil aviation authority is to enforce safety, security and privacy rules [27–29]. An internationally consistent collaborative traffic management concept and better globally shared meteorological data for near-real-t information on aircraft conditions during ascent/descent are needed [26].

However, despite the increased discussions worldwide, there is yet no legal specification particularly regulating the operation of HAPs anywhere currently [30], which is hampering the development and deployment of the system. Liu and Tronchetti [30] propose a new categorization of near space, spanning from 18–100 km, as the exclusive utilisation space. The model suggests differentiating the legal status of the near space group from that of national airspace and outer space. The authors suggest abrogating near space, which will be regulated by a set of rules targeted at profit maximization and sustainability, from the national sovereignty of the underlying states. However, it proposes to maintain the sovereign rights of the underlying states by allocating them the right to decide the conditions for use, negotiate the conditions to be complied with by foreign bodies, and regulate/enforce safety and security issues as concerns the states.

Radio Regulation

Irrespective of platform characteristics, broadband communications from a HAP is based on a set of spectrum globally regulated by the ITU-R [15]. Over the past 20 years, different spectrum bands both in the millimetre wave (mm-Wave) and international mobile telecommunications (IMT) bands such as 2.1 GHz, 6.5/6.6 GHz, 47/48 GHz, etc. have been assigned for HAP use in different regions by the ITU-R. At the 2019 world radio congress (WRC-19), the ITU-R released more regulations on the requirements for the maximum transmit equivalent isotropically radiated power, antenna beam pattern, power flux density level, separation distance between radio astronomy station and the nadir of a HAP for operation in HAP spectrum bands [31]. The allocated HAP spectrum bands in different regions are highlighted in [15, 28].

Integrating HAPs into the global communication systems require appropriate radio regulations. Unfortunately, there are some regulatory challenges facing the use of HAPs for wireless com-

munications such as the inability to allocate different spectrum especially in tropical areas, which is required due to the significant attenuation experienced in the allocated 47/48-GHz and 38–39.5-GHz bands. Unfortunately, the 28/31–GHz and 21.4–27.5 bands with significantly lower rain attenuation is not available in some regions. Another main challenge is the varying aeronautic regulations in different countries. Studies in [29] look at some regulatory challenges affecting the use of HAPs for wireless broadband service provisioning such as the geographical limitation of some spectrum bands identified for HAPs use by ITU-R.

2.2.3. Projects

A number of projects, some of which are specified below, are based on supporting wireless communications applications from HAPs. While many of the projects focus on wireless communication applications, some focus on the development of platforms for varying applications such as Airbus' Zephyr-S/T [32], BAE Systems' PHASA-35 [33], Boeing's Phantom Eye [34], Stratobus Airship [35], and Sceye [23] projects. Some of the popular wireless communications HAP-related projects, which includes two current and two past projects are highlighted below.

- **SPL HAP**: Stratospheric Platforms Limited is developing a hydrogen-powered HAP with a 60 m wingspan to deliver reliable 5G communications. The platform, which can support a 140 kg payload, is designed to be used for filling mobile coverage gaps and providing rural home broadband services. The system is projected to be ready for commercial release in 2024 [7].

- **HAPSMobile**: this joint venture announced in early 2018 [36] aims at providing wireless communications services from HAP for commercial purposes. It includes the development of a solar-powered platform called Hawk30 [37], which is specially designed for stratospheric communication systems with a 78 m wingspan.

- **Loon**: this project investigated using a constellation of free-flying stratospheric balloons that can support a payload weight of about 10 kg for over 100 days to deliver internet services within a given location and powered solely by solar panels [8, 38, 39]. Unfortunately, this project was discontinued in January 2021.

- **Aquila**: similar to Loon in terms of goals, Aquila intended to deliver high speed broadband services [8,14]. It planned to develop and use a solar-powered platform with a 42 m wingspan. This project was also discontinued but with plans to continue partnering with Airbus on enabling HAP connectivity and technologies.

2.3. HAP Wireless Communications Background

The idea of using HAPs for wireless communications, and some other applications, started becoming mainstream when Djuknic et al. [19] alluded to this possibility. Currently, HAPs are being considered for wireless communication delivery over a wide area for a long duration. However, there are a number of technical challenges, based on platform capability and wireless communication system limitations, to overcome. Over the years, there have been consistent improvements in the research and development of appropriate wireless communications design and techniques for HAPs. There are existing studies in the HAP-related literature on network architecture and topology design, antenna beam placement and cell formation [6, 40], adaptation of conventional communication techniques such as coordinated multipoint (CoMP) and inter-cell interference coordination (ICIC) for HAPs [41], the use of free space optical(FSO) communication for inter-platform links connecting these systems, etc. With the advent of 5G and the futuristic 6G, which are popularising advanced techniques such as software-defined networking (SDN) [42], network function virtualisation (NFV) [43], network slicing [44], reconfigurable smart surfaces [45], some of the issues currently limiting HAPs deployment such as miniaturisation of payloads for reduced weight, backhaul technology, payload power consumption can be mitigated. Some of these underlying approaches and strategies that facilitate HAPs for wireless communications are discussed in this section.

2.3.1. Network Topology

The ideal HAP network topology aims to achieve high connectivity, speed, and reliability as well as low latency and power consumption. Whether this is feasible depends on the HAP application, deployment scenario, and coverage area, which determines the features and capacity of components required on the platform and possibly on the ground. The simplest HAP topology involves a transparent transponder in a standalone HAP with a single beam footprint [46].

Different HAP topologies, commonly categorised as hybrid or non-hybrid, have been extensively studied [15,47–49]. In hybrid topologies, HAPs form part of a HAP-Terrestrial or HAP-Terrestrial-Satellite system [15, 48, 49] as shown in Figure 2.1. The HAP-Terrestrial topology involves terrestrial base stations (BS) serving higher user density areas while the HAPs serve the lower user density areas and temporary hotspots [48], providing both access and backhaul links to the core network. Other emerging use cases include HAPs supporting and managing internet of things (IoT) services, other aerial networks, and intelligent transport systems as well as serving as aerial data centres. Extended versions of the topologies in Figure 2.1 can include

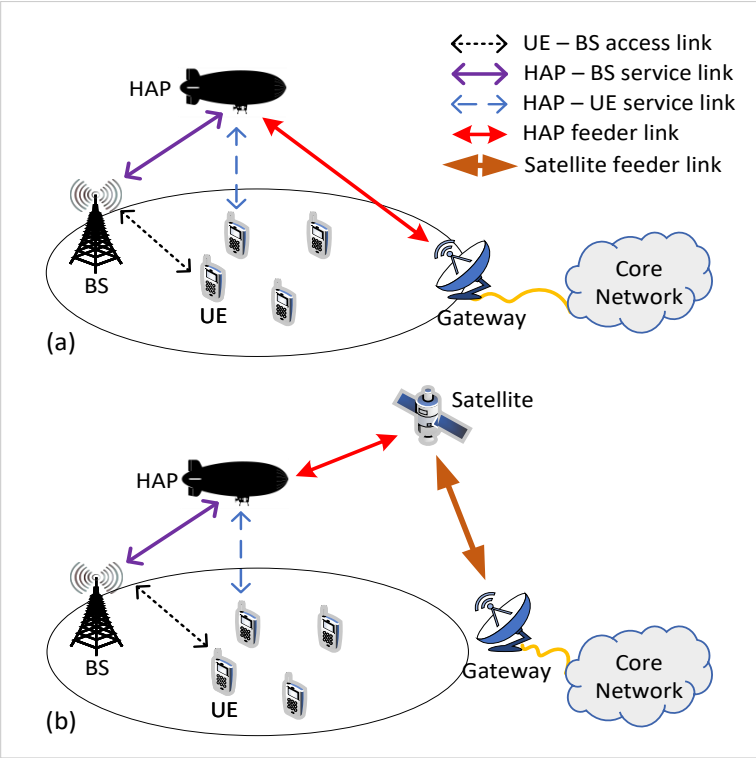

Figure 2.1: Hybrid topologies for (a) HAP-Base station and (b) HAP-Base station-Satellite.

low altitude platforms, which can be unmanned aerial vehicles (UAVs) or tethered platforms, as part of a larger vertical heterogeneous network [21]. The terrestrial system can potentially be backhauled to the HAP, which can be beneficial in areas without core network infrastructure nearby. The HAP-Terrestrial-Satellite topology can be used to increase fault tolerance, ensure redundancy, and prevent system failures [48] with the possibility of backhauling from HAP to satellite. A HAP-Satellite topology is discussed in [15] although the authors acknowledge that the significant distance between the HAP and satellite can be problematic especially given the low latency requirements of next generation (NextGen) communication systems.

For the non-hybrid topologies shown in Figure 2.2, the HAP can either be configured as a single BS or as a part of a mesh network of HAPs with capability for extended coverage and enhanced capacity [48, 49]. The standalone-HAP in Figure 2.2a requires both radio and baseband units to be located in the HAP with a backhaul link to the core network. On the other hand, the multi-HAP topology in Figure 2.2b can be used to extend the wireless network [15]. Studies in [50, 51] investigate the use of multi-HAP topology to improve the system performance. The system capacity of broadband services from HAPs can be improved through diversity by using multiple platforms sharing the same spectrum and exploiting user antenna directivity as shown in [50]. The diversity performance of multiple HAPs using virtual multiple input multiple output (MIMO) with different modulation schemes is studied in [51] with up to 10 dB

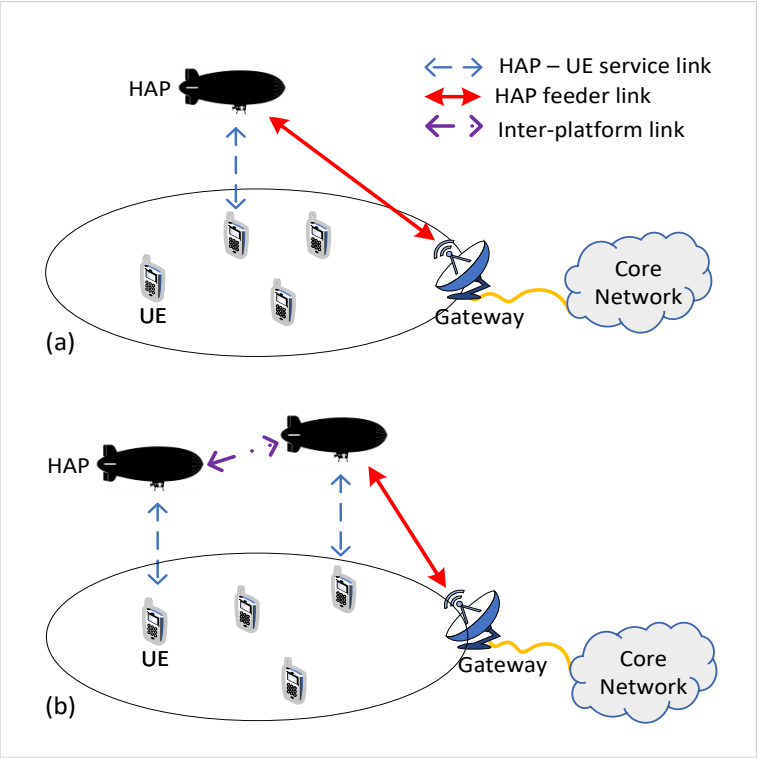

Figure 2.2: Non-hybrid Topologies with (a) Standalone-HAP and (b) Multi-HAP topologies.

diversity gain shown to be achievable using a 4x4 MIMO antennas.

2.3.2. Antenna Beam-pointing and Cell Formation

Similar to terrestrial systems, delivering wireless communication services from HAPs require the formation of cells on the ground as depicted in Figure 2.3. This typically involves the simul-

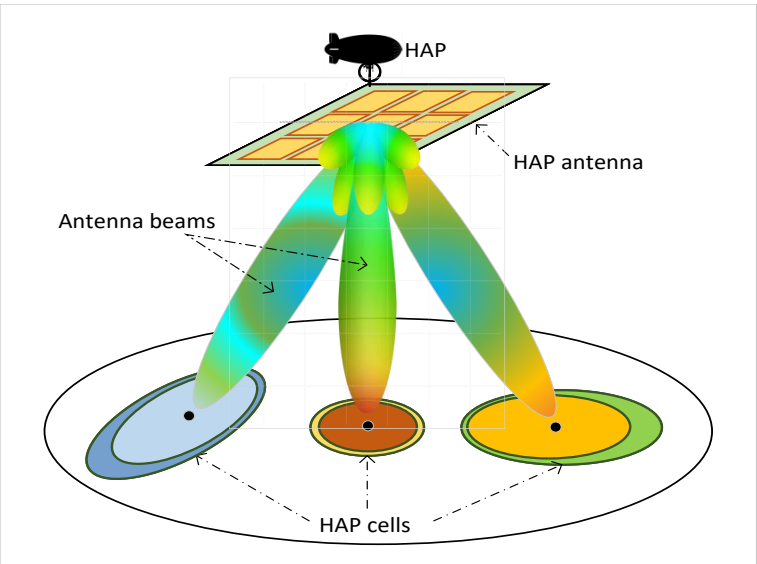

Figure 2.3: HAP multi-beam formation for cellular communication.

taneous formation and deployment of multiple antenna beams, which are used to form cells. However, there are some challenges in forming these cells, which can either be contiguous over the service area or provide islands of coverage based on the spatio-temporal distribution of users or traffic, especially for extended coverage. Firstly, HAP cells pointed at significant distances from the sub-platform point (SPP) broaden in shape and size due to the limitations of antenna beamforming [52]. The lack of consideration of broadening leads to a severe overlap between neighbouring cells, which in turn results in significant ICI. Secondly, the antenna radiation pattern producing the beams, which form cells, directly influences ICI due to the behaviour of the sidelobes. Therefore, in order to ensure an efficient HAP cellular communication system especially within the context of an extended service area, cell formation must be carefully considered, also taking the effects of beam broadening and antenna radiation pattern into account. The majority of the existing studies on HAP cell formation and deployment such as [6, 19, 53–56] consider a limited service area with radius of around 30 km, which this thesis aims to extend. These prior studies are further discussed in the following paragraph.

In [19], Djuknic *et al.* propose the use of scanning beams that scan across an arrangement of cells randomly to activate them. The proposed scheme requires traffic buffering, which is inefficient, increases the system complexity, requires modification of the air interface, and significantly increases the scanning time especially for wide area networks, making multiple simultaneous service beams necessary. Studies in [53] propose the use of a uniform hexagonal cell grid architecture over an area of 30 km radius using antenna beams illuminating the hexagonal cells, with the authors limiting the HAP's elevation angle to a minimum of around 30°. This ensures the formation of approximately circular cells on the ground and regular tessellation, which are challenging at lower elevation angles. In [54, 56], HAP cell footprints are described mathematically as functions of antenna beamwidth, elevation and azimuth angles, but these functions are only valid within a limited area of 30 km radius due to increased inaccuracies and approximation errors at extended distances. Zakaria *et al.* [6,57] studied the use of intelligent beamforming strategies for HAP coverage and capacity while providing protection to terrestrial system users. They proposed a beam-pointing scheme where users are clustered using k-means and cells pointed at the centroid of the clusters with the number of clusters determined a priori. The k-means based scheme achieves a 95% coverage, outperforming the regular and random schemes. However, they only considered a 30 km radius coverage area, which is common for most of the studies in the literature. Since, some terrestrial macrocells can provide coverage within an area of up to 30 km radius [18], and the higher elevation angles to HAPs result in better propagation, it is reasonable to study HAP beam-pointing performance

beyond a 30 km radius as discussed in Section 2.4, which is important for cost-effective communications in low user density areas.

For non-contiguous beam-pointing particularly in spatio-temporal user or traffic and hotspot scenarios, user localisation is important to effectively deploy cells for improved system performance. Accurately localising users becomes more important in NextGen systems such as 6G involving beam tracking of high mobility users. In [58], a two-stage two-dimensional direction-of-arrival algorithm with up to 40 m resolution is proposed. Then, to deploy targeted beams, the use of user clustered opportunistic beamforming is proposed [22] with beams targeted at clusters of users that are formed using only received signal information from users. Since beams are deployed based on clusters, it is likely that some users in a cluster are not properly served. In this case, beam-pointing can be optimised by using techniques such as gradient descent, genetic algorithm, simulated annealing (SA), etc. Furthermore, cells may have to be shaped or deployed smartly to account for arbitrary cluster shapes.

2.3.3. Cellular Architecture

The HAP cellular architecture has evolved over the past two decades. Earlier studies [19,59,60] proposed a number of architectures, which include ring-shaped cell clustering and cell scanning. As highlighted in Section 2.3.2, both solutions are rigid considering the spatio-temporal variation of users and traffic characteristics and ad-hoc deployments expected of NextGen systems. A different architecture based on sectorisation of concentric rings of cells with each sector, which is analogous to a typical cell, illuminated by a separate orthogonal antenna beam is proposed in [59, 60]. This is similar to the proposed architecture in the high altitude long operations (HALO) network [59] made up of 6 rings of cells from 125 antenna beams. It involves creating fixed cells with no overlapping areas between adjacent cells, which is difficult to achieve without introducing significant coverage holes due to beamforming limitations. Falletti *et al.* [60] studied the performance of the sectored cell architecture using 3 rings sectored cells. Assuming that all sectors have equal area, they show that downlink capacity increases while traffic channel density reduces with increasing number of rings and sectors. An interesting architecture, which involves creating both macrocells and microcells, was proposed in [54]. The macrocells are constantly maintained while the microcells formed using a different antenna will shift in position by following the traffic in the network. This complicates the antenna system and introduces additional weight for the platform to deal with.

The performance of the ring-shaped cell clustering solution proposed in [19] is investigated in [61] with the author proposing an improved vertical antenna array beam shape of the ring-cells,

which results in a significantly improved carrier-to-interference ratio. However, the study also did not address the concerns about the effect of beam broadening and user or traffic variation on the system performance. Some more flexible solutions considering irregular cell shapes, user clustering, and antenna beam broadening are proposed in [6, 40]. In [40], the HAP coverage area is divided into pixel spots, grouping these pixels accordingly to obtain the desired cell shape. An advantage of the proposed architecture is that cells can be formed to follow the arbitrary shape of any cluster of users using an appropriate antenna. Alternatively, studies in [6], propose an ad-hoc architecture where beams are deployed to serve clustered users with beams pointed at the cluster centroids. While cells can be pointed to the clusters to better serve the clusters, depending on the size of a cluster, more than one cell may be needed to appropriately serve the clusters. Unfortunately, this was not considered by authors.

2.4. HAP Extended Cellular Coverage

Most of the studies [6, 41, 54, 55, 62] on wireless communications coverage architecture using a HAP are limited to within 30 km radius coverage area mainly to maintain a minimum elevation angle of around 30°, and have relatively equal sized cellular footprint. However, this is pessimistic given a HAP's significantly better propagation environment compared with both satellite and terrestrial systems. Expectedly, a HAP should provide coverage over an area of radius significantly more than 30 km. In fact, ITU-R recommends that a HAPS can have a footprint covering an area of up to 400 km diameter [63].

ITU-R recommendation SF. 1843 [63] considers a HAPS service area of approximately 400 km diameter. It partitions the coverage zones of HAPS stationed at an altitude of 21 km into urban area coverage (UAC) with radius between 0–36 km from the SPP (i.e. the point directly beneath the HAPS), suburban area coverage (SAC), which is between 36–76.5 km, and finally rural area coverage (RAC) between 76.5–203 km as illustrated in Figure 2.4. On the one hand, this recommendation recognises the possibility of providing extended coverage using HAPS. On the other hand, the extension is too optimistic, as it will be difficult to achieve practically due to the significant path loss and shadowing effects at extended distances among other factors. One other problem with extended coverage is beam broadening when beams are pointed at distances considerably away from the SPP. This causes cells deployed at the centre of the service area to differ in shape and size with those towards the edge, and it results in the degraded performance especially at the edge of coverage because the further away from the centre of coverage, the tighter the link budget becomes. To compensate for the broadening,

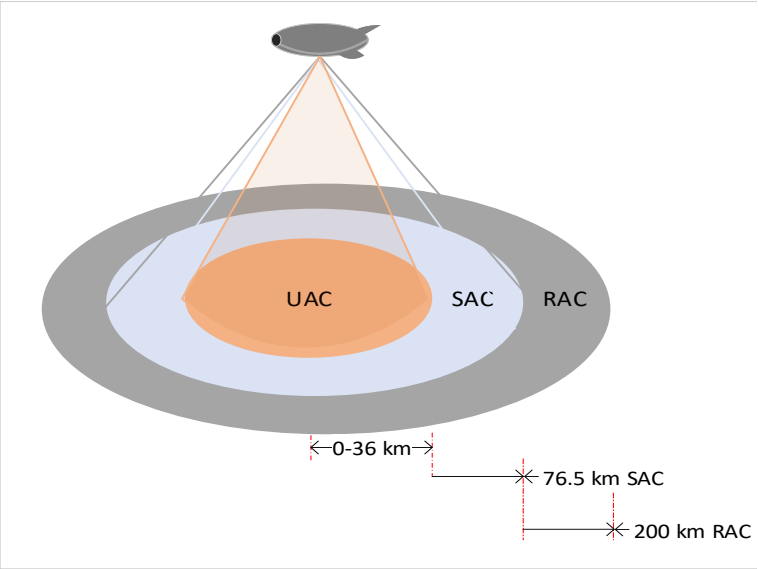

Figure 2.4: ITU-R defined high altitude platform station coverage zones.

proper pointing and shaping of beams are needed in addition to appropriate antenna system design. Power control and flexible dynamic resource management are some of the other techniques that can possibly be exploited to improve edge performance. The variation of data rate achievable between the centre and edge can be reduced for fairness by the use of adaptive modulation and coding, although, that is beyond the scope of this thesis. Consequently, in order to achieve extended coverage using HAPs, flexible beam deployment, effective radio resource and interference management are required among other factors.

2.4.1. Factors Affecting Coverage Extension

There are a number of factors affecting the achievable coverage extension from an aerial platform. This section highlights some of these factors.

Platform Placement

An aerial platform's operating altitude can influence its footprint and hence coverage on the ground. Therefore, optimizing the altitude of a HAP directly optimises its ground coverage. A few studies in the literature have investigated how the optimisation of the flying altitudes of platforms in general influences coverage and capacity. In [64–66], the authors present models for coverage optimisation based on platform placement, showing the relationship between an aerial platform and the maximum achievable coverage. The analytical frameworks presented show that the optimal altitude of an aerial platform, and by extension a HAP, is a function of the maximum allowable path loss. Increasing altitude results in a higher probability of line-of-sight (LoS) transmission, however, it also leads to increased path loss [66]. Therefore, it is important

to determine the optimum altitude, which minimises path loss while maximising the achievable coverage. However, the achievable coverage is directly dependent on the characteristics of the propagation environment, which varies [65].

Furthermore, HAP coverage is extendible by a mesh network of HAPs. Sharma and Kim derive a model for the coverage probability of a UE in a multiple UAV network [67]. In deriving the model, a ground-based UE is used with multiple UAVs placed in three-dimensional locations. One UAV provides coverage to the UE while the others seen as interferers are located along the boundary of a circle a distance R away from the reference UE with dynamic altitude control. The results show the impact of various system and channel parameters on received signal quality and coverage probability. Interference from the other UAVs in the mesh network can be managed with good platform placement and intelligent reuse schemes. The study in [68] analyses the coverage problems for the multi-UAV system, presenting an algorithm for optimum UAV placement and resource allocation. To extend HAP coverage, array antennas can be used to exploit the spatial characteristics of waves. Antenna radiation pattern, signal propagation, transmission parameters, resource and topology management are some other important factors that can be exploited to achieve a HAP coverage extension. A mesh network of HAPs using the multi-HAP architecture can achieve regional coverage.

Antennas

Long distance communication from HAP requires a very directive high gain antenna system with multi-beam and dynamic beam-pointing capability. The multi-beam horn and electronically steered antennas consist of individual horns and microstrip patch antennas respectively as antenna elements with prototypes for both types of antenna arrays developed in [69]. The low profile, suitability for both planar and non-planar surfaces, ease and lower cost of production with printed-circuit technology [5] are some of the factors that make electronically steered antenna systems ideal for HAPs. Large antenna arrays as in [70] result in an increased directivity and gain but also with increased sidelobe level. A compromise between the number of antenna elements and size, weight, cost, directivity and gain, mainlobe beamwidth, sidelobe level, and power consumption is required. Hence, appropriately dimensioning the antenna system is necessary. In phased antenna arrays, the total antenna gain can be estimated as the product of the array factor and an individual element gain. The total array factor, on the other hand, is the product of elements in the x- and y- axes for planar arrays. The detailed mathematical derivation of an antenna array factor is presented in [5]. Obviously, increasing the number of elements of the phased array consequently increases the antenna gain, but at the cost of an

increased antenna size.

Over the years, researchers have studied the effects of antenna characteristics on HAP communication systems. Thornton et al [71] compare the performance of circular and elliptical beam antennas in terms of carrier-to-interference ratio in a HAP cell operating in the mm-Wave bands. It is highlighted that elliptical beams result in optimised power at the edges, thereby providing better coverage than circular beam antennas. Also, having a steep roll-off factor results in quantifiable improvements in carrier-to-interference ratio, with the optimum value between 10–35 dB [62]. In [72], two techniques of HAP antenna array steering are proposed. While one technique involves individual steering of the antennas to maintain the boresight at the cell centre, the second dynamically compensates for aperture pointing errors resulting from platform displacement.

Propagation Channel

In comparison with terrestrial and satellite systems, radio communication links between a HAP and the ground are characterised by considerably lower propagation loss and multipath fading [73]. However, realistic and accurate propagation channel models are required for accurate predication of the HAP communication link. A number of the channel models in the literature are analytical [74, 75] with some empirical models available [76–78]. However, there are also measurement campaigns for empirical models in diverse environments and at different frequency bands, as reported in [79]. Note that the models reported are based on measurement campaigns carried out using low altitude platforms. Therefore, there is the inherent concern about whether these models can be used directly for high altitude communications considering the considerable differences in their propagation environments. Also, models derived from measurement campaigns are environment-specific. Nevertheless, actual measurement campaigns have not been conducted with a HAP at stratospheric altitudes to the best of knowledge of the author. Satellite-based models are dominated by path loss due to LoS and shadowing.

This thesis assumes a HAP-to-ground channel dominated by large scale fading. A path loss model that is comprised of free space path loss and shadowing loss, which is in line with the 3rd generation partnership project (3GPP) non-terrestrial network (NTN) channel model reported in 3GPP TR 38.811 [80, 81], is used as described in Section 4.2 of Chapter 4. Multipath fading is neglected as the focus here is on the HAP cellular structure and the long term mutual interference effects of the cells on each other. A rural African scenario with Rician fading channels and a largely open service area having limited or no scatterers is considered mainly. Therefore, the availability of a strong line of sight path is logically assumed at all times.

2.4.2. Radio Resource and Interference Management

Irrespective of the coverage area, it is important that HAP communication systems are designed to be spectrally efficient and coexist with terrestrial systems, which are highly likely to exist within the footprint of a HAP, given its wide coverage and the allocation of terrestrial spectrum for HAP use by the ITU-R. Adequate resource and interference management (RIM) facilitates this coexistence by mitigating against destructive interference, which improves the overall performance. With the increasing densification in wireless networks, the relevance of RIM for improved capacity and QoS cannot be overemphasised [82].

Ensuring the coexistence of HAPs and terrestrial systems is a major aspect of the RIM especially given the vertical heterogeneous network vision of future communication networks. Considering the reuse of the terrestrial frequency spectrum by HAPs, it is essential for both wireless systems to coexist with minimal harmful cross-interference between the networks. Perhaps, cross-interference between HAP and terrestrial systems is one of the major impediments to HAP deployment. In order to avoid excessive interference on the terrestrial system [1, 62], HAP cells must be pointed at an appropriate inter-cell distance away from the terrestrial cells. However, due to a HAP beam's fast received signal power roll-off with distance compared with a terrestrial system as shown in Figure 2.5, the inter-cell distance can be relatively short without causing excessive interference on the terrestrial system. Thus, adequate control of beam placement is necessary to mitigate cross-interference between the systems.

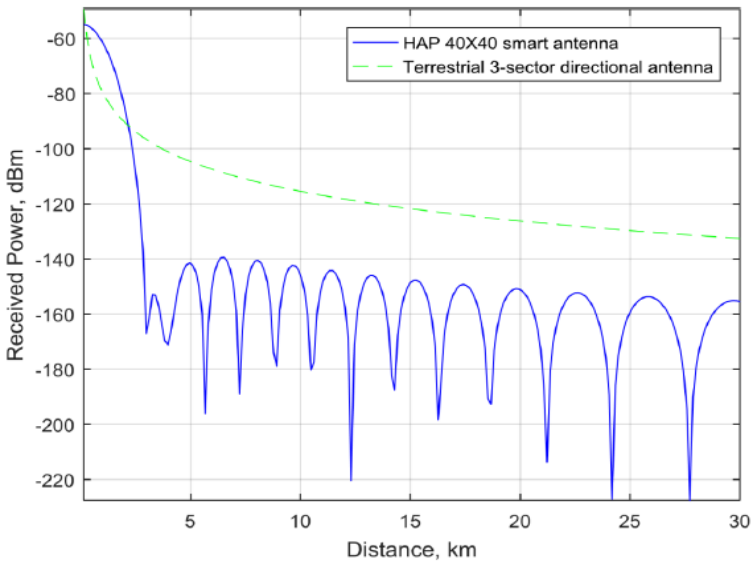

Figure 2.5: Power profile of terrestrial and HAP wireless systems (reproduced from [1]).

The beamwidth, mainlobe and sidelobe attenuation level of the HAP antenna beam pattern in Figure 2.5 are modifiable with tapering/windowing functions. This influences the size of

the cell formed by the beam and also the inter-cell interference, which affects HAP coexistence with terrestrial systems [1]. Windowing functions are mathematical functions applied in designing beamforming weights $w[n]$ to significantly reduce sidelobe levels by orders of magnitude at the cost of a slightly wider mainlobe [83, 84]. There are several windowing functions such as Blackman, Hamming, Hanning (Hann), etc. [85]. The work in this thesis uses Hann window due to its considerable relative sidelobe attenuation of about -32dB and steep roll-off, which significantly reduces ICI. Hann window is expressed as follows [86],

$$w[n] = 0.5\left(1 - cos(\frac{2\pi n}{N}\right), \quad 0 \leq n \leq N, \tag{2.1}$$

where $w[n]$ is the weight of each of the total N antenna elements in the antenna array.

The majority of HAP-Terrestrial system coexistence studies focus on interference management, resource allocations, and the individual system performance. Guaranteeing improved system capacity with limited interference on other systems needs advanced techniques such as diversity, advanced RIM, smart antennas, etc. [15]. The use of smart antennas and advanced RIM to ensure coexistence constitute the main techniques. Adopting cognitive radio concepts in formulating the dynamic spectrum management strategy can potentially facilitate coexistence [15]. While spectrum management avoids interference by appropriate utilisation of spectrum, smart antennas allow for spatial beamforming and isolation, which minimises interference in a particular direction. Furthermore, coexistence can be ensured based on ITU-R recommendation of keeping interference-to-noise ratio (INR) below a defined threshold at the primary receiver. In [87] INR and CINR based spectrum etiquettes are proposed for HAP systems. Both spectrum etiquettes use the INR or CINR level of an incumbent user respectively as a reference level, to manage a newly activated system downlink transmit power. Considering a HAP as the newly activated user, the studies ensure that interference on the terrestrial incumbent user does not cause a drop in INR or a change of the modulation scheme in use. Similar underlying ideas presented in [6, 88, 89], propose the implementation of an appropriate separation distance between the systems and antenna beam adjustment, as strategies that can be adopted to improve performance and facilitate the coexistence of the systems. The following sections present more discussions on some of the RIM techniques highlighted earlier.

Channel Allocation

Channel allocation has evolved from the very simplistic fixed allocation schemes, which involve artificially partitioning the available spectrum into chunks for fixed assignment, to dy-

namic spectrum allocation schemes, which allocate spectrum based on the state of the network. Fixed allocation can be implemented in HAPs by partitioning the available spectrum and allocating specific chunks between the HAP cells. However, although less complex compared to other allocation schemes, the fixed allocation scheme is rigid in dynamic traffic and network scenarios thus resulting in poor QoS and spectrum utilisation [90]. In dynamic allocation, which is the commonest in modern wireless systems, channels are assigned dynamically to users from a pool, based on request, availability, and network condition [90, 91]. For instance, the same resources in use in one cell can be reallocated to a user in another cell as long as the minimum interference threshold or maximum loading condition is satisfied. This is typical in communication systems with full spectrum reuse, where cells are assigned the same set of spectrum, which they dynamically allocate to their users based on traffic demand, interference from other tiers, etc. Resources can be dynamically allocated to users by cells, with or without inter-cell coordination, to facilitate traffic load balancing, mitigate against interference and underutilisation. This flexibility significantly improves the spectral efficiency and performance of the overall system. The work in this thesis use dynamic resource allocation.

Depending on the network entity in charge of the allocation, dynamic allocation schemes can be further sub-categorised as either distributed or centralised. Centralised schemes, which are suitable for HAPs, require complex computations for optimal performance, however, this complexity is minimised in distributed schemes but without the guarantee of optimality [92]. In [93], a centralised scheme that uses a harmony search to handle interference between nodes, which requires coordination for optimum network performance, is presented. Alternatively, studies in [94] propose a distributed scheme where channels are associated with prices. Users, which have allocated budgets for spectrum and can be assigned more than one channel, request for a set of channels based on their current prices and the users' utility. The complexity of the proposed algorithm is shown to be low.

Dynamic allocation schemes are consistently being enhanced with advanced schemes using genetic algorithm [95], game theory [96], and machine learning [97]. In [95], the proposed algorithm takes the channel active time, busy time and transmission time indicating the duration of time the channel is active, busy or employed for data transmission, into consideration. A model that considers a heterogeneous network made up of primary and secondary users, where a primary user has channel allocation priority over a secondary user, is proposed in [96]. In this case, the HAP can be the secondary user sharing spectrum with the primary terrestrial system. Studies in [97] highlight the challenges and opportunities of reinforcement learning for RIM. They propose an architecture to facilitate machine learning-based solutions in a radio access

network (RAN) in addition to their proposed reinforcement learning-based framework, with the ability to independently generate targeted RIM algorithms directly from measured data.

Cell Coordination

In modern cellular systems, coordination techniques are used to mitigate ICI and enhance system performance. For cross-interference mitigation between HAP and terrestrial systems, both systems can exploit coordination techniques. Coordination signalling and information exchange between cells can be carried through the backhaul links or X2 interface in long term evolution (LTE) BSs. Current state-of-the-art in cell coordination includes the use of ICIC and CoMP transmission. Major challenges of cell coordination techniques include synchronisation error and latency, which negatively affect the performance of the network [98].

ICIC: The goal of ICIC is to reduce interference between neighbouring cells by enabling frequent signalling exchange [99]. Based on the signalling highlighted in Figure 2.6, a BS informs its neighbours about the parts of the spectrum it is likely to interfere with them on. This information is then used by these neighbours to avoid resources used by the neighbour, which might cause harmful interference to those cells if used for high power transmission.

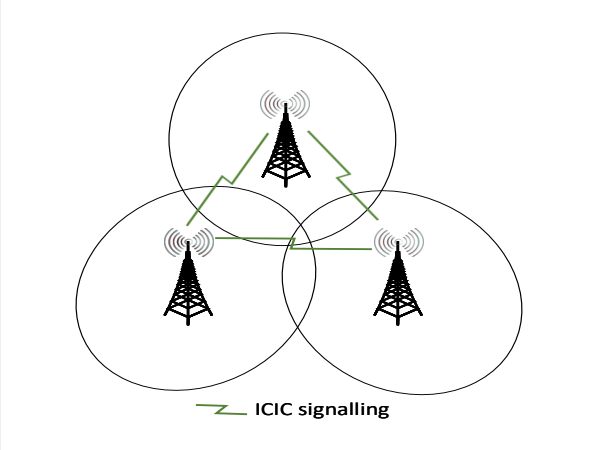

Figure 2.6: Inter-cell interference coordination signalling.

ICIC uses relative narrowband transmit power indicator (RNTP). It contains a bitmap indicating the resource blocks (RBs) a BS plans to transmit at high power on, which implies the RBs it may interfere with in neighbouring cells. In deciding whether a particular transmit power is high or low, a threshold derived from the average transmit power in the cell and the RNTP threshold is used. RNTP threshold is standardised and can take the following values measured in dB relative to the average transmit power in a given cell [100].

$$RNTP_{thr} \in \{\infty, -11, -10, -9, \cdots, 1, 2, 3\}\ dB. \tag{2.2}$$

In [101], centralised and distributed ICIC schemes that enable optimised cell associations and power allocations over multiple RBs in a UAV-assisted cellular network are proposed. The centralised ICIC scheme requires significant information exchange between BSs, which can be a challenge in a large network. On the other hand, the decentralised ICIC scheme lowers complexity and signalling overhead considerably for practical implementation.

CoMP: This enables BSs to cooperatively support data transmission to receivers and to constantly change channel state information, which improves cell and user performance. Multiple BSs coordinate with each other such that the signals from neighbouring BSs mitigates interference or are combined to improve the effective received signal quality. CoMP operation highlighted in Figure 2.7 require strict synchronisation between the coordinating BSs [2].

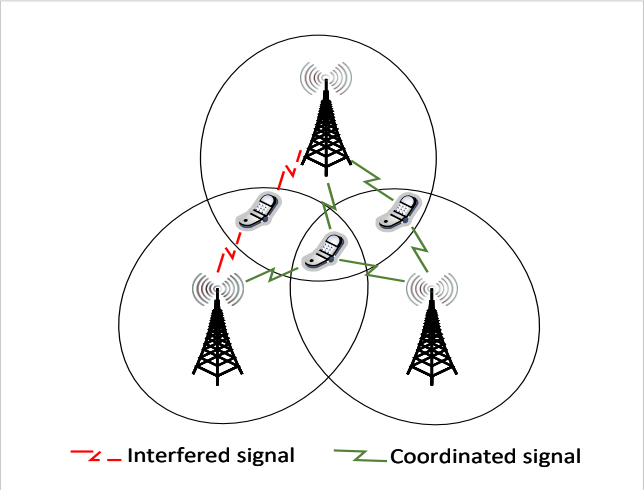

Figure 2.7: Coordinated multipoint operation (modified from [2]).

Pioneering the application of CoMP in HAPs are the studies in [41,57] where joint transmission JT-CoMP is used to improve performance. The proposed JT-CoMP requires coordination between overlapping cells to transmit the same signal simultaneously to a user in a CoMP region typically at cell edges, using an appropriate bandwidth allocation scheme. The authors classify cell edge users into non-CoMP, 2-way, and 3-way CoMP users based on the power level difference between the received signals. Participating CoMP cells reserve the same portion of their allocated bandwidth, which cannot be reused in any other region of the cells outside the CoMP region. The reserved bandwidth is allocated to the CoMP users depending on their location within the CoMP region. The results show improved CINR especially for users at the edge and overlapping regions. In [102], a proposed Q-learning-based channel allocation scheme uses

ICIC between neighbouring cells, which share their ICIC messages on a regular interval, to mitigate ICI, enhance system performance and speed up convergence.

Antenna Beam Management

Adequately predicting interference and managing antenna beams can help minimise interference. For instance, overlapping mainlobes and sidelobes of HAP antenna beams using the same radio channel introduces interference to other beams, therefore, adequately predicting interference is key to managing beams to ensure coexistence. In terrestrial systems, there is a difficulty in interference prediction at different locations due to the dissimilarity between terrains and fast changing propagation environment. On the other hand, HAPs can predict resultant interference at various locations because the propagation is mainly LoS [69].

Thornton et al. [53,71] propose a method of predicting co-channel interference based on curve-fit approximations of elliptic antenna beam radiation patterns illuminating cell edges with optimum power. They propose a means of evaluating optimum antenna beamwidths for each cell of a regular hexagonal layout. In [53], it is shown that heavy frequency reuse worsens interference at cell edges and this can be improved with sidelobe suppression. Adaptive beamforming and shaping, high minimum elevation angle, antenna radiation pattern improvement, and dynamic channel assignment are some beam management based interference mitigation techniques. In [40], the authors propose a method of creating irregular-shaped HAP cells by dividing the coverage area into a grid of small pixel spots grouped into the desired shape. With this approach, cells can be shaped in a way that avoids interference in areas of interest.

Exploiting Radio Environment Maps

Radio environment maps (REM) can potentially be used for enhancing resource utilisation, interference prediction and management in cellular networks [103]. It is important in the creation of dynamic interference maps of a particular location at a given frequency based on information gathered from measurement capable devices [104]. Some of the information are about the radio elements such as UEs and radio environment, differing from a geolocation database by its ability to process information gathered from measurement capable devices in conjunction with regulatory policies. While the radio elements maintain information of radio devices like location, mobility and transceiver, the radio environment characteristics contain the interference data [104]. For instance, UEs can report specific radio measurements with respect to their location, which can be used for optimisation. A proposed technique in [105] uses regression Kriging to reconstruct radio maps with mean received power at points with no available

measurements. Information like node location, BS height, transmission power provided by radio environment maps can potentially be exploited by a HAP using smart antennas for beam pointing, to avoid destructive interference on the terrestrial systems at given locations.

2.5. Challenges of Wireless Communications Using HAPs

One of the most fundamental challenges confronting the delivery of wireless communications services from HAPs is the design and development of aeronautic platforms, with the appropriate form factor capable of maintaining its station at stratospheric altitudes over a long duration, while carrying wireless communications payload and providing sufficient energy. Particularly, addressing the energy constraint issue, which is further discussed in Section 2.6, is important to ensure the feasibility and operability of solar-powered HAPs. Apart from this, some other challenges are faced in the development of appropriate wireless communication approaches and techniques such as cell formation, handover, backhaul and inter-platform communication, etc. This section presents some of these issues and challenges.

2.5.1. Extended Coverage Cell formation

Similar to terrestrial systems, providing wireless communication services from HAPs require the formation of cells on the ground. There are some challenges in forming these cells, which can be contiguous over the service area or provide islands of coverage based on the spatio-temporal distribution of users or traffic, especially for extended coverage. Firstly, as stated earlier, HAP cells pointed at distances from the SPP broaden in shape and size due to the limitations of antenna beamforming. The lack of consideration of broadening leads to a severe overlap between neighbouring cells especially in extended coverage scenario, which in turn results in significant ICI. Secondly, the antenna radiation pattern producing the beams, which are used to form cells, directly influences the ICI performance of the system. Therefore, in order to ensure an efficient HAP cellular communication system, cell formation must be careful considered, taking the effects of beam broadening and antenna radiation pattern into account.

2.5.2. Handover

Handover in HAPs may occur either due to user mobility or platform motion and instability. With the increasing intolerance for latency in communication systems, strict constraints on the handover process must be satisfied by the HAP system. This is particularly important in HAP extended coverage scenario where handover process could potentially introduce latency fundamentally due to distance and signalling. Handover in HAPs can be classified into

intra/inter-HAP and multi-layer HAP handover with intra/inter-HAP handover as the commonest. Intra-HAP handover occurs when a user moves between cells within the coverage of the same HAP, while inter-HAP handover occurs when a user moves between different HAPs. On the other hand, multi-layer HAP handover refers to user handover between HAPs and satellites or terrestrial networks, which can only happen in an integrated terrestrial satellite network. Typically, handovers are bound to occur in HAPs, which could be as result of platform or user motion or instability. Therefore, appropriate schemes that minimises the amount of signalling and latency are required in order for HAPs to satisfy the stringent latency requirements of NextGen systems. One possible approach is for the HAP to be able to predict when handover might occur so that the process could start beforehand by using changes in received signal strength of a mobile user as proposed in [106]. However, this approach still raises questions about the accuracy of the predictions especially considering platform displacements, which may be difficult to predict and also affect handover performance [107]. On the other hand, multi-layer handover refers to a handover occurring between a HAP and other communication systems such as terrestrial and satellite networks. This can be quite complicated to implement especially when a mobile user is moving between systems owned or operated by different platform operators.

2.5.3. Backhaul and Inter-platform Communication

Backhaul and inter-platform communication are extremely important in HAP wireless communication to enable the integration of HAPs into the wider communication network as they are not expected to operate in isolation in most cases. A general HAP communication scenario consisting of user access links, backhaul and inter-platform links and alternative backhauling via satellite is shown in Figure 2.8. Backhaul communication network is needed to integrate a HAP into the wider core network on the ground either directly or via a satellite, while inter-platform links interconnect multiple HAPs together to form a mesh network. Backhaul and inter-platform communication can be implemented through radio or FSO communication link, each of which has its advantages and disadvantages. The most suitable approach depends on many factors such as geographic location, available spectrum, bandwidth, radio regulations, prevalent weather condition, etc. Using radio communication for backhaul requires prohibitively high bandwidth at mm-Wave bands and transmit power, which is usually not available or allowable. Alternatively, FSO provides high bandwidth at a fraction of the power required for radio communication, however, it can be severely affected by weather conditions and pointing error [108]. For redundancy, a hybrid of radio and FSO links have been proposed.

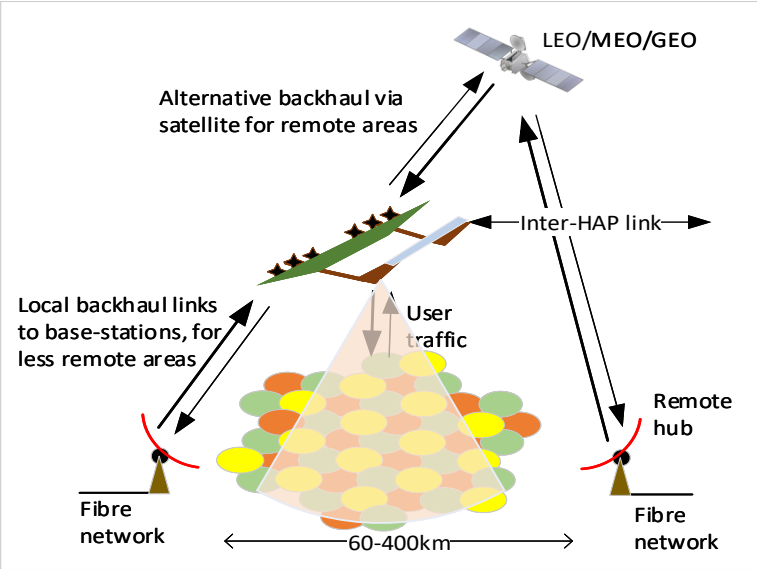

Figure 2.8: HAP communication links (modified from [3]).

However, this comes at the cost of increased weight and power consumption, both of which are limited on the HAP.

2.5.4. Networking

In the NextGen communication systems, user traffic will be completely based on internet protocols. Therefore, communication from HAPs requires the capability of routing internet protocol data packets from the source to destination. It is likely that most routing algorithms developed for terrestrial systems can not be directly applied in HAPs as they are based on the concept of shortest path between source and destination. In HAPs, this path can change due to either user or platform motion or instability, resulting in topology information rapidly becoming obsolete [69]. Therefore, appropriate consideration of the consistently changing condition of a HAP network in the design of networking protocols, algorithms and topologies is required.

2.6. Energy Feasibility Considerations For Solar-Powered HAPs

Energy feasibility is a necessary requirement for the use of HAPs for wireless communications. While numerous studies in the literature discuss the design and performance of solar-powered HAPs, most focus on aeronautics. Only a few studies [16, 17] consider solar energy harvesting in relation to solar-powered HAPs. However, while the emphasis of [16] is on aerodynamic structural design of a solar powered HAP, studies in [17] focus on the solar energy performance of a very small aircraft of 1 m wingspan. Assessing the feasibility of HAPs requires the understanding of their energy collection and consumption behaviour, which involves en-

ergy modelling. Early studies [109,110] on electrical systems and components at high altitude, considering the most significant energy consuming subsystems, suggest that solar cells at high altitude can potentially harvest enough energy to support an aircraft. Using a tethered aerostatic high altitude platform, a solar energy harvesting model is proposed in [111]. The harvested energy is transmitted to the ground using a power cable attached to the platform tether. The proposed model is based on tropospheric altitudes and expresses instantaneous solar intensity as a function of the atmospheric extinction parameter. This model cannot be straightforwardly used for HAPs operating in stratospheric altitudes due to the dissimilarity in the platform and altitude characteristics. Conversely, an energy model for stratospheric altitudes is developed in [112]. The model originally assumes that the aircraft is at ground level and then the resulting model is adjusted for stratospheric altitude, requiring parameters like refractive index and optical depth. The extrapolation of these parameter values can potentially introduce considerable inaccuracies in the results. Therefore, a reasonably more accurate, realistic and less complex energy collection model based on extra-terrestrial radiation suitable for stratospheric applications is required.

The major components of a HAP energy ecosystem are shown in Figure 2.9. The total energy that can be collected is limited by the wing area of the aircraft, which limits the effective area for the deployment of solar cells, in order not to undermine the aircraft's aeronautical design integrity [109]. On the other hand, the energy consumption involves the platform propulsion, avionic and payload systems. Consumption by the propulsion system is affected by factors

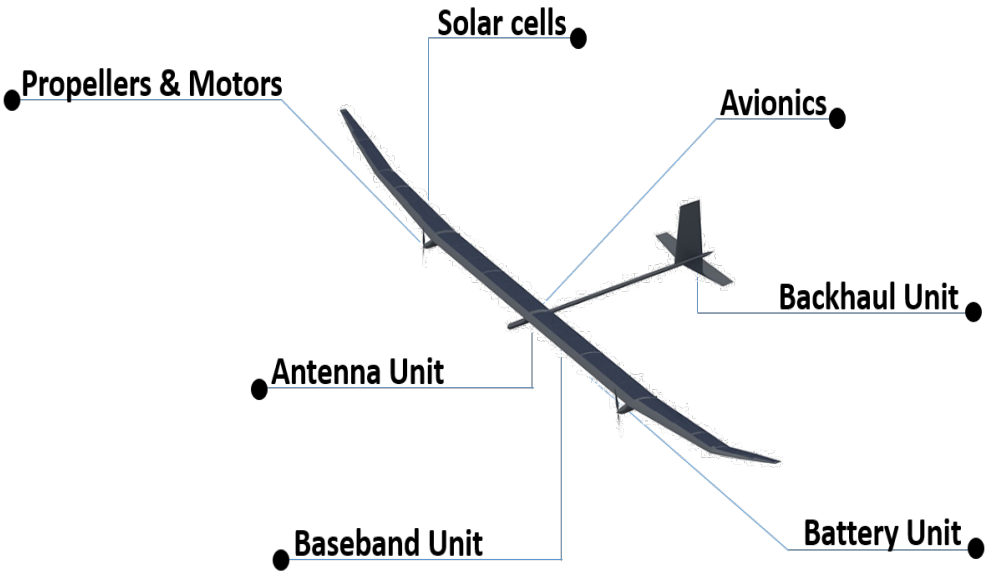

Figure 2.9: Components of solar-powered HAP energy ecosystem contributing to the total energy consumption of a solar-powered HAPs for wireless communication.

including payload and platform weight. The baseband, radio, antenna, energy storage, and backhaul units all contribute to the payload weight, while the platform frame and the avionic system contribute to the platform weight. The feasibility of the platform is not significantly affected by the weight of the solar cells [109] especially as high efficiency ultra-thin solar cells [113] are already available. For the backhaul network, FSO and mm-Wave technologies can potentially be used, therefore, the choice of technology depends on a compromise between size, weight, power, and required data rate. For instance, the antenna unit of mm-Wave technology comprises antenna arrays operating at V-band frequencies (e.g. 47/48 GHz allocated for HAPs by ITU-R [114]), which are expectedly small in size compared to other components in the ecosystem. A typical implementation of a HAP backhaul network at V-band frequency can therefore integrate the radio, baseband, and antenna units in a packaging.

Furthermore, the modularity of the components in the ecosystem is important for multi-mission solutions that require minimal reconfiguration for each mission. In [115], the authors propose a flexible and scalable architecture for the electrical power system of a UAV, which separates the power system into solar, storage, and payload modules. The solar and energy storage modules, which are some of the main components of the solar-powered HAP energy ecosystem, are briefly discussed below.

1. Ultra-thin solar cells - single crystalline based solar cell technology has been used for many years. The cost of Silicon, which dominates single crystalline based solar cells, catalysed the studies for alternatives. This resulted in two-compound technologies such as Gallium Arsenide. Thin film solar cells are a good technology for weight limited HAP applications. Different technologies such as the ultra-thin crystalline silicon [113] have also facilitated the production of ultra-thin solar cells.

2. Energy Storage - the most common energy storage systems proposed for HAPs are batteries and fuel cells. Typical examples of these systems usable in solar-powered HAPs are Lithium-Sulphur batteries and hydrogen fuel cells (HFCs). The ideal storage system for a given mission depends on location, platform characteristics, weight, specific energy, and deployment scenario, etc. Presently, battery-based storage systems are often one of the limiting factors for long endurance wireless communication missions due to their low specific energy. Alternatively, HFCs, which can provide continuous energy in the presence of hydrogen fuel, with significantly higher specific energy, can be used but at an increased cost due to the necessary replenishment of the hydrogen fuel. Adequate power for a given application is delivered by an appropriate stack of HFCs [116]. The

technology has been researched over the years and has been deployed in extra-terrestrial missions, with a hydrogen-powered HAP currently being developed [7]. In [117], the optimal sizing of a photovoltaic/hydrogen energy system that is ideal for HAP applications using a particle swarm optimiser is proposed.

The energy storage system is an essential component in solar-powered aircraft-based HAPs and must be critically considered. Since continuous operation of the wireless communication system is required, an adequate amount of energy is needed especially for night-time operation. This can potentially increase the required energy storage weight, which means a larger wingspan, is necessary. Therefore, appropriate sizing of the storage system can allow for the use of a smaller aircraft and an increase in mission endurance. It is shown in [117, 118] that methodically sizing the energy storage system increases the entire system efficiency. Lee et al. [119] apply an active power management method, which determines the power output for aerial vehicles powered by solar cells, fuel cells and batteries. The proposed method facilitates adequate power supply during the continuously changing flight conditions, keeping power sources within their proper operational bounds. In general, effective energy management strategies can ensure better energy system performance of a solar-powered aircraft-based HAP. This can involve path planning by allowing the aircraft to climb to a maximum altitude during sunrise, storing energy and gliding after sunset as shown in [109, 117, 120]. However, this will undoubtedly increase the complexity of the wireless communication system, as the effects of the platform mobility need to be taken into consideration.

2.7. Conclusion

This chapter presented the fundamental information that forms the basis of the work presented in the rest of the thesis. As highlighted in the chapter, providing wireless communications using HAPs have some comparative advantages over alternative terrestrial and satellite systems. However, there are some issues and challenges confronting the deployment of HAPs for wireless communications. A number of studies have investigated some of these challenges and proposed different solutions and techniques to facilitate the use of HAPs to provide coverage. Nevertheless, these state-of-the-art studies focus on a coverage area of within 30 km radius, which is pessimistic given the good propagation environment and elevated look angle of HAPs. Therefore, this thesis focuses on approaches to significantly extending this coverage area, which will maximise their benefits. This chapter gave an overview of HAPs, which includes their classification, radio and aeronautic regulations as well as some relevant past and

current projects. Furthermore, the background of providing wireless communications using a HAP was introduced. Particularly, discussions were presented on network topology, cellular architecture, antenna beam pointing radio resource and interference management as well as challenges of HAPs wireless communications systems. This chapter, which mostly focused on HAPs use for coverage extension and capacity enhancements, and also highlighted some the factors affecting coverage extension. In addition, the energy feasibility of HAPs within the context of energy collection and consumption, given that they are part of the main constraints of HAPs was discussed.

Chapter 3. Energy Management of A High Altitude Platform for Wireless Communication

Contents

3.1. Motivation

Over the years, a considerable amount of work has been carried out on the research and development of solar-powered HAPs for long endurance wireless communications missions, as highlighted in Chapter 2. However, a significant majority of these completely neglect the feasibility of solar-powered platforms for such missions in terms of solar energy availability and usage by both the platform and telecommunication payload. Energy management of these systems is a fundamental task that needs careful consideration to ensure technical and operational feasibility. In fact, it determines whether the solutions and schemes developed for continuous HAP operation in the rest of the thesis can be practically and feasibly deployed. For continuous HAP operation throughout the desired mission duration, the energy requirements of the subsystems govern how much solar energy is needed to maintain flight and support the payload during the day, as well as the needed energy storage for night-time usage [121]. These are determined by the technical specifications of the subject platform.

Given that the most common focus on energy source for HAPs is on solar, the purpose of this chapter is to facilitate the understanding of how the design and operation of a solar-powered HAP might affect the feasibility of providing wireless communications over varying areas with respect to power and weight limitations of the platform considering the current state-of-the-art. This requires adequate solar energy collection and consumption modelling for both platform and payload, as well as showing the necessary calculations required to highlight the methodology of using solar-powered HAPs, the lack of which in a wider sense has led to the failure of some HAP projects in the past. Specifically, the chapter focuses on energy management of HAPs deployed as base stations (BSs) at different latitudes for wireless communication to ground-based users, over the much studied coverage area of 30 km radius and the extended area of 60 km radius as proposed in Chapter 4. While significant work on the design of HAPs for wireless communications is ongoing in the industry, there is a lack of details of some of the technical specifications like chord length, wing area, platform and payload power consumption in open literature, which this chapter aims to address.

3.2. Energy Collection Model

The intensity of solar radiation impinging on a horizontal surface is dependent on the operating latitude and the sun's position in the sky. Then, the total energy collected by the HAP is dependent on the incident angle of the sun's rays on the solar cells placed on the platform, the duration of daylight, and air mass [122]. While observing the sun from the earth, it appears to move along a great circle on the celestial sphere concentric to the earth called the ecliptic as shown in Figure 3.1. The celestial sphere is an imaginary sphere of an arbitrarily large radius. Realistically, the earth is rotating around an axis z', tilted by an angle of obliquity ε = 23.44° with respect to the ecliptic frame [123]. On the ecliptic frame, the sun's latitude is 0°.

The extra-terrestrial solar radiation, which is the solar radiation detected above the atmosphere, is higher than that at the surface of the earth due to gaseous and water vapour absorption of radiation at some wavelengths in the atmosphere. Fortunately, in the stratosphere, the amount of water vapour is less than 0.05% of the amount on the earth's surface [112]. The expression for the instantaneous extra-terrestrial solar intensity, G_i is given as [117]

$$G_i = G_{sc}\left(1 + 0.033\cos\left(\frac{360d}{365}\right)\right)\left(\sin(L)\sin(\delta) + \cos(L)\cos(\delta)\cos(h)\right), \tag{3.1}$$

where G_{sc} = 1366.1 W/m^2 is a standard solar constant at zero air mass defined by the American

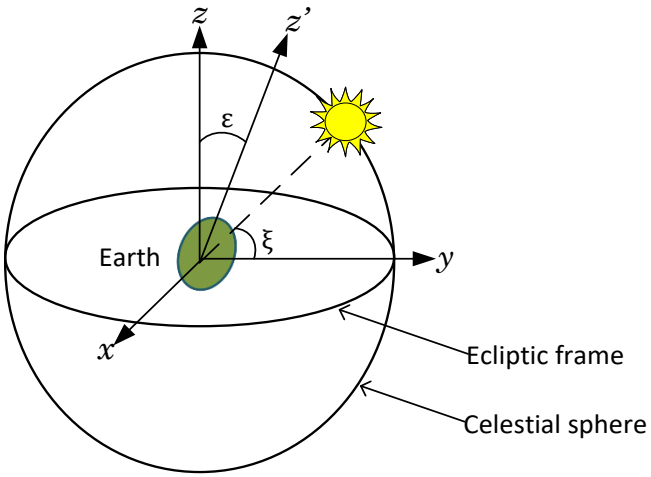

Figure 3.1: The earth and the sun on the celestial sphere. ξ is the slowly changing elevation angle of the sun as observed from earth while ε is the angle of obliquity. $x,\ y,\ z$ are coordinate axes at an observation point.

Society for Testing and Materials (ASTM) standard E490, δ is the solar declination angle, which is a function of day d of the year, L is the latitude and h is the solar hour angle.

Assuming that the HAP is maintained horizontally throughout its mission duration, the solar irradiance on a plane horizontal to the surface of the earth is the sum of the direct, diffuse and reflected components of the solar radiation intensity. In the stratosphere, with very low levels of aerosols and water vapour resulting in a negligible probability of scattering [124], the Rayleigh distributed diffuse components can be neglected. It is only logical that the solar cells are deployed on the parts of the platform facing upwards, hence the reflected component is also neglected. Therefore, only the direct component is considered subsequently. Additionally, it is assumed that the HAP is horizontal and parallel to the surface of the earth, and that the sun 'rises' from one side of the horizontal plane with respect to the HAP and 'sets' on the other side of the plane as depicted in Figure 3.2. The sun is seen to be rising at position 1, its trajectory is depicted in between its setting position at 6. Note that the motion of the sun is continuous, the numbering is just for depiction. The angle θ is the solar zenith angle, which is a function of the position of the sun. It is important to note that the sun never really gets to an elevation angle of 90° (90°-θ, zenith angle of 0°) in most cases. This only occurs at the equator.

The radiation intensity expressed in (3.1), which is a function of the solar hour angle requires computation of the intensity at any particular time, and the solar hour angle must be known. This increases the necessary computation. To simplify, the maximum radiation intensity G_{mi} that can be experienced at any latitude is detected at the equator at solar noon (i.e. hour angle,

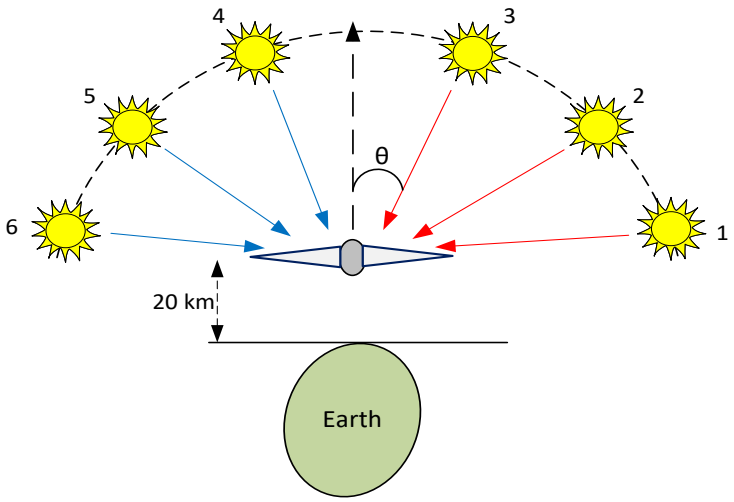

Figure 3.2: The movement of the sun around the platform from sunrise to sunset. The elevation and zenith angles of the sun are slowly changing at the same rate from sunrise to sunset.

$h = 0°$) and given as [118]

$$G_{mi} = G_{sc}\left(1 + 0.033\cos\left(\frac{360d}{365}\right)\right)\left(\sin(L)\sin(\delta) + \cos(L)\cos(\delta)\right), \tag{3.2}$$

where the solar declination angle δ, expressed as a function of day of the year, is given in radians as

$$\delta = 0.4093\sin\left(\frac{2\pi\,(d - 79.75)}{365}\right). \tag{3.3}$$

Introducing Lambert's cosine law, which states that the irradiance measured on a plane will vary with respect to the cosine of the angle between the optical axis of the source and normal to the detector as shown in Figure 3.3, combining (3.1) and (3.2), the following is obtained.

$$G_i = G_{mi}\cos\left(\frac{\pi}{2} - \xi\right), \tag{3.4}$$

where ξ is the elevation angle of the sun.

The radiation intensity of the sun on the plane of the platform varies with the incidence angle, while the total energy collected on a given day of the year is also dependent on the number of daylight hours. The total irradiance per m^2 on a given day is evaluated between sunrise and sunset. Since the elevation angle, as observed at the HAP, varies with time resulting in a

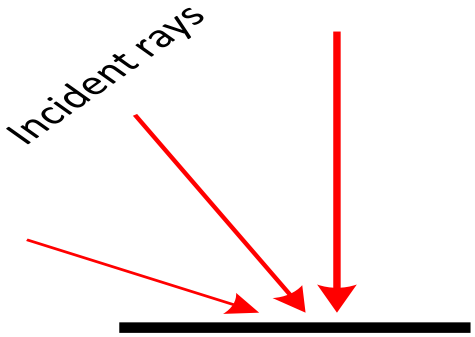

Figure 3.3: The depiction of Lambert's law. Intensity at the plane is maximum when the source is perpendicular. Intensity decreases with increasing angle of the source from the perpendicular axis as indicated by the size of the arrows.

time-varying solar radiation intensity, the sun's elevation angle as a function of time duration since sunrise on a typical day, which can be referred to as elevation angle time series function, is defined. An approximate elevation angle time series is derived to facilitate the formulation of an expression to calculate the total solar irradiance per m^2 on a given day. In deriving the elevation angle time series function for a typical day, it is assumed that at sunrise the sun's elevation angle as observed from the HAP is 0° as the sun is in line with the HAP and that it is consistently increasing afterwards. There is a direct proportionality between time duration since sunrise t and the elevation angle at that particular duration $\xi(t)$. For instance if sunrise is 06.00 am, then $t = 0$ at this time while 08.00 am $t = 2$ hours. The elevation angle continues increasing slowly with t until noon at which point the maximum elevation angle for the day has been attained. Then, it starts decreasing at a similar rate until sunset when it equals 0° again. This results in a triangular elevation angle time series function given in (3.5) and pictorially represented in Figure 3.4. Actually, in [4], it is shown that the elevation angle

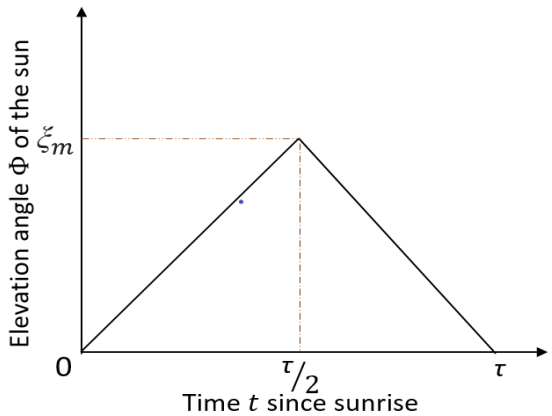

Figure 3.4: The sun's elevation angle time series function for a typical day. ξ_m is the maximum altitude of the sun in radians. τ is the time (in hours) from sunrise to sunset at the given latitude. The rising part of the triangle indicates sunrise while the falling part indicates sunset.

time series, heuristically determined for some select days in a year, is not exactly represented by a triangular function, but rather, the function is almost parabolic in shape as highlighted in Figure 3.5. However, (3.5) is a pessimistic approximation of the elevation angle time series function as the sun does not immediately start descending after reaching the peak for the day. Although, the accuracy of the function increases with increasing maximum elevation angles as Figure 3.5 also highlights. This is particularly significant as HAPs are expected to operate mainly in regions with high maximum elevation angle. It is important to note that the triangular function potentially leads to a point of singularity at noon. However, this does not affect total solar irradiance calculation, which is analogous to evaluating the area under the triangle. This can be achieved straightforwardly irrespective of the singularity.

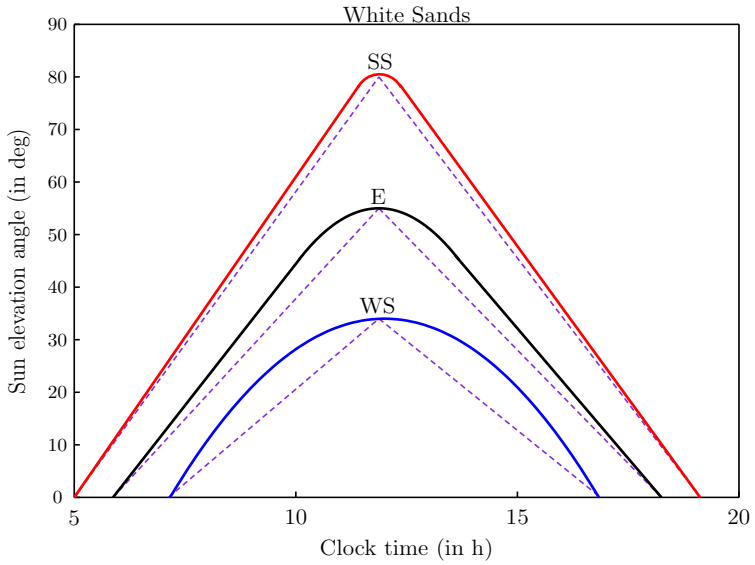

Figure 3.5: Elevation angle time series for spring equinox (E), winter and summer solstice (WS & SS) from [4] annotated with the proposed triangular time series function. The dashed line shows the proposed triangular function with increasing accuracy with respect to the increasing maximum elevation angle.

The elevation angle time series function is expressed mathematically as follows,

$$\xi(t) = \begin{cases} \frac{2\xi_m}{\tau}t, & 0 < t \leq \frac{\tau}{2} \\ 2\xi_m - \frac{2\xi_m}{\tau}t, & \frac{\tau}{2} < t \leq \tau \end{cases}. \tag{3.5}$$

Substituting (3.5) into (3.4) yields

$$G_i = G_{mi} \begin{cases} \cos\left(\frac{\pi}{2} - \frac{2\xi_m}{\tau}t\right), & 0 < t \leq \frac{\tau}{2} \\ \cos\left(\frac{\pi}{2} - \left(2\xi_m - \frac{2\xi_m}{\tau}t\right)\right), & \frac{\tau}{2} < t \leq \tau \end{cases}. \tag{3.6}$$

The total extra-terrestrial solar irradiance G_{Ti} per m^2 on the surface of a HAP at stratospheric altitude on a given day of the year at a particular latitude is therefore computed by integrating (3.6) over the duration between sunrise and sunset as shown below.

$$G_{Ti} = \int_0^{\frac{\tau}{2}} G_{mi} \cos\left(\frac{\pi}{2} - \frac{2\xi_m}{\tau}t\right) \delta t + \int_{\frac{\tau}{2}}^{\tau} G_{mi} \cos\left(\frac{\pi}{2} - \left(2\xi_m - \frac{2\xi_m}{\tau}t\right)\right) \delta t. \quad (3.7)$$

Integrating the triangular function over the entire range $0 - \tau$ is equivalent to integrating only half, e.g. 0–$\frac{\tau}{2}$, and multiplying the result by 2. Therefore, (3.7) can be simplified as follows.

$$G_{Ti} = 2G_{mi} \int_0^{\frac{\tau}{2}} \cos\left(\frac{\pi}{2} - \frac{2\xi_m}{\tau}t\right) \delta t. \quad (3.8)$$

Evaluating the integral in (3.8) yields the expression for G_{Ti}, which is given as

$$G_{Ti} = \frac{G_{mi}\tau}{\xi_m} \left(1 - \cos(\xi_m)\right). \quad (3.9)$$

Using (3.9) requires the knowledge of the time duration τ from sunrise to sunset at a particular latitude, and the maximum elevation of the sun on a given day of the year. In [123], closed form expressions of the daylight time duration, $\tau(L, d)$ as a function of the azimuthal angle, $\varphi(d)$ of the sun and mean anomaly $M(d)$ (the angular distance of the sun) are given as follows,

$$\tau(L, d) = 24\left(1 - \frac{1}{\pi}\cos^{-1}(\tan(L)) \times \left(\frac{\sin(\varepsilon)\sin(\varphi(d))}{\sqrt{1 - \sin^2(\varepsilon)\sin^2(\varphi(d))}}\right)\right). \quad (3.10)$$

$$\varphi(d) = -1.3411 + M(d) + 0.0334\sin(M(d)) + 0.0003\sin(2M(d)). \quad (3.11)$$

$$M(d) = -0.041 + 0.017202d, \quad (3.12)$$

The maximum sun's elevation ξ_m on a given day of the year is a function of latitude and declination angle [123]. It varies with day of the year and is estimated in radians using [125]

$$\xi_m(L_r, \delta) = \frac{\pi}{2} + L_r - \delta, \quad (3.13)$$

where L_r is the latitude in radians, which can be positive or negative, depending on whether it is referencing the northern or southern hemisphere respectively.

3.2.1. Energy Collection Model Application

Using equations (3.12)–(3.13), the daylight time duration at different locations with varying latitudes over the duration of 1 year are evaluated and plotted as shown in Figure 3.6. The locations near the equator have an approximately constant time duration of daylight from sunrise to sunset. Moving north of the equator, the longest duration of daylight is witnessed on the day of the mid-year solstice which occurs on 21st June (i.e. d = 172). This can be as long as 17 hours of daylight in York, United Kingdom (Lat: 53.96°N) for instance. Conversely, the shortest day witnessed at all locations north of the equator occurs on the year-end solstice (i.e. 21st December, d = 355). Also in York for instance, this could be as short as just 7 hours. The total solar irradiance per m^2, G_{Ti} at locations as in Figure 3.6 is computed using (3.9) and plotted for different days of the year.

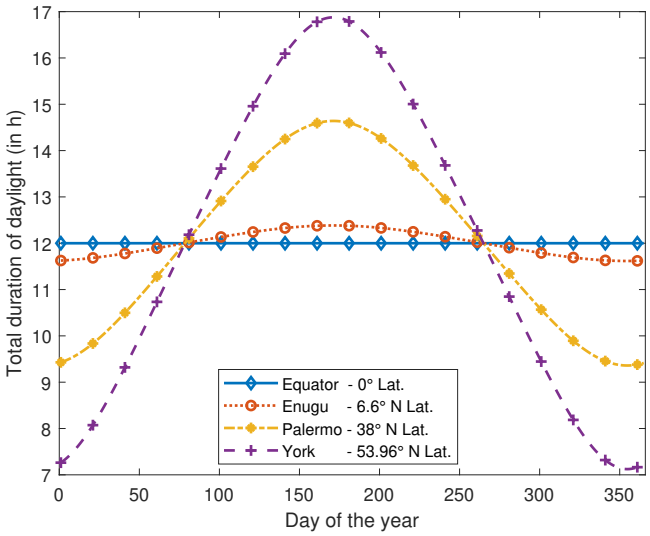

Figure 3.6: Duration of daylight for given days of the year. This varies as the distance of sun from earth varies throughout the year.

Figure 3.7 shows that G_{Ti} in equatorial regions ranges between 7–8 kWh/m^2/day around June/July when the sun is furthest from the equator. This value increases to approximately 11 kWh/m^2/day towards the beginning and the end of the year when the sun is closest to the equator. On the contrary, G_{Ti} attains its maximum value of approximately 14 kWh/m^2/day in regions further up the northern hemisphere during the middle of the year. The main reason being the significantly longer duration of daylight experienced during this time of the year [126]. Unfortunately, the G_{Ti} drops to its minimum value of approximately 1.5 kWh/m^2/day on De-

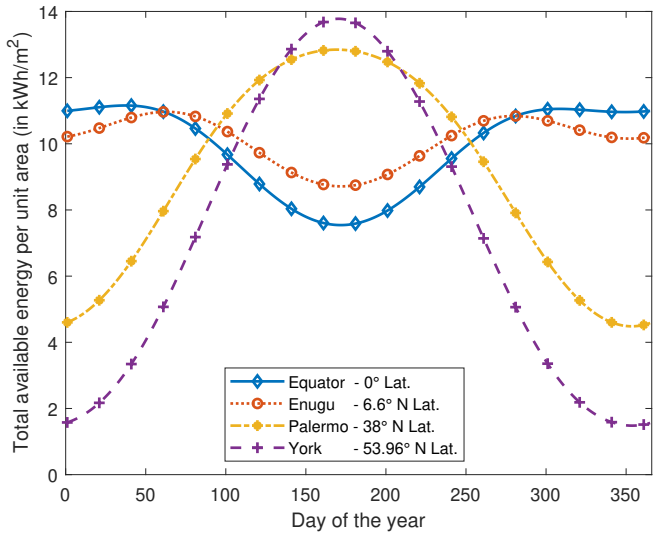

Figure 3.7: Total solar irradiance per m^2 at various locations for a given day of the year highlighting the longer days in northern hemisphere.

cember the 21st in the northern hemisphere, due to the short duration of daylight. It is assumed that battery storage, which is currently the most technologically matured energy storage system, is used to store excess generated energy for use at night. However, the effect of possibly using storage systems with high specific energy such as fuel cells is investigated and analysed later in Section 3.4.

Therefore, continuous year-round operations of a solar-powered HAP in the northern hemisphere are constrained by the duration of daylight on the year-end solstice. To assess the operational feasibility, the total energy, E_{Tc} collected by the platforms on the year-end solstice is evaluated with the total energy consumption by the platform and payload to ensure adequate availability of energy. Considering the total area of solar cells deployed on the platform and the efficiency of the solar cells, E_{Tc} is expressed as follows [117],

$$E_{Tc} = \eta_{pv} A_{pv} G_{Ti}, \tag{3.14}$$

where η_{pv} and A_{pv} are efficiency and area of the photovoltaic system.

The analysis in the rest of this chapter assumes that energy can only be stored for 24 hours [127], which is logical because the HAP is SWaP constrained. Due to the constraints of HAP, designing a HAP that can operate continuously over a period of time is based on the worst day, which is the day with the lowest duration of sunlight. This determines the available power and therefore the minimum weight of the energy storage system for the HAP. If the

HAP can be designed to work on this day, then, it can work on any other day that would have more generated power and require less storage for instance. Increasing the storage capacity beyond the minimum required for the worst day will also increase the weight of the storage system, which will in turn increase the energy consumption of the platform. In fact, to store extra energy long enough to be significant will require doubling the minimum weight at least, which will likely violate the SWaP constraints. For instance, assuming there is excess energy available in summer that could be stored for use in winter, the amount of energy required to make a difference in winter would be significant and require more weight for storage, which is a severe penalty for the HAP. Therefore, due to the SWaP constraints, it is more logical to design for the worst 24 hours to obtain the minimum power and weight needed by the HAP, which also works on every other day.

3.3. Energy Consumption Model

The energy consumed by a HAP can further be subdivided into the energy consumed by the following subsystems:

1. Avionics (platform control)
2. Propulsion
3. Payload

The total energy consumption of the aircraft E_{tot} is therefore modelled as the sum of the power consumption due to the propulsion P_{req}, avionics P_{Av}, and payload P_{Pl} multiplied by the duration of operation T as given in the following.

$$E_{tot} = (P_{req} + P_{Av} + P_{Pl})\, T. \tag{3.15}$$

where T is the duration of operation of the HAP under consideration.

3.3.1. Propulsion System Energy Consumption

The propulsion system's energy consumption is modelled based on the orientation of the HAP and its flight pattern. Two models are developed with the consideration of the following distinct flight patterns and HAP orientations:

1. Steady Horizontal flight: where the platform maintains a steady horizontal flight in the stratosphere moving in the direction of thrust.

2. Steady Circular Flight: where the platform is in a circular flight, operating within a position circle of a given radius and at a fixed banking angle.

Horizontal Flight

Firstly, in order to evaluate the energy consumption by the propulsion system of a HAP, it is assumed that the aircraft is maintaining a steady horizontal flight. The principal forces acting on the platform in flight are depicted in Figure 3.8.

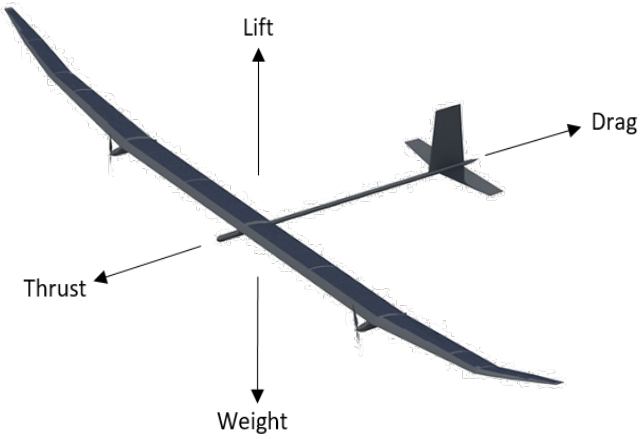

Figure 3.8: Forces acting on a HAP in flight.

In a steady horizontal flight state, the platform lift force, F_l and weight, F_w are equal. Also, the forward thrust, F_t equals its drag F_d. The lift force, F_l is generally expressed as [128]

$$F_l = F_w = \frac{1}{2}\left(\rho V^2 A_w C_L\right), \tag{3.16}$$

where ρ is the air density, V is the velocity of the aircraft, A_w is the platform's wing surface area and C_L is the coefficient of lift. Rearranging (3.16),

$$\frac{F_w}{C_L} = \frac{1}{2}\left(\rho V^2 A_w\right). \tag{3.17}$$

Similarly, the forces of thrust, T_t and drag, D_g can be expressed as a function of the drag coefficient C_D in the following equation.

$$F_t = F_d = \frac{1}{2}\left(\rho V^2 A_w C_D\right). \tag{3.18}$$

Thus, substituting (3.17) into (3.18) gives

$$F_t = F_w \left(\frac{C_D}{C_L} \right). \tag{3.19}$$

The power P_{req} required to maintain the steady flight, as defined in the following equation, is a function of thrust F_t, velocity V, and propeller efficiency η_{prop} of the HAP [129].

$$P_{req} = \frac{F_t V}{\eta_{prop}}. \tag{3.20}$$

Solving for V in (3.16) and substituting both (3.19) and the solution for V into (3.20), P_{req} in Watts, which is used in Section 3.4.2 to estimate the required power by the representative example platforms, can be expressed as

$$P_{req} = \left(\frac{C_D}{\eta_{prop} C_L^{\frac{3}{2}}} \right) \sqrt{\frac{2F_w^3}{\rho A_w}}, \tag{3.21}$$

where $F_w = m_T g$ is the weight of the platform, m_T is the total mass of the platform, and g is acceleration due to gravity.

In the rest of this chapter, the propeller efficiency η_{prop} used for calculations is 80%. This is based on the comprehensive study of UAV propeller performance in [130]. A perfect electric motor with 100% efficiency is assumed to simplify the analysis. In practice, electric motors for light aircraft such as unmanned HAPs have efficiencies up to 98% [131,132], which is high enough to make the assumption logical.

Steady circular flight with banking

Assuming the HAP follows a circular trajectory within a circle of radius R, at a bank angle α as in Figure 3.9. ITU-R and the Helinet project [133] recommend different values for R.

For the HAP in Figure 3.9 to maintain steady circular flight at the same altitude, the vertical component of the lift force, F_{lv} must be equal to the total weight F_w.

$$F_{lv} = F_l \cos(\alpha) = F_w = m_T g, \tag{3.22}$$

Furthermore, F_C is the centripetal force pulling the aircraft towards the centre of the circular flight path. It is given as,

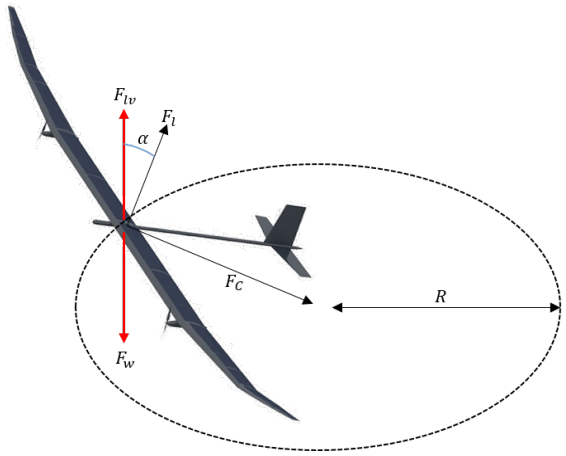

Figure 3.9: Steady circular motion of a HAP flying at a banking angle α.

$$F_C = F_l \sin(\alpha) = \frac{m_T V^2}{R}, \tag{3.23}$$

Combining (3.22) and (3.23), the banking angle α is obtained as follows.

$$\alpha = \tan^{-1}\left(\frac{V^2}{gR}\right). \tag{3.24}$$

The turn radius can be defined based on (3.24) as $R = \frac{V^2}{g\tan(\alpha)}$. Since the directions of the centripetal force and acceleration are towards the centre of the circle, the tangential acceleration is equal to zero. In this case, the significant forces acting on the HAP are the centripetal force, the total weight of the platform and the lift force. Following similar assumptions as in the case of the steady level flight above,

$$F_w = F_l \cos(\alpha), \tag{3.25}$$

$$\frac{F_l}{F_w} = \frac{1}{\cos(\alpha)} = L_f, \tag{3.26}$$

where L_f is referred to as the aircraft load factor. The power required P^c_{req} to maintain a steady circular flight at a bank angle α is a product of the square of the load factor at a given bank angle and the power required for a steady level flight [134]. Therefore,

$$P^c_{req} = \left(\frac{1}{\cos(\alpha)}\right)^2 P_{req}. \tag{3.27}$$

3.3.2. Payload Energy Consumption

The HAP payload is the BS apparatus that provides wireless access to the users and a backhaul link from the HAP to the infrastructure on the ground. Detailed studies have been carried out on the energy efficiency of a terrestrial long term evolution (LTE) BS by the EARTH project [135, 136]. The project proposed an energy efficiency evaluation framework (E^3F), which can be applied in estimating the power consumption of a terrestrial BS. The framework facilitates the estimation of the power consumption of a BS as the summation of the power consumed by the BS baseband unit, the power amplifier (PA), radio frequency (RF) chains and overheads. The model highlights that baseband power is dependent on variables such as the system bandwidth, modulation scheme, coding rate, number of antennas, time-domain duty-cycle and frequency-domain duty cycle. Time-domain duty-cycle is the fraction of time during which the baseband circuitry is active, while the frequency-domain duty cycle is the proportion of the frequency resources in use. Each of these parameters are weighted and scaled according to their contribution to the overall baseband consumption. The weighting and scaling are based on similar parameters obtained from a reference terrestrial BS whose power was obtained by measurements. Studies in [136] give the power consumption of the different sub-components of the RF chain under transmitting and receiving scenarios.

Using the EARTH model developed for terrestrial BS power consumption estimation, a HAP BS power consumption is estimated. The EARTH project [135, 136] proposed an Energy Efficiency Evaluation Framework (E^3F). The framework estimates the power consumption of a BS as the summation of the power consumed by the base station sub-systems including the RF chain. The power is dependent on variables such as the system bandwidth BW, modulation scheme M, coding rate R, number of antennas Ant, time-domain duty-cycle D_t and frequency-domain duty cycle D_f. Time-domain duty-cycle is the fraction of time during which the baseband circuitry is active, while the frequency-domain duty cycle is the proportion of the frequency resources in use. Each of these parameters are weighted and scaled according to their contribution to the overall baseband consumption. According to [136], power consumption can be modelled as,

$$P_{RF} = \sum_{i \in I_{RF}} P_{i,ref} \prod_{x \in X} \left(\frac{x_{act}}{x_{ref}} \right)^{s_i,x} \tag{3.28}$$

where I_{RF} represents a set of the sub-system components, $x = [BW, Ant, M, R, D_t, D_f]$ is the set of parameters affecting each subcomponent, $s_i = [1, 1, 1, 1, 1, 1]$ is the correspond-

ing scaling factor. Note the values for reference $x_{ref} = [20, 1, 6, 1, 100, 30]$ for BW, Ant, M, R, D_t and D_f respectively. Assuming the HAP BS to be operating with a 20 MHz system bandwidth, 1600 antenna elements, PA efficiency of 47%, with 100% time-domain and frequency-domain duty-cycle, the total estimated HAP BS power consumption obtained using (3.28) is in excess of 60 kW. See [136] for a more detailed explanation. This is very pessimistic and the significantly high power consumption results from the model's lack of consideration of antenna array scenario. In addition, HAP applications were not considered in the development of the model. These illustrate the shortcomings of trying to evaluate a HAP based system using models optimised for terrestrial operation. Given that the EARTH model was developed for terrestrial BS, it is inappropriate to apply it for a bespoke systems such as a HAP BS.

Since there is no appropriate model for HAP BS power consumption estimation in the literature, an alternative model is therefore derived herein. If a HAP BS consisting of N antenna elements forms a maximum of C cells on the ground, it can be assumed that each cell has a maximum RF output power P_o. Considering the PA efficiency η_{PA}, the power input to the PA is estimated as

$$P_{PA} = \frac{CP_o}{\eta_{PA}}. \tag{3.29}$$

In the EARTH model, the PA is estimated to contribute approximately 57% of the total BS power consumption. The total power consumption by a HAP BS can therefore be expressed as a function of the fraction of the consumption by its PA as follows

$$P_{Pl} = \frac{CP_o}{\gamma\eta_{PA}}, \tag{3.30}$$

where γ is the fraction of total power consumption by the BS fronthaul power amplifier. The whole communication system payload power consumption P_{Pl} can therefore be expressed as

$$P_{Pl} = \frac{CP_{out}}{\gamma\eta_{PA}} + P_{bh}, \tag{3.31}$$

where $P_{bh} = \frac{C_b P^b_{out}}{\gamma^b \eta^b_{PA}}$ is the backhaul power consumption with C^{bl} backhaul links, P^b_{out} backhaul link output power, the fraction of total backhaul power consumption by the backhaul power amplifier with efficiency η^b_{PA}. Backhaul is typically implemented in mm-Wave frequencies with a few links, therefore, compared with the fronthaul with many cells and kilowatts of

power consumption, the backhaul with consumption ideally within tens of watts of power consumption is neglected in this thesis to simplify (3.31). This is justified by the fact that BS fronthaul is responsible for over 80% of the BS power consumption [137].

In order to understand the cellular performance within a HAP service area, estimation of the number of cells needed to provide contiguous coverage over a service area is important. The geometry of the cells vary from relatively circular at the SPP to elliptical further away. This is based on the radiation pattern of the antenna in the HAP [62]. Its shown in Chapter 4 that an estimated 265 cells are required to provide contiguous coverage over a service area of 60 km radius for instance. The result is based on a 1600 element planar phased array antenna with transmit power of 2 W, 2.1 GHz carrier frequency and using a free-space path loss with log-normally distributed random shadowing propagation model. The antenna elements are spaced at half-wavelength intervals apart and the estimated area of the array is under 15 m^2. Assuming a minimum RF output power for each cell of 2 W and a PA efficiency of 47%, the power consumed by the access network of the communication system is computed using (3.31). Considering that in [135, 136] it is estimated that the PA consumes about 57% of the total consumption, the contribution of the PA is conservatively assumed to be 50%, i.e. γ = 0.5. The total power consumption of the BS payload estimated using (3.31) is $\approx$2.3 kW. Therefore, the 60 kW value obtained using the EARTH model presents a hard upper limit of power consumption. Realistically, the power consumption for the stated scenario is more likely to be much less at around the 2.3 kW range than the 60 kW range. However, this is directly dependent on the contribution of the power amplifier to the total BS power consumption.

3.4. Case Study Platforms

An analysis of the possible energy collection and consumption for different solar-powered HAPs is presented in this section. This is based on the year-end solstice (i.e. 21st December, day = 355), which is the shortest day in the northern hemisphere representing the worst case scenario. The following analysis is based on a worst case scenario where the batteries supporting the platform at the start of the year-end solstice have zero charge. Therefore, enough solar energy needs to be harvested to support the platform and payload throughout the day considering two locations - Enugu, Nigeria (Lat: 6.60°N) and York, United Kingdom (Lat: 53.96°N). These represent equatorial African and northern European scenarios. Focusing on the current state-of-the-art in solar powered HAPs, the analysis considers some example solar-powered aircraft-based HAP platforms with wingspans ranging between 25–35 m, which represent only

the aircraft in the development stages close to commercial production. Some relevant parameters of the example platforms considered are highlighted in Table 3.1.

Table 3.1: Case study examples of solar-powered high altitude platforms [11, 12].

Wingspan (m)	**Platform mass** (kg)	**Payload mass** (kg)	**Example platforms**
25	75	5	Zephyr-S
33	140	20	Zephyr-T
35	150	15	PHASA-35

3.4.1. Energy Collection

Consider the earlier Zephyr-6, its wing aspect ratio AR = 9.5 with a wingspan S_w = 18.28 m and chord k_{cd} = 2.13 m. Due to the absence of public or published data on the representative example platforms, the chords of the 25 m, 33 m and 35 m wingspan aircraft are extrapolated linearly based on the known Zephyr-6 information. This is logical since extrapolation is a valid method of predicting behaviour beyond that of a particular UAV model [138]. Thus, the parameters of the example platforms are extrapolated using the following expression.

$$K_{cd}^{i} = \frac{2.13 S_{w}^{i}}{18.28}, \tag{3.32}$$

where $S_{w}^{i} \in \{25,\ 33,\ 35\}$.

Using the (3.32), the estimated chords of the example platforms are approximately 2.9 m, 3.9 m and 4.1 m for the 25 m, 33 m, and 35 m wingspan platforms respectively. The wing surface area is then approximated as the product of chord and wingspan, i.e $A_w = S_w^i K_{cd}^i$. Thus, the surface area for the platforms with different wingspans are estimated as given Table 3.2,

Table 3.2: Example platform wing areas.

Wingspan (m)	**Wing Area** (m^2)	**Allowance** (m^2)	**Solar Area** (m^2)
25	73	3	76
33	127	6	133
35	143	0	143

Wingspan extrapolation with (3.32) does not take into account the additional surface area in the tail stabiliser of the platform where solar cells can also be deployed. Therefore, an allowance of 3 m^2 per tail area is added, considering the characteristics of the example platforms, to obtain the total area available for solar panel deployment. The current state-of-the-art in solar cells is an ultra-thin Gallium Arsenide solar cell with a reported power conversion efficiency of 37.5%

and specific power of 3 kW/kg [139]. Using the proposed energy collection model given in (3.14) and assuming a solar cell efficiency of 37.5%, the total energy collected per wingspan on the 21st of December at Enugu and York are evaluated and presented in Figure 3.10. The total solar irradiance per unit area for York and Enugu used in the evaluations are 1.5 kWh/m^2 and 10 kWh/m^2 respectively with the values obtained from Figure 3.7.

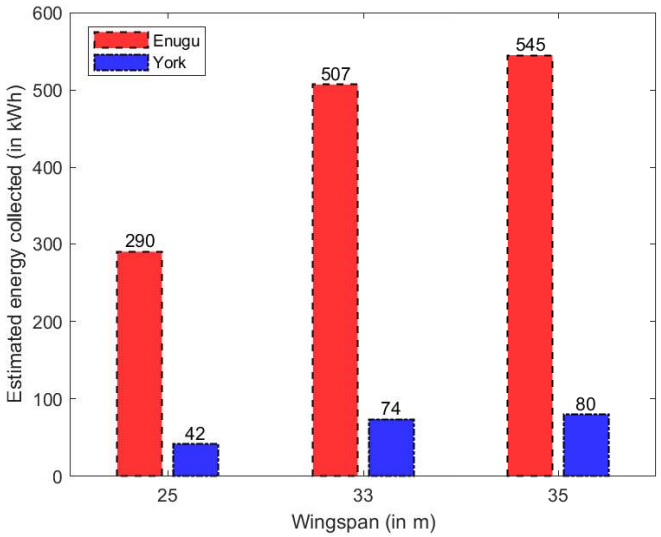

Figure 3.10: Solar energy collected by platforms with different wingspans and at different locations on the year-end solstice.

Notice the significant difference in the solar irradiance between Enugu and York. Typically, this conforms with the expectation of receiving lesser solar irradiance when moving further away from the equator and towards the North Pole. Combining (3.14), (3.21) and (3.31), an inequality that must be satisfied in order to overcome the energy constraint and close the energy budget for a platform with a particular wingspan is expressed as follows.

$$E_{Tc} \geq (P_{req} + P_{Pl} + P_{Av})\, T. \tag{3.33}$$

Expression (3.33) states that the total energy E_{Tc} available to the HAP must be greater than or equal to the product of the HAP operation duration T and the sum of the power required to maintain the platform in flight P_{req}, power consumption of the payload P_{Pl} and avionics P_{Av}. Rewriting (3.33), and substituting (3.21) and (3.31), the following expression is obtained.

$$\frac{\eta_{pv} A_{pv} G_{Ti}}{T} - \frac{C P_{out}}{\gamma \eta_{PA}} - P_{Av} \geq \left(\frac{C_D}{\eta_{prop} C_L^{\frac{3}{2}}} \right) \sqrt{\frac{2F_w^3}{\rho A_w}}. \tag{3.34}$$

It is assumed that the area of solar cell on the platform A_{pv} is equal to the area of the wing A_w. Hence, the area of the wing can be expressed as $S_w k_{cd}$. Therefore, (3.34) becomes

$$\frac{\gamma \eta_{PA} \eta_{pv} S_w k_{cd} G_{Ti} - TCP_{out} - \gamma T \eta_{PA} P_{Av}}{\gamma T \eta_{PA}} \geq \left(\frac{C_D}{\eta_{prop} C_L^{\frac{3}{2}}} \right) \sqrt{\frac{2F_w^3}{\rho S_w k_{cd}}}. \tag{3.35}$$

The resulting expression (3.35) is complex with numerous inter-related variables, however, ensuring a closed energy budget for the platform requires a solution to the expression. Considering the wingspan, among other variables, the longest wingspan that forms part of the solution of the equation is the most desirable. This is because the longer the wingspan, the bigger the wing area and hence the area available for solar cells, which will result in more energy being collected. Additionally, more payload can also be supported. Any chosen method of solving (3.35) must consider the inter-relationship between the variables of the equation. Considering the left hand side (l.h.s) and right hand side (r.h.s) of the inequality as different functions of S_w, the l.h.s becomes a linear function of S_w while the r.h.s becomes a cubic function of S_w. The most desirable wingspan is the highest wingspan at a point of intersection of the two functions for a particular set of variables. A point of intersection is a point where the inequality condition is satisfied. Figure 3.11 presents an indicative plot of the functions showing one point of intersection and the corresponding wingspan.

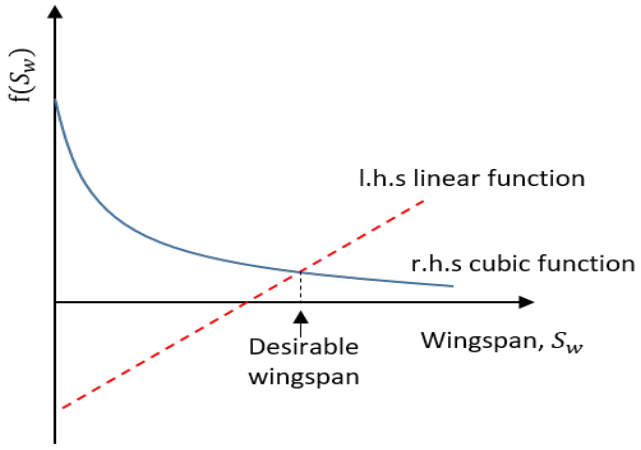

Figure 3.11: Indicative plot of the functions of HAP wingspan. The desirable wingspan that satisfies the energy requirements of the HAP is determined at the equal point of the functions.

3.4.2. Energy Consumption

The following subsections provide estimates of the energy consumption for propulsion, wireless communication and avionics payload respectively.

Propulsion

In order to estimate the power required by each of the platforms using the power equation, the lift and drag coefficients are necessary. The lift coefficient is estimated building from (3.16) as

$$C_L = \frac{2F_w}{\rho V^2 A_w} = \frac{2m_t g}{\rho V^2 A_w}, \tag{3.36}$$

Assuming the true airspeed V = 20 m/s, air density ρ = 0.08891 kg/m^3 at 20 km altitude [140] and g = 9.8 m/s^2, C_L is 0.54, 0.60, and 0.57 for typical aircraft with wingspans of 25 m, 33 m and 35 m respectively.

Wireless deployment of cells from different aircraft flying at different airspeeds over an area may differ as a result of cell displacement due to the airspeed difference. Thus, the comparative analysis of the wireless communication system power consumption becomes more complex. Consequently, a constant airspeed approach is used in this chapter for simplified analysis and to facilitate a direct comparison between the different aircraft. However, the example platform designs may have different true airspeeds in practice, but it is thought that HAPs are intended to be quasi-stationary, hence, the constant speed approach is plausible. Drag coefficient values are obtained using a drag polar, which is a tabular or graphical relationship between lift and drag coefficients for different aerofoils. All platforms under study use laminar aerofoils. In [141], the drag polar of two typical laminar aerofoils are shown. Drag coefficients corresponding to the lift coefficients of the 25 m, 33 m and 35 m wingspan aircraft evaluated above are estimated from the drag polar to be approximately 0.0070, 0.0075 and 0.0071 respectively. The power required by the example platforms is estimated using (3.21) and given in Table 3.3. The estimated values are realistic for the lighter weight platforms given that a 300 kg UAVs are estimated to typically need about 1.5 kW to fly [138].

Table 3.3: Estimated parameters of platforms with different wingspans.

Wingspan (m)	Wing Area (m^2)	Power P_{req} (W)	Lift Coefficient C_L	Drag Coefficient C_D
25	73	243	0.54	0.0070
33	127	429	0.60	0.0075
35	143	459	0.57	0.0071

The power required to maintain a steady flight at a bank angle for the different platforms obtained using (3.27) is shown in Figure 3.12. The figure shows that for steady flights at bank angles less than 30°, the power required is close to that required for a steady horizontal flight. Assuming the Helinet and ITU-R proposed position cylinders of radius 4 km and 0.61 km for

HAPs [62], the bank angle required to maintain an aircraft within the cylinders are about 1° and 4° respectively. At these bank angles, which can be obtained using (3.24), it is observed in Figure 3.12 that the difference in the required power compared with the power required for steady level flight is negligible. Thus, all calculations and analysis also apply to HAPs flying in the defined position cylinders. From (3.27),

$$\lim_{\alpha \to 90} \left(\left(\frac{1}{\cos(\alpha)} \right)^2 P_{req}^c \right) = \infty \tag{3.37}$$

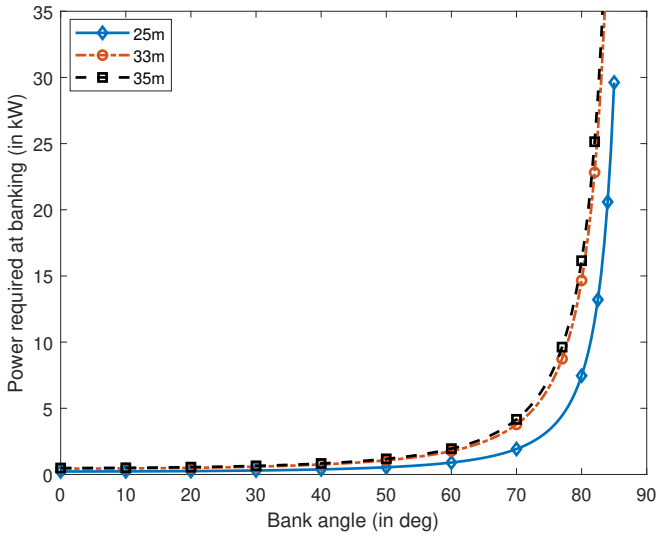

Figure 3.12: Power required for steady banked flights at different banking angles the for example platform wingspans.

Communication and Avionics

Considering a service area radius of 60 km, it is determined through heuristic evaluations, discussed in more detail in Chapter 4, that contiguous coverage can be provided by deploying a minimum of 265 cells given certain propagation and antenna characteristics as highlighted in Section 3.3.2. If each cell transmits a maximum of 2 W with PA efficiency of 50%, the power consumption by the communication payload can be estimated using (3.30) as approximately 2.3 kW. If a smaller service area radius of about 30 km is considered, contiguous coverage can be provided by deploying at least 187 cells using similar assumptions for the 60 km radius service area case. Therefore, the power consumption for a 30 km radius service area is estimated to be approximately 1.6 kW using (3.30).

The consumption by the avionics system can be estimated using parameters in [142] where the avionics mass-to-power ratio is given as 6 W/kg and weight of the avionic sub-system is

22 kg. The approximate total avionics power consumption P_{Av} = 132 W for an estimated 22 kg avionic system.

3.5. Energy Analysis

The total power consumption, $P_{tot} = P_{req} + P_{Av} + P_{Pl}$ of the different platforms and payload is evaluated and the result given in Figure 3.13. The feasibility analysis of the case study platforms in terms of energy harvesting and consumption based on the proposed models are discussed in two categories. The first highlights the feasibility of wireless communication service provisioning in a conservative coverage area of 30 km radius state-of-the-art while the other underscores the feasibility in an extended coverage scenario.

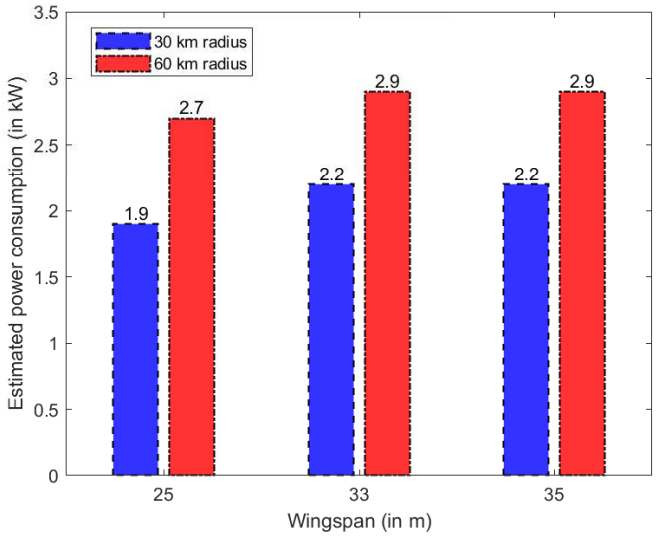

Figure 3.13: Estimated power consumption of the platforms with different coverage radius.

3.5.1. 30 km Radius Service Area

Most HAP studies [39, 53] consider a service area of 30 km radius. In order to conform to the majority of the studies, a service area of 30 km radius is also considered here. The total energy required for a 24-hour continuous operation of the HAP system at full capacity (evaluated using TP_{tot}, where $T = 24$) is approximately 46 kWh for 25 m wingspan HAP, and 53 kWh for both the 33 m and 35 m wingspan HAP. The total available energy in York, which is evaluated using (3.14) and presented in Figure 3.10, is approximately 42 kWh, 74 kWh and 80 kWh respectively. Clearly, the energy harvested by the 25 m platform is below its requirement, while the 33 m and 35 m platforms can potentially harvest sufficient energy to support their operation. Figures 3.14 and 3.15 show the estimated total harvested energy by

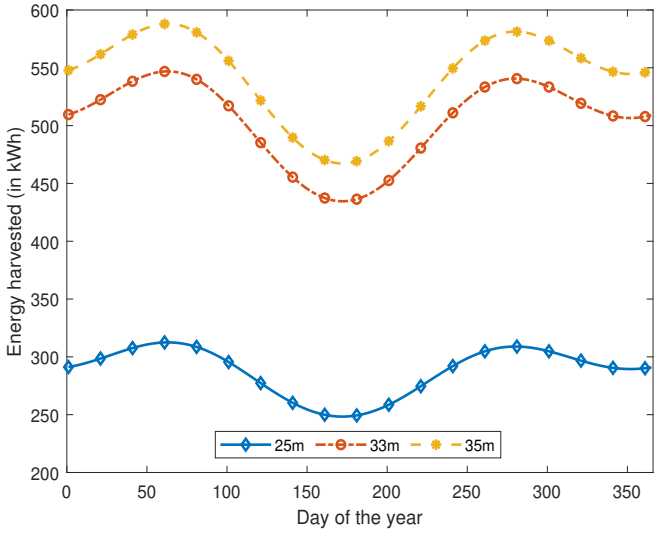

Figure 3.14: Estimated total harvested energy for different platform wingspans – Enugu (Lat: 6.6°N).

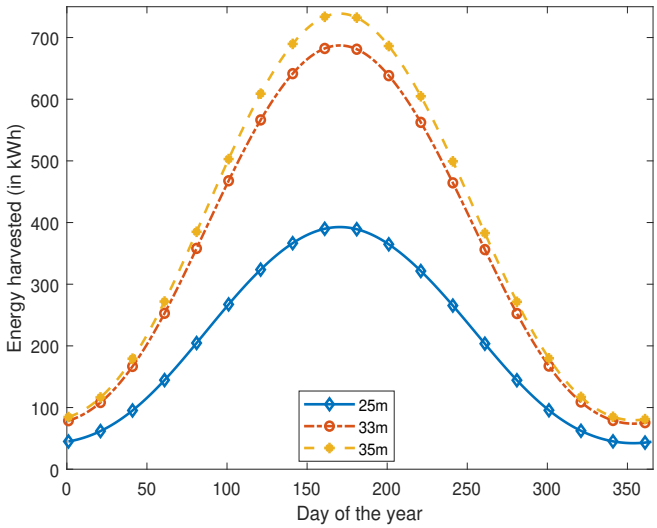

Figure 3.15: Estimated total harvested energy for different platform wingspans – York (Lat: 53.96°N).

the example platforms of varying wingspans (thus different solar cell area) in Enugu and York respectively. Clearly, enough energy is generated in equatorial regions for an unconstrained operation. However, energy management is needed to operate in regions further up the northern hemisphere. Considering the minimum 9 kWh, 13.5 kWh and 14.2 kWh (evaluated using $T(P_{req} + P_{Av})$, where $T = 24$) required energy for continuous flight for a day by the respective platforms, the total energy harvested by the platforms is sufficient to at least keep them in flight. However, 24-hour wireless communications from a HAP over a service area of 30 km radius is not feasible with the 25 m platform due to the insufficient energy available to the communication payload. However, considering the 1.6 kW payload power consumption, sufficient

energy is available to provide around 20 hours of wireless communications with some energy management techniques.

On the other hand, configuring the 25 m platform to operate at full capacity for half of the day and half capacity for the second half of the day by gliding can increase the energy available to the communication payload. Consequently, the total energy required for continuous flight of the 25 m HAP can be reduced to 6.4 kWh (evaluated using $\frac{3T}{2}(P_{req} + P_{Av})$, where $T = 12$). Energy can be saved by allowing the HAP to increase its altitude slowly during the day while harvesting energy, and gliding at night to minimise energy consumption [120] for instance. However, this comes at a cost of increased system complexity due to the platform motion. This configuration allows for an increase in the duration of wireless communication from the 25 m wingspan HAP to around 22 hours. The use of HFC to complement the available solar energy is an alternative to improve the available power to both platform and payload, thereby increasing the duration of operation. A conventional HFC can be used with sufficient hydrogen fuel to last through a required period as shown in Figure 3.16. Practically, the HFC is needed for only a few days or weeks of the year when the total solar irradiance is at the minimum. Note that the power consumption of the backhaul network has been neglected in the analysis thus far because communications payload consumption is significantly dominated by the fronthaul. However, backhaul power consumption should be considered for realistic estimations in practical deployments. The HAP operator may decide how best to deploy the HAP system to provide services for a determined duration of time per day. Considering the

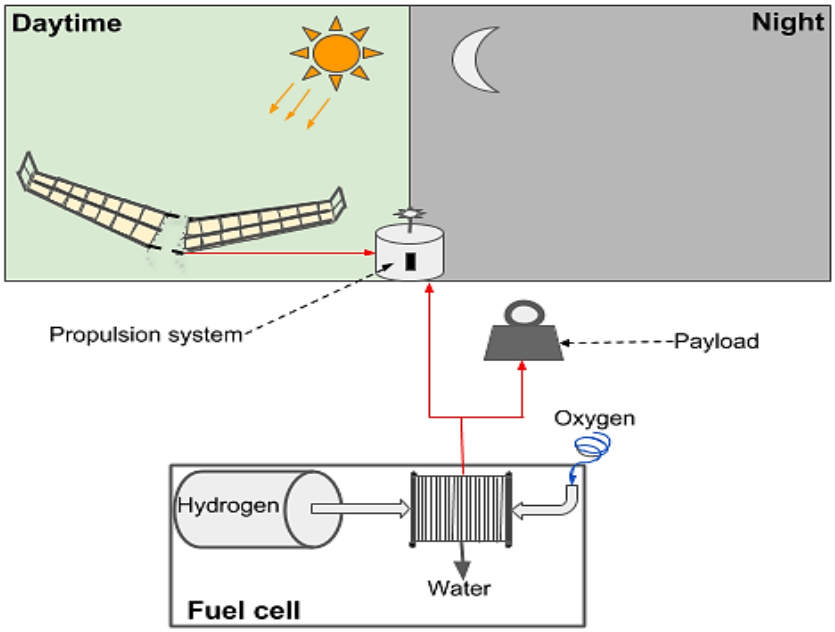

Figure 3.16: Proposed energy subsystem using hydrogen fuel cells.

33 m and 35 m platforms, continuous 24-hour full operation is borderline possible because of the availability of sufficient energy. However, any increase in power consumption by the subsystems can quickly make these platforms infeasible as alternative source of energy may be needed. Although most of the popular HAPs are solar-powered, designers are now beginning to look at hydrogen-powered platforms such as presented in [7]. The downside is that they will operate for shorter mission duration due to the need to bring down the platform for hydrogen refill, which increases operational cost compared with a solar-powered platforms.

Apart from energy availability, the feasibility of a platform for wireless communication is further limited by the maximum payload weight it can support. This is important when considering the weight of the energy storage and other subsystems for the HAP mission. As highlighted earlier, the hard constraint here is the available payload weight allowance on the HAP. A typical specific energy of hydrogen storage tank varies between 1–2 kWh/kg [143–145] with the current state-of-the-art. The estimated payload energy consumption in 24 hours is approximately 38 kWh (evaluated using TP_{Pl}, where $T = 24$). Therefore, using a storage system with 2 kWh/kg specific energy based on the current state-of-the-art, the minimum weights for the energy storage system payload needed for 24-hour continuous full-capacity HAP operation is 19 kg (i.e. $\frac{38\ kWh}{2\ kWh/kg}$). Hence, perpetual delivery of wireless communications is not possible using any of the example HAPs. However, the possibility of a high power fuel cell with specific energy of >3.3 kWh/kg is shown by Eickhoff and Klein in their patent [146]. Futuristically, assuming the proposed HFC by Eickhoff and Klein with specific energy of 3.3 kWh/kg can be commercialised by overcoming its oxygen accumulation problem, the energy storage system weight becomes approximately 11.5 kg (i.e. $\frac{38\ kWh}{3.3\ kWh/kg}$). Considering the weight of other necessary payloads, in both cases the total payload weight limit of the 25 m wingspan platform, as given in Table 1, is exceeded. Therefore, the 25 m wingspan platform is not suitable for contiguous wireless communications coverage over the given service area. However, if the proposed HFC by Eickhoff and Klein is used, the 33 m and 35 m wingspan platforms can certainly support the 11.5 kg energy storage system with 8.5 kg and 3.5 kg available for other payload respectively based on Table 3.1. The 33 m and 35 m wingspan platforms with less payload constraints are certainly feasible in both Enugu and York considering their significantly higher available solar energy.

3.5.2. 60 km Radius Service Area

Due to the increased energy consumption resulting from extending the service area, the 25 m wingspan HAP can potentially harvest enough energy to maintain operation for about 15 hours

in York. This is based on the 42 kWh generated energy and 2.7 kWh total energy consumption as highlighted in Figure 3.10 and Figure 3.13 respectively. It is shown in the previous section that the 25 m wingspan platform is not suitable for contiguous coverage from a HAP over a 30 km radius area mostly due to its weight limitations. Therefore, the analysis in this section largely focuses on the 33 m and 35 m wingspan platforms for an extended wireless coverage significantly beyond an area of 30 km radius. Based on the HAP cell deployments in described in Chapter 4, wireless service provisioning using the 33 m and 35 m wingspan platforms in an area of 60 km radius for a whole day requires approximately 70 kWh of energy (evaluated using TP_{tot}, where $T = 24$). In both Enugu and York, sufficient energy is available to the platforms for 24-hour continuous wireless communications service delivery as highlighted in Figure 3.10. This confirms the feasibility of the 33 m and 35 m platforms subject to overcoming the weight constraint.

Assuming half the energy is expended for daytime operation, the energy storage needed for night-time operation must have the capacity of storing approximately 35 kWh (i.e. $\frac{70\ kWh}{2}$) of energy for both platforms. With the HFC developed by Eickhoff and Klein, HAP extended coverage in equatorial regions is potentially feasible in the near future subject to the weight of other necessary subsystems. The feasibility of the 25 m wingspan platform is power and weight limited in the United Kingdom and other places further up northern hemisphere while it is mostly only weight constrained in equatorial regions. Conversely, the 33 m and 35 m wingspan platforms are potentially feasible in terms of energy storage weight especially if storage systems with high specific energy are used. However, the weight of other payloads like the access and backhaul subsystems may affect the feasibility of both platforms. Realistically, it is difficult to overcome the limitations of both energy and weight requirements for continuous whole-day operations on the 21st of December with the currently available platforms.

Nevertheless, with improvements in HAP design such as wingspan, these limitations can be overcome in the future. Certainly, airships with higher payload and power allowance present a viable alternative for all scenarios discussed. Significant development in aeronautics will expectedly allow for the development of platforms with higher form factor. Thus, the capability of lifting heavier payloads may be realised promptly with continuous development in aeronautic technology. For instance, Boeing 737 (B-737) aircraft has a wingspan of 35.79 m and can lift a maximum of 186 passengers. The Boeing 747 (B-747) with higher form factor has a wingspan of 60 m, lifting a maximum of 660 passengers. The ratio of the wingspan of B-747 to B-737 is approximately 1.68. Similarly, the ratio of B-747 payload to B-737 is approximately 3.55. Hence, the payload increase factor is approximately 5.96 (i.e. 1.68 x

3.55). Therefore, doubling the wingspan for instance can allow for up to six times the amount of payload an aircraft can support (see (3.17) for the relationship between wingspan and lift force for any given aircraft). Doubling the wingspan of PHASA-35 to get 70 m for instance could probably enable the new aircraft to lift 90 kg worth of payload. It is worth noting that increasing wingspan of an aircraft is complicated as a result of the size, weight and power phenomenon. However, a longer wingspan is aerodynamically achievable as demonstrated by the NASA Pathfinder and Centurion aircraft. NASA Pathfinder, a solar-powered aircraft with a wingspan of 29.5 m supports a payload of 45 kg. On the other hand, Centurion with double the wingspan at 61.8 m supports a payload of up to 270 kg, which is exactly six times that of Pathfinder. Unfortunately, even though Pathfinder only flew for a few hours at high altitude [147] and Centurion only achieved low altitude validation flights [148], the fact remains that longer wingspan results in higher payload support. This is achievable as demonstrated by these platforms and their limitations can be overcome with advancements in aeronautical engineering, which is beyond what was obtainable years ago. An interesting platform under development specially for wireless communications is the 78 m wingspan solar-powered Hawk30 with potential large enough power and payload capability to overcome the constraints discussed above [37]. Similarly, a hydrogen-powered HAP [7] under development has a 60 m wingspan and can potentially support up to 140 kg of payload proving 20 kW of power. With the large wingspans of these platforms, they can overcome both the power and weight constraints and increase the feasibility of HAPs for wireless communication in line with the above analysis. Although these platforms show potential, commercial production is unlikely to happen within the next 5 years. Apart from increasing form factor of aircraft, network splitting, can be used to minimise energy consumption and payload weight by having different hardware located either on the aircraft or on the ground by using Cloud-based RAN functional splitting, albeit at the cost of an increased backhaul power consumption. Cloud RAN systems adapt network functional splitting by separating the functionality of the baseband and radio units (RU) [149].

3.6. Conclusion

In this chapter, the energy management of solar-powered aircraft-based HAP for wireless communications is studied as a way of placing later work in context. Models for estimating energy collection and consumption in a HAP have been proposed. It has been shown that with the current platforms, having wingspans ranging between 25–35 m, wireless communications can be provided from a HAP for a limited duration of 15–24 hours per day depending on the platform,

latitude of operation, operation configuration, and radius of coverage with appropriate energy management. Continuous operation at full capacity without energy management, where the HAP operates at reduced capacity, is infeasible with the example platforms particularly at latitudes significantly away from the equator. It is also infeasible even at locations closer to the equator considering the extended 60 km service area radius.

However, for continuous operation for a whole day using platforms with wingspan less than 35 m, feasibility can be enhanced by the use of HFC systems to supplement the energy collected by the solar cells in order to overcome the energy and weight constraints. Different capabilities and configurations of HAPs are needed for operations in different locations considering the latitude dependent solar irradiance, which determines the available energy of a platform at any given latitude. For latitudes further away from the equator and extended service area operability, airships or hydrogen-powered HAPs will be needed for feasibility. Although the feasibility of current platforms is both energy and power limited, it is suggested that doubling the wingspan of aircraft will provide sufficient gains in both payload carrying capability by up to a factor of 6 and energy collection since surface area for solar cells increases. This will significantly enhance feasibility and facilitate deployment. Interestingly, stakeholders are already looking at this with the over 70 m wingspan Hawk30 HAP under development. Furthermore, using communication techniques like Cloud-based radio access network and functional splitting proposed for next generation wireless systems can increase the feasibility of solar-powered aircraft-based HAPs for wireless communications since fronthaul payload power consumption is minimised, although with a marginal increase in backhaul payload power consumption.

Chapter 4. Delivering Extended Cellular Coverage and Capacity Using High Altitude Platforms

Contents

4.1. Motivation

Having discussed the feasibility of using solar-powered HAPs, which is the most common and technologically matured, for long endurance wireless communications missions in Chapter 3, this chapter proposes an actual beam deployment technique for cellular wireless communication service delivery over an extended coverage area. In addition, mathematical models that can be used for estimating the limits of coverage extensions are also derived. Considering the good propagation environment of HAPs, and that a significant proportion of the world's population is still unconnected [150], extending the coverage achievable using a HAP is key to increasing connectivity and maximising utility especially in areas with low user densities. A number of studies have investigated the deployment of HAP beams for cellular communications. For instance, in [54, 56], HAP cellular footprints are described mathematically as functions of antenna beamwidth, elevation and azimuth angles. In [6,57], the use of intelligent beamforming strategies for HAP coverage and capacity is investigated and a beam-pointing scheme based on k-means clustering proposed. However, these proposed techniques are only

valid within an area of around 30 km radius due to increased inaccuracies and approximation errors at extended distances, which is common in most of the other prior studies.

This chapter proposes an improved approach to beam deployment using a HAP for contiguous cellular coverage over an extended coverage area, which is significantly beyond the 30 km radius area common in the literature, and analysing the achievable capacity of such extended coverage system. It also provides a detailed theoretical framework and performance analysis of the proposed approach based on simulations and derived models. The main challenge of forming HAP beams in an extended coverage scenario is the significant broadening of beams especially at low elevation angles, which results in significant cell overlap due to the disproportionately sized beam projections on the ground and affects coverage and capacity of the HAP system. Effectively, beam broadening increases the ellipticity of the cells further away from the centre of the service area. To address this challenge, a recursive beam-pointing and tessellation algorithm, which forms cells over an extended coverage area using multiple beams from the HAP, is proposed. The algorithm compensates for broadening, starting from the sub-platform point cell (SPPC) and provides flexibility on cell size variations and the level of overlap needed between the cells. In Chapters 5 and 6, the proposed approach is used to deploy a tier-based HAP architecture to provide service for fixed and variable user densities respectively.

4.2. HAP Beam-pointing

A HAP communication system forms beams on the ground at a given elevation angle to deliver both coverage and capacity to users [54]. These beams are used to form cells that are isolated by the HAP antenna radiation pattern [62], and are limited by the beamforming techniques implemented, which affects the achievable coverage and capacity. Ideally, each antenna beam delivers uniform illumination to its corresponding cell, ensuring through a steep roll-off that no power is detected outside the cell boundaries. Unfortunately, this is impractical due to the imperfect roll-off of practical antenna beams, which introduces inter-cell interference (ICI) worsened by beamforming limitations [71], especially with improper beam-pointing. Improving coverage and capacity while minimising ICI requires appropriate beam-pointing and overlap control with an adequate antenna system and characteristics.

4.2.1. System Model

The HAP extended coverage system can be used to provide wireless coverage and capacity over a significantly wide area, and it can also be used to provide umbrella cells for other wire-

less systems. This section introduces the HAP extended coverage system architecture and highlights the important parameters and models for efficient beam-pointing using the proposed algorithm. The metrics for performance analysis are also presented. Subsequently, all discussions and analysis throughout this thesis are based on the HAP downlink.

Beam Deployment

Consider a quasi-stationary HAP located at an altitude h_p, and at the centre of a service area of radius R. The HAP supports a radio unit (RU) with a multi-element uniform planar phased array antenna of $M \times N$ antenna elements. Multiple beams are formed from the antenna unit such that the footprints of the main lobes provide coverage to a set of users $\mathcal{I}$ within the service area as shown in Figure 4.1. These beams, which are pointed at a set of coordinates $\mathcal{C}$, such that the resulting footprints produce a regular tessellated structure of contiguous cells $c \in \mathcal{C}$ over an extended HAP service area. Each user associates to a cell that maximises its signal quality. The HAP transmit antenna gain profile G_i^t for signal quality evaluation, as observed by user $i \in \mathcal{I}$, is given as follows [5],

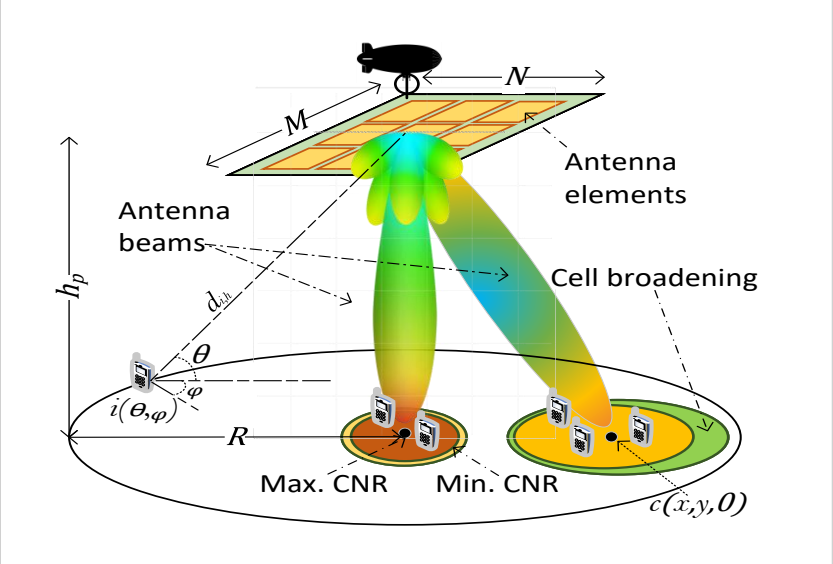

Figure 4.1: HAP phased array antenna beamforming for cellular coverage.

$$G_i^t = g_e AF_i^x AF_i^y, \tag{4.1}$$

where g_e is the gain of an isotropic antenna element, array factors AF_i^x and AF_i^y are given as

$$AF_i^x = \sum_{n=1}^{N} I_{n1} e^{j(n-1)(kd^x \sin(\theta_i)\cos(\varphi_i)+\beta_i^x)}, \tag{4.2}$$

$$AF_i^y = \sum_{m=1}^{M} I_{1m} e^{j(m-1)(kd^y \sin(\theta_i)\sin(\varphi_i)+\beta_i^y)}, \tag{4.3}$$

where angular wave number $k = \frac{2\pi}{\lambda_s}$, λ_s is the wavelength, d^x and d^y are the inter-element spacings in the x- and y-axes of the antenna array with array factors AF_i^x and AF_i^y at user i respectively. I_{nj} and I_{jm} represent the excitation amplitudes of the antenna elements as depicted by Figure 4.2, θ_i and φ_i define user elevation and azimuth angles as seen by the HAP. $\beta_x = -kd^x \sin(\theta_i^0)\cos(\varphi_i^0)$ and $\beta_y = -kd^y \sin(\theta_i^0)\sin(\varphi_i^0)$ are phase shifts with θ_i^0 and φ_i^0 being the boresight elevation and azimuth angles respectively.

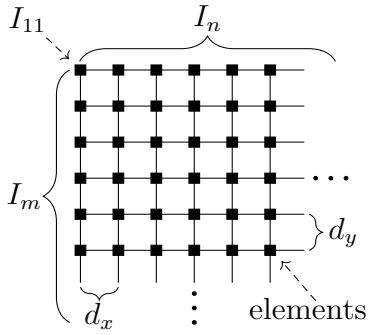

Figure 4.2: Antenna element excitation for an $M \times N$ antenna array. Elements in rows and columns are referred to as $x-$ and $y-$ axes elements with d_x and d_y distances apart respectively. There is proportionality between the excitation amplitudes of the elements in both x, y axes. The $(m, n)^{th}$ element excitation amplitude is expressed as $I_{mn} = I_{mj} I_{jn}$ [5].

Coordinates of the multiple beam boresight points on the ground, which are obtained as a set of cell coordinates $\mathcal{C}$ from the proposed algorithm and then converted into their corresponding elevation and azimuth angles of the HAP from the ground, are supplied to the beamformer implementing (4.1) to obtain the distribution of HAP antenna transmit gain on the ground. The beams formed are then used to create HAP cells on the ground. In defining a cell in this thesis, let $\tilde{\mathcal{A}}_b$ be the footprint of the b^{th} beam on the ground, p be any interior point in $\tilde{\mathcal{A}}_b$ and Γ_p be the carrier-to-noise ratio (CNR) at point p, the cell c is a bounded region around the beam boresight with boundary $\partial c := p \in \tilde{\mathcal{A}}_b \ : \ \forall p, \Gamma_p \geq 9$ dB. The 9 dB value is a global system for mobile communications (GSM) standard for cell delineation [151], which ensures that every user in a cell can receive signal with good quality. Γ_i for user i is evaluated as

$$\Gamma_i = \frac{P_i^t G_i^r G_i^t}{N L_{i,h}}, \tag{4.4}$$

where P_i^t is the HAP transmit power, G_r^i and N are the user receive antenna gain and noise power respectively. The path loss $L_{i,h}$ between user i and the HAP h is modelled as free-

space path loss with log-normally distributed fading due to shadowing [6], which follows the 3rd generation partnership project (3GPP) non-terrestrial network (NTN) channel model [80] and allows for a realistic large-scale representation of the HAP propagation channel. $L_{i,h}$ is expressed as

$$L_{i,h} = \left(\frac{4\pi d_{i,h} f}{v}\right)^2 X_\sigma, \tag{4.5}$$

where $d_{i,h}$ is the slant distance between user i and the HAP h in km, f is the carrier frequency in GHz, v is the speed of light in m/s and X_σ is a log-normally distributed random variable with zero mean and standard deviation σ_x of 4 dB, representing fading due to shadowing [152]. Small-scale fading is not considered in this line-of-sight (LoS) scenario since the focus is on cellular structure in general and the long term mutual interference effects of the cells on each other. It is also because the HAP is assumed quasi-stationary and users are fixed, therefore, no small-scale fading is experienced. This is validated by results from a practical HAPS flight test reported in [26].

Cell pointing to achieve good tessellation depends initially on the radius of the SPPC (i.e. the HAP antenna broadside cell), which is defined by the angle ρ subtended by the edge of the SPPC to a plane vertical to the HAP. This depends on the antenna mainlobe beamwidth and the minimum CNR Γ requirement at the edge of the SPPC. Hann windowing described in Section 2.4.2 of Chapter 2 is used in generating weights for the beamforming due to its capability of suppressing sidelobes by as much as 32 dB down from the mainlobe, which reduces ICI. Here, ρ is defined such that at all locations within a cell, $\Gamma_p \geq 9$ dB. The HAP antenna array forms multiple beams, which are pointed at a set of coordinates $\mathcal{C}$, such that the footprints create a regular tessellated structure of contiguous cells over the entire HAP service area. The cell centre coordinates are supplied by the proposed beam-pointing algorithm discussed in detail in Section 4.3. The resulting cells in the extended service area are disproportionate in size and shape due to the effect of broadening mainly at low elevation angles as seen in Section 4.5.

User Distribution

Considering a user density of λ users/km^2, a set of users $\mathcal{I}$ with 2D coordinates are randomly distributed over the HAP service area of A km^2. The users in $\mathcal{I}$ are independently and identically distributed over the space according to a Poisson distributed random variable with mean λA. The number of users is $|\mathcal{I}|$, where $|.|$ denotes cardinality. This follows a bivariate Poisson point process (PPP) $\Phi_p \in \mathbb{R}^2$.

User Association

For evaluating the performance of the proposed model, the distributed users are associated to the cells formed using the coordinates obtained from the proposed algorithm. It is important to note that the user distribution does not affect the tessellation performance of the proposed pointing algorithm. The tessellation is aimed at providing contiguous coverage of cells over the entire service area considering the footprints of the HAP antenna beams. Users are distributed over the service area to assess both the CNR and CINR distribution on the ground resulting from the algorithm's cell tessellation.

The users associate with cells based on their perceived CNR. User i associates to cell c_i where it achieves the highest CNR, which must be at least equal to a given threshold. If a user's CNR is less than the threshold in any of the cells, the user is not connected in that cell. Furthermore, if user j is located in a region where two or more cells overlap, it is assigned to the cell that maximises its received power. Invariably, users in an overlap region can detect all the overlapping cells $\{c_l\}$. The set of coordinates of the overlapping cells $\mathcal{C}_o$ for all users an overlap region is obtained as follows,

$$\mathcal{C}_o = \{c_l \mid l \in \{1, 2, ..., N\}\}, \forall j \mid l \neq j, \tag{4.6}$$

where N is the number of overlapping cells. Then, the set of received powers $\mathcal{P}^j$ for user j is defined in (4.7). User j associates to cell $c_{l^*(j)}$ such that $l^*(j)$ realises the maximum in P_r^l.

$$\mathcal{P}^j = \{P_r^l \mid l \in \{1, 2, ..., N\}\}. \tag{4.7}$$

4.2.2. Performance Metrics

In deciding the cell a user should associate with after the formation of cells, and whether service can be provided at a given performance level, the user's received power, expressed mathematically below, is used. Additionally, other metrics for performance evaluation are presented below. They are also used for evaluations both in Chapters 5 and 6.

1. *CINR γ_i:* For users already associated to cells, their performance with the proposed algorithm is evaluated using CINR γ, which expresses the ratio of their carrier power to both interference and noise. For user i, γ_i is defined as [6, 57]

$$\gamma_i = \frac{P_i^r}{\sum_{j=1}^{J} P_i^j + N_i}, \tag{4.8}$$

where the received power P_i^r of user i from its associated cell expressed as

$$P_i^r = \frac{P_i^t G_i^r G_i^t}{L_{i,h}}, \tag{4.9}$$

and $\sum_{j=1}^{J} P_i^j$ is the summation of the interference power from the other cells $j = 1, 2, 3 \cdots$, $J \mid P_i^r \neq P_i^j$ as a result of the sidelobes of their beams, and $J = |\mathcal{C}| - 1$ is the number of interfering cells for each user.

2. *Throughput T_i:* This is evaluated per user in bits/s/Hz using the Truncated Shannon Bound expression as given below [153].

$$T_i = \begin{cases} 0, & \gamma_i < \gamma_{min} \\ \alpha \log_2(1 + \gamma_i), & \gamma_{min} \leq \gamma_i \leq \gamma_{max} \\ \alpha \log_2(1 + \gamma_{max}), & \gamma_i > \gamma_{max}, \end{cases} \tag{4.10}$$

where $\alpha = 0.65$ is the implementation loss, $\gamma_{min} = 1.8$ dB is the minimum allowed CINR and $\gamma_{max} = 22$ dB is the CINR resulting in the maximum achievable throughput [57]. γ_{min} is assumed for an acceptable LTE signal quality [154], while γ_{max} relates to the LTE signal quality mapping with modulation and coding scheme, where maximum capacity is achieved at around 22 dB CINR [155].

3. *Average Spectral Efficiency η_i^f:* Let the average capacity per user in cell i and the bandwidth allocated to a user in the cell pointed at any given distance be $\tilde{C_{pu}^i}$ (expression to be derived later) and B_i respectively. Thus, $\eta_i^f \triangleq \frac{\tilde{C_{pu}^i}}{B_i}$.

4. *Average Area Spectral Efficiency $\tilde{S_{\eta_f}^i}$:* This is the ratio of average spectral efficiency η_f^i to the area of the cell A_i in bit/s/Hz/km^2 for cell i [156, 157]. Thus, $\tilde{S_{\eta_f}^i} \triangleq \frac{\eta_f^i}{A_i}$.

4.3. Beam-pointing for Extended Coverage

In order to achieve the desired extended coverage and capacity from HAP, an improved beam-pointing algorithm, which gives the total number of cells N_c and a set of their boresight co-ordinates required to deliver contiguous coverage, is proposed. The algorithm explicitly takes

beam broadening and overlap into consideration to achieve good tessellation and minimise ICI of the resulting HAP cellular footprint on the ground. This section provides a detailed discussion of the algorithm and its theoretical framework.

4.3.1. HAP Beam Geometry

The algorithm considers the geometry of a HAP beam footprint on the ground, which is dependent on the HAP antenna profile and varies in size with the pointing angle. Beams pointing away from the sub-platform point (SPP) broaden in size due to the limitations of beamforming at lower elevation angles. The HAP cell geometry is described in Figure 4.3. In this chapter, it is assumed that the beam footprint is elliptical in shape to simplify the mathematical analysis. Figure 4.3 highlights the geometry of the cells, one pointing at the SPP and the others at distances away. Note that the elliptical shape assumption is plausible. Imagine that a beam from the antenna is approximately conical in shape, then cutting through the cone at an angle using a straight surface gives an elliptical footprint on the ground. The effect of the curvature of the earth is ignored here.

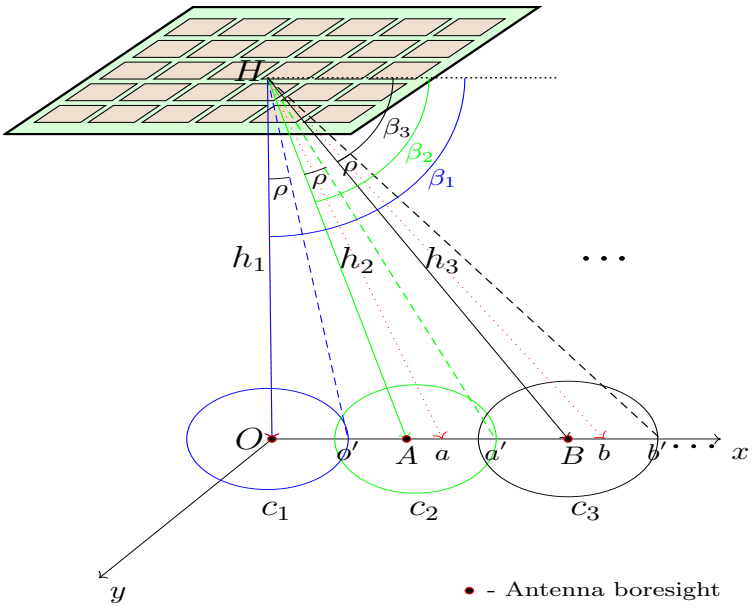

Figure 4.3: The HAP cell geometry. Dotted arrow lines show the initial cell boresight when neighbouring cells only touch each other. Solid arrow lines show the new boresight after adjusting the initial pointing angle to add overlap between neighbouring cells. The angle between the solid and dashed lines for each cell highlight the angle ρ between the boresight and cell edge, which is constant for all cells.

4.3.2. Beam-pointing Algorithm

The proposed algorithm involves five steps, which are described as follows:

Step 1: Initially, the boundary of the SPPC c_1 is defined by the angle ρ it subtends at the HAP with its centre as highlighted in Figure 4.3. Also, the look angle of the cell boresight from the HAP horizontal plane β is set to 90°. The angle between the cell centre and boundary is assumed to be equal for all cells. Similarly, the initial path distance d_1 between the HAP and the SPPC centre c_1 is set to h_p. A second cell c_2 points along the x-axis of the SPPC by making the angular distance between the centres of c_1 and c_2 equal to 2ρ and d_1 updated to d_2. A set of the tessellated cell centre coordinates $\mathcal{C}_T$ along the x-axis is maintained and updated for each new coordinate. The update of $\mathcal{C}_T$ for subsequent cell centres $c_{i+1}, \cdots, c_n$ along the x-axis of c_i is carried out using an update function $U^d(\rho, \beta_i)$ which is derived as follows.

In order to derive $U^d(\rho, \beta_i)$, the geometry in Figure 4.4 is obtained from Figure 4.3. The solid lines in Figure 4.4 represent the cell boresight and the dashed lines represent the cell edges. Let $d_1,\ d_2,\ d_3,\ \cdots$ be the distances of cell boresights from the SPP along an axis. The distance of cell boresight $\overline{Oa}$, i.e. $Oa = d_2 - d_1 = d_2$ since $d_1 = 0$.

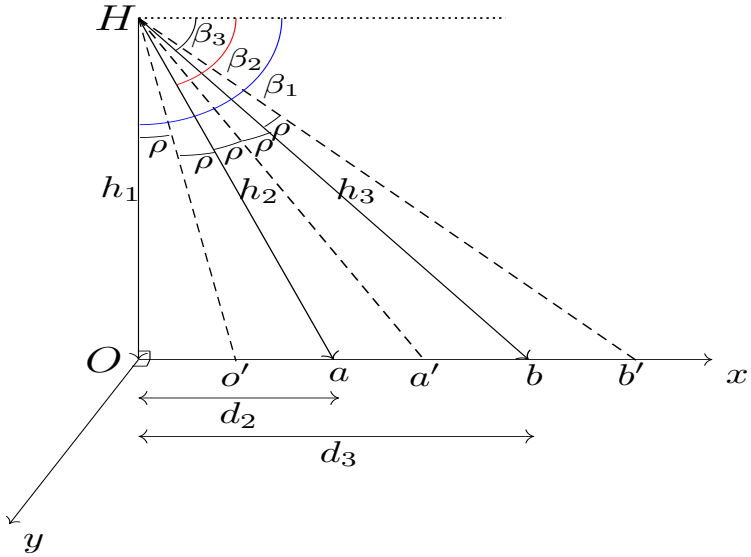

Figure 4.4: The HAP cell boresight geometry.

Using sine rule on ΔOHa,

$$\frac{d_2}{\sin(2\rho)} = \frac{h_1}{\sin(\beta_1 - 2\rho)}. \tag{4.11}$$

Therefore,

$$d_2 = h_1 \left(\frac{\sin(2\rho)}{\sin(\beta_1 - 2\rho)} \right) + d_1. \tag{4.12}$$

Similarly, the distance of cell boresight $d_3 = d_2 + \overline{ab}$. From ΔaHb,

$$\frac{\overline{ab}}{\sin(2\rho)} = \frac{h_2}{\sin(\beta_2 - 2\rho)}. \tag{4.13}$$

Thus,

$$\overline{ab} = h_2 \left(\frac{\sin(2\rho)}{\sin(\beta_2 - 2\rho)} \right). \tag{4.14}$$

This implies that,

$$d_3 = \overline{ab} + d_2 = h_2 \left(\frac{\sin(2\rho)}{\sin(\beta_2 - 2\rho)} \right) + d_2. \tag{4.15}$$

Therefore, a generic equation with a common update function $U^d(\rho, \beta_i)$ can be obtained from (4.12) and (4.15) as

$$d_{i+1} = h_i U^d(\rho, \beta_i) + d_i, \tag{4.16}$$

where,

$$U^d(\rho, \beta_i) = \frac{\sin(2\rho)}{\sin(\beta_i - 2\rho)}. \tag{4.17}$$

Note that $d_1 = 0$ implies that the distance of the SPPC boresight from the SPP is zero and SPPC is the cell at the centre of the HAP service area.

In order to include some level of overlap between c_i and c_{i+1}, as depicted in Figure 4.3, to minimise outage for mobile users moving between cells, the distance between c_i and c_{i+1} is reduced by a distance corresponding to an overlap angle Φ, which is a function of the overlap rate ε and inter-cell distance. The value of Φ between cells c_i and c_{i+1} is derived as follows. Consider c_1 and c_2 in Figure 4.3 with the geometry partially reproduced in Figure 4.5.

Let $\Phi_{i,i+1}$ be the angle of overlap between neighbouring cells c_i and c_{i+1}. Since the coordinate of a is determined based on the size of the SPPC centred at O, h_2 can be evaluated while h_1 is the platform height. Overlap distance $\overline{Aa}$ is determined by a given overlap rate ε and the distance between the neighbouring cell centre boresight from the SPPC boresight. If d_2 is the distance of the neighbour to the SPPC and d_1 is the distance of the SPPC boresight from the SPP, then

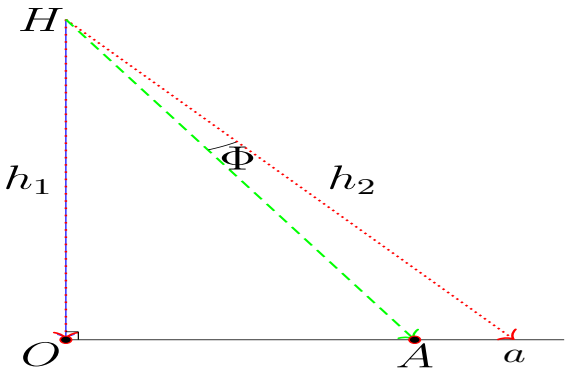

Figure 4.5: The HAP cell overlap geometry.

$$\overline{Aa} = \varepsilon(d_2 - d_1). \tag{4.18}$$

Hence, using sine rule on ΔOHA and ΔOHa,

$$\frac{h_2}{\sin(\frac{\pi}{2})} = \frac{\varepsilon(d_2 - d_1)}{\sin(\Phi_{1,2})}. \tag{4.19}$$

Therefore,

$$\Phi_{1,2} = \sin^{-1}\left(\frac{\varepsilon(d_2 - d_1)}{h_2}\right). \tag{4.20}$$

Extending (4.20) to subsequent cells results in

$$\Phi_{i,i+1} = \sin^{-1}\left(\frac{\varepsilon(d_{i+1} - d_i)}{d_{i+1}}\right). \tag{4.21}$$

This process, as described above and depicted in Figure 4.3, continues for subsequent cells along the x-axis. At the end of this first step, a set of tessellated cell centres $\mathcal{C}_T = \{(0, d_1, 0), (0, d_2, 0), (0, d_3, 0), \cdots, (0, d_n, 0)\}$ is obtained, resulting in the structure shown in Figure 4.6a. Thus, $\mathcal{C}_T$ is defined as follows,

$$\mathcal{C}_T = \{c_i \mid c \in \mathbb{R}^3\} \ \ \forall i = 1,\ 2,\ 3,\ \cdots,\ n. \tag{4.22}$$

Step 2: In the second step, all $c_i \in \mathcal{C}_T$ is rotated by 60° as shown in Figure 4.6b to obtain another set of cell boresight coordinates $\mathcal{C}_R$. The 60° rotation allows for the exploitation of the good tessellation properties of hexagonal geometry. The path vectors between the cells

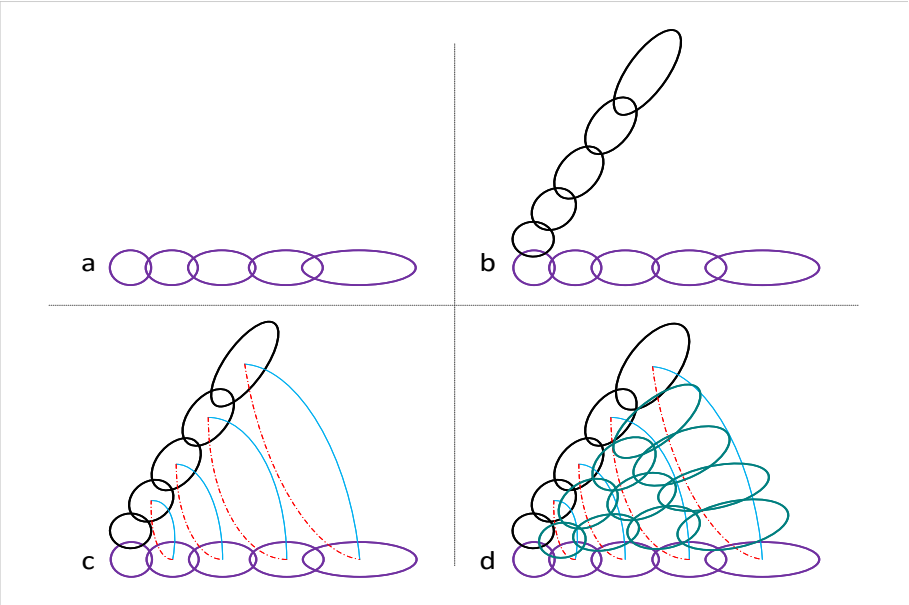

Figure 4.6: The cell tessellation processes. a) The first step with cells pointing at increasing distance on the x-axis. b) The rotation of the first cells from (a) yielding another set of cells in the second step. c) and d) Show the third and fourth steps where new cells are deployed between the structure in b).

in $\mathcal{C}_T(i)$ and their corresponding rotated copies in $\mathcal{C}_R(i)$ are obtained using spherical linear interpolation defined by the following expression [158].

$$V_I = \frac{\sin{(1-s)\theta}}{\sin(\theta)} V_i + \frac{\sin{(\theta s)}}{\sin{(\theta)}} V_f, \tag{4.23}$$

where θ is the angle between corresponding i^{th} cells in $\mathcal{C}_T$ and $\mathcal{C}_R$, V_i and V_f are the start and final vectors of the corresponding cells with reference to the SPPC. While V_I represents the path vector between the cells with s determining the steps in the path.

Steps 3-4: The solid curve between corresponding cell centre coordinates, shown in Figure 4.6c, highlights the paths between the cells and their rotated copies. Ideally, increasing number of cells will be formed azimuthally on the paths depicted by the solid lines between corresponding cells. These paths are divided into $k+1$ equal path lengths. Here, k indicates the position of the paths, shown using the solid lines in Figure 4.6c, on the cellular arc/path starting from $k = 0$ at the SPPC. k new coordinates (equidistant from each other) are introduced starting from arc/path $k = 1$ to the last in the structure. However, introducing new cells azimuthally on the solid arcs/paths results in minimal or no overlap between neighbouring cells towards the SPPC. This is due to azimuth antenna pattern (AAP) distortion, which occurs particularly in aerial systems when beams are formed or steered electronically in azimuth using

a planar phased array antenna [159, 160]. AAP distortion results in steering angle quantisation, grating lobes and main lobe gain reduction [159], which minimises the overlap between neighbouring cells as highlighted earlier. To mitigate against AAP distortion and compensate for the mainlobe gain loss, the new cells are pointed on the mirror images of the paths, which are the dashed convex paths, rather than the original solid concave paths. The mirror image approach, which works by packing the new cells formed azimuthally closer together, is a heuristic and less complex method that ensures proper overlap between neighbouring cells towards the SPPC. The optimality of this heuristic method has not been investigated on this occasion, however, it is shown to enhance the tessellation. A set of the new cell centre coordinates are assigned to $\mathcal{C}_P$ to conclude the third and fourth steps with the resulting cellular structure when cells are pointed at these coordinates shown in Figure 4.6d.

Step 5: Finally, in order to achieve full 360° coverage, the union of the set of cell boresight coordinates $\mathcal{C}_T$, $\mathcal{C}_R$ and $\mathcal{C}_P$, which yields the tessellation in Figure 4.6d, are rotated 5 times. The resulting coordinates from the rotation are contained in another set $\mathcal{C}_F$. In Figure 4.7, a set $\mathcal{C}$ of the entire cell centre coordinates is therefore obtained as $\mathcal{C} = \mathcal{C}_T \cup \mathcal{C}_R \cup \mathcal{C}_P \cup \mathcal{C}_F$ and the cell footprint is shown. The resulting total number of cells $N_c = |\mathcal{C}|$. Note that the proposed algorithm, described in Algorithm 4.1, is applicable for other antenna beam patterns as long as h, ρ, ε, and β at the SPPC can be defined. The algorithm considers the resulting beam footprints on the ground and not the antenna producing the beams. While the proposed algorithm enables extended HAP coverage, it is important to understand the limits of coverage extension based on the capacity and spectrum performance of the system especially at the edge of coverage. The following section provides a capacity analysis of the extended HAP coverage.

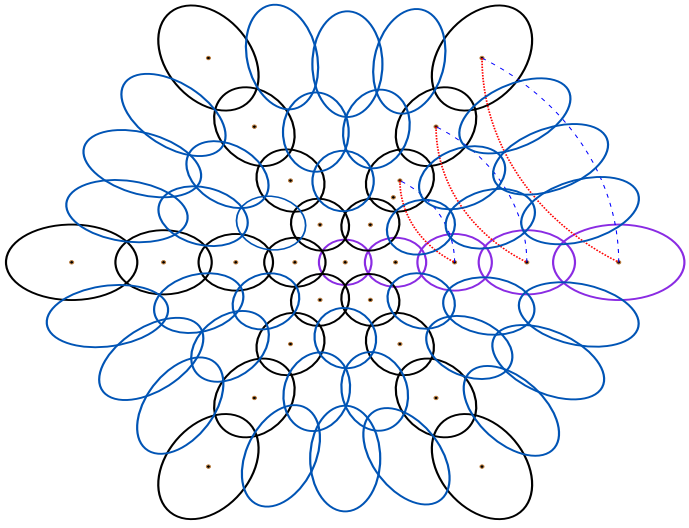

Figure 4.7: The full cellular structure for the HAP extended coverage.

Algorithm 4.1: Cell Pointing Algorithm

Declare h, ρ, ε, β: HAP height, broadside cell subtended angle, overlap ratio, and SPPC boresight angle.
Set $i = 1$ and $d_1 = 0$.
Initialise $\beta_1 = 90^\circ$, $h_1 = h_p$, $\mathcal{C}_T = \{(0, d_1, 0)\}$, $\mathcal{C}_P = \emptyset$ and Φ.
while $\beta_i > 0$ **do**
    Compute $d_{i+1} := h_i U^d(\rho, \beta_i) + d_i$.
    Update $\mathcal{C}_T \mid \mathcal{C}_T := \mathcal{C}_T \cup (0, d_{i+1}, 0)$.
    Update $\beta \mid \beta_{i+1} := \beta_i - 2\rho$.
    Update $h \mid h_{i+1} := \frac{h_i}{\sin(\beta_{i+1})}$.
    $i \leftarrow i + 1$.
end while
Adjust $\mathcal{C}_T$ for overlap, i.e. $\forall d_i \neq 0 \mid i \in \{1,\ 2,\ 3,\ \cdots\} \wedge d_i \in x_i \subseteq \mathcal{C}_T,\ d_i := d_i - h_i \sin(\Phi_i)$.
Obtain new set $\mathcal{C}_R := 60^\circ$ rotation of $\mathcal{C}_T$.
Set ring $j = 0$.
for each $m_i \subseteq \mathcal{C}_T$ and corresponding $n_i \subseteq \mathcal{C}_R \mid i \in \{1,\ 2,\ 3,\ \cdots\}, m_i \neq \emptyset, n_i \neq \emptyset$ **do**
    $V_I(j+1)$:= Interpolation between $m_i(j+1)$ and $n_i(j+1)$ using (4.23).
    Set $V_I^*(j+1)$:= mirror image of $V_I(j+1)$.
    Obtain $\mathcal{C}_P := \mathcal{C}_P \cup \frac{V_I^*(j+1)}{j+1}$ as new cells equidistant from other cells in each ring j.
    $j \leftarrow j + 1$.
end for
Evaluate $\mathcal{C}_F :=$ 5 sets of coordinates where each is a 60° step rotation of $\mathcal{C}_T \cup \mathcal{C}_R \cup \mathcal{C}_P$.
Obtain a set of cell coordinates $\mathcal{C} := \mathcal{C}_T \cup \mathcal{C}_R \cup \mathcal{C}_P \cup \mathcal{C}_F$.
Collect $\mathcal{C}$: Coordinates for beamforming.

Interestingly, Algorithm 4.1 has much less asymptotic time complexity of $\mathcal{O}(c)$ in comparison with that of the state-of-the-art in [6], which is $\mathcal{O}(cik)$, where c, i, k denote the numbers of cells or k-means centroids declared in [6], users, and k-means iteration respectively.

4.4. Spectral Efficiency and Capacity Analysis

In order to objectively assess the limits of extension of the proposed HAP extended coverage, it is important to highlight the theoretical performance bounds of the system. In this section, an expression for the average per-user capacity $\tilde{C_{pu}^i}$ achievable in cell i pointing at any given distance away from the SPP within the HAP service area is derived. Furthermore, an expression for the average area spectral efficiency $\tilde{S_{\eta_i}^f}$ in cell i is obtained as a function of $\tilde{C_{pu}^i}$. Calculating the average area spectral efficiency of a cell pointed at a given distance away from the SPP requires the area of the cell, taking the cell geometry and broadening into consideration. In deriving an expression for HAP cell area, it is assumed that the maximum power point is at the centre of the cell. Ideally, this is not exactly the case, especially for cells formed at low elevation angles. The maximum power point, which might not be at the centre of the beam

depending on the boresight point distance from SPP, is skewed towards the HAP antenna. This is also exacerbated by the broadening effect.

For an elliptical cell i pointing at a distance d_i from the SPP, let x_i and y_i represent its semi-major and semi-minor axes respectively. If h_p is the height of the platform with the other variables retaining their definition from the previous section, the area A_i of cell i is derived as follows. Consider Figure 4.8 partially reproduced from Figure 4.3. From ΔOHP, α_i and h_i are given as follows.

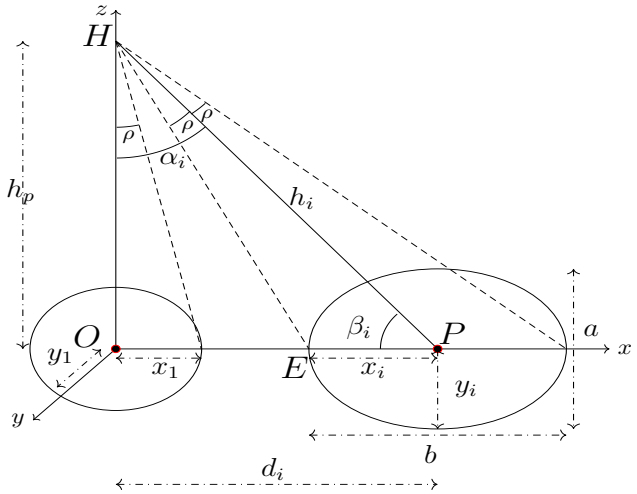

Figure 4.8: HAP elliptical cell geometry (semi-major axis)

$$\alpha_i = \tan^{-1}\left(\frac{d_i}{h_p}\right), \tag{4.24}$$

$$h_i = \sqrt{h_p^2 + d_i^2}. \tag{4.25}$$

Using the sum of angles in a triangle rule in ΔOHP, the elevation angle of the HAP β_i from the centre of cell i is,

$$\beta_i = \frac{\pi}{2} - \alpha_i. \tag{4.26}$$

Hence, using sine rule in ΔHEP, x_i is expressed as follows.

$$x_i = \frac{h_i \sin(\rho)}{\sin(\rho + \beta_i)}. \tag{4.27}$$

Substituting for h_i in (4.27),

$$x_i = \left(\frac{\sin(\rho)}{\sin(\rho+\beta_i)}\right)\sqrt{h_p^2+d_i^2}. \tag{4.28}$$

Using trigonometric identity, (4.28) can be re-expressed as

$$\frac{1}{x_i} = \frac{\sin(\rho)\cos(\beta_i)+\sin(\beta_i)\cos(\rho)}{\sin(\rho)\sqrt{h_p^2+d_i^2}}. \tag{4.29}$$

Expanding (4.29),

$$\begin{aligned}\frac{1}{x_i} &= \left(\frac{\sin(\rho)\cos(\beta_i)}{\sin(\rho)}+\frac{\sin(\beta_i)\cos(\rho)}{\sin(\rho)}\right)\frac{1}{\sqrt{h_p^2+d_i^2}} \\ &= \left(\cos(\beta_i)+\frac{\sin(\beta_i)}{\tan(\rho)}\right)\frac{1}{\sqrt{h_p^2+d_i^2}}.\end{aligned} \tag{4.30}$$

Therefore, the semi-major axis x_i is expressed as follows.

$$x_i = \frac{\sqrt{h_p^2+d_i^2}}{\cos(\beta_i)+\sin(\beta_i)\cot(\rho)}. \tag{4.31}$$

Furthermore, the semi-minor axis y_i can be derived from Figure 4.9 redrawn from Figure 4.8 for convenience.

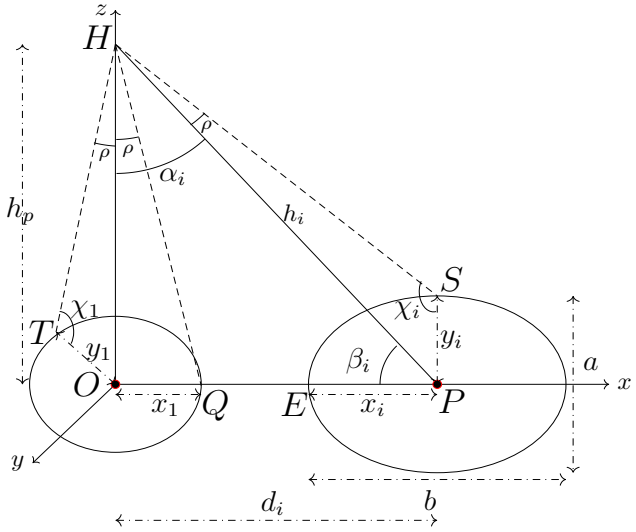

Figure 4.9: HAP elliptical cell geometry (semi-minor axis)

Applying the sum of angle in a triangle and sine rules in ΔHSP, $\angle HSP$ yields

$$\chi_i = \frac{\pi}{2} - \rho, \tag{4.32}$$

and,

$$\frac{h_i}{\sin(\frac{\pi}{2} - \rho)} = \frac{y_i}{\sin(\rho)}. \tag{4.33}$$

Thus,

$$y_i = h_i \tan(\rho). \tag{4.34}$$

Therefore, substituting (4.25) into (4.34) gives,

$$y_i = \left(\sqrt{h_p^2 + d_i^2}\right) \tan(\rho). \tag{4.35}$$

Using (4.31) and (4.35), the expression for surface the area of cell i can be obtained. Recall that the area of an ellipse is given as $\pi x_i y_i$. Thus,

$$\begin{aligned} A_i = \pi x_i y_i &= \pi \left(\frac{\sqrt{h_p^2 + d_i^2}}{\cos(\beta_i) + \sin(\beta_i)\cot(\rho)} \right) \left(\left(\sqrt{h_p^2 + d_i^2}\right) \tan(\rho) \right) \\ &= \frac{\pi \left(h_p^2 + d_i^2\right) \tan(\rho)}{\cos(\beta_i) + \sin(\beta_i)\cot(\rho)}. \end{aligned} \tag{4.36}$$

Multiplying (4.36) by $\frac{\cot(\rho)}{\cot(\rho)}$, and using the trigonometric identity $\tan(\rho)\cot(\rho) = 1$, the area A_i of cell i at a given distance d_i from the SPP is derived as

$$A_i = \frac{\pi(h_p^2 + d_i^2)}{\cos(\beta_i)\cot(\rho) + \sin(\beta_i)\cot^2(\rho)}. \tag{4.37}$$

With the expression for the area of a HAP cell i derived, the average capacity per-user $\tilde{C_{pu}^i}$ and the average area spectral efficiency $\tilde{S_{\eta_f}^i}$ in the cell can now be derived. Let B_i be the user-allocated bandwidth. If P_i^t, G_i^t, G_i^r are the HAP antenna transmit power, transmit gain and user receive gain respectively, λ_s is the signal wavelength, $r(\varphi)$ is the distance of any arbitrary point from the cell centre at an angle φ from its x-axis, N_o is the receiver noise density, and

r_{max} is the distance from the boundary to the centre of cell i, $\tilde{C_{pu}^{i}}$ is then derived. Considering a uniform distribution of users over cell i, a normalised bandwidth density $\tilde{b_i}$ (in Hz/unit area) detected at position (x, y) is expressed as

$$\tilde{b_i}(x, y) = \frac{B_i}{A_i}. \tag{4.38}$$

Hence, the expected capacity density $\tilde{c_{A_i}}$ (in bit/s/unit of surface) for a given realisation is defined using the Shannon capacity equation as [161],

$$\tilde{c_{A_i}}(x, y) = \tilde{b_i}(x, y) \log_2(1 + \gamma_i(x, y)), \tag{4.39}$$

where $\gamma_i(x, y)$ is the CINR at point (x, y) in cell i. Considering a user distribution with density $u_i(x, y)$ in cell i, the average per-user capacity (in bit/s/user) in cell i is expressed as [161]

$$\tilde{C_{pu}^{i}} = \frac{1}{A_i} \iint\limits_{\mathcal{A}} \frac{\tilde{c_{A_i}}(x, y)}{u_i(x, y)} dxdy, \tag{4.40}$$

where $\mathcal{A}$ denotes the i^{th} cell region. Assuming constant user and bandwidth densities u_i and B_i respectively throughout the cell and substituting (4.39) in (4.40),

$$\begin{aligned} \tilde{C_{pu}^{i}} &= \frac{1}{A_i} \iint\limits_{\mathcal{A}} \frac{B_i}{u_i} \log_2(1 + \gamma_i(x, y)) dxdy \\ &= \frac{B}{A_i} \iint\limits_{\mathcal{A}} \log_2(1 + \gamma_i(x, y)) dxdy. \end{aligned} \tag{4.41}$$

With the assumption of constant user density throughout the cell, the term u_i is eliminated in (4.41). Thus, the concept of per-user capacity becomes the same as capacity density [161]. Due to the elliptical geometry of the HAP cell, it is more logical to express (4.41) in polar coordinate form. This involves specifying $\gamma_i(x, y)$ and $dxdy$ in their polar coordinate form. Considering free-space path loss, $\gamma_i(x, y)$ can be expressed as follows:

$$\gamma_i(x, y) = \gamma_i(r(\varphi), \varphi) = \frac{P_i^t G_i^t G_i^r \lambda^2}{N_o (4\pi D_i)^2}. \tag{4.42}$$

Therefore,

$$\tilde{C_{pu}^{i}} = \frac{B_i}{A_i} \iint_{\mathcal{A}} \log_2 \left(1 + \frac{P_i^t G_i^t G_i^r \lambda^2}{N_o (4\pi D_i)^2}\right) r(\varphi) dr(\varphi) d\varphi, \tag{4.43}$$

where D_i is the distance of an arbitrary user in cell i to the HAP. This is based on a single cell scenario, which noise limited and gives the maximum achievable capacity. D_i needs to be resolved as a function of d_i and h_i, and represented in polar coordinates as derived below. This can be achieved by considering the geometry in Figure 4.10.

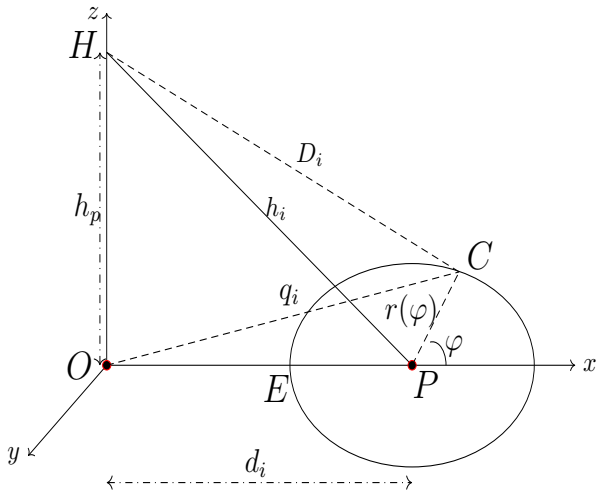

Figure 4.10: HAP elliptical cell geometry (polar coordinates)

Imagine that the elliptical cell is flat on the ground and is bounded within a 2D space of radius $r(\varphi)$. The radius varies with respect to the azimuth angle φ. With origin at the centre of the cell, therefore, $r(\varphi)$ is defined as [162]

$$\begin{aligned} r(\varphi) &= \frac{x_i y_i}{\sqrt{(y_i \cos(\varphi))^2 + (x_i \sin(\varphi))^2}} \\ &= \frac{y_i}{\sqrt{1 - (e \cos(\varphi))^2}}, \end{aligned} \tag{4.44}$$

where e is the cell eccentricity, x_i and y_i are the semi-major and semi-minor axes defined in (4.31) and (4.35) respectively. e is expressed as

$$e = \sqrt{1 - \left(\frac{y_i}{x_i}\right)^2}. \tag{4.45}$$

By applying cosine rule on ΔPOC in the above figure, D_i can be expressed as a function of φ. Firstly, the distance q_i of an arbitrary user from the SPP is obtained as follows.

$$q_i^2 = d_i^2 + r(\varphi)^2 + 2r(\varphi)d_i\cos(\varphi). \tag{4.46}$$

Then, considering ΔHOC and substituting for q_i,

$$D_i^2 = h_p^2 + q_i^2 = h_p^2 + d_i^2 + r(\varphi)^2 + 2r(\varphi)d_i\cos(\varphi). \tag{4.47}$$

Finally, substituting for D_i in (4.43), the average per-user capacity $\tilde{C_{pu}}$ is given as

$$\tilde{C_{pu}^i} = \frac{B_i}{A_i}\int_0^{2\pi}\int_0^{r_{max}} \log_2\left(1 + \frac{P_i^t G_i^t G_i^r \lambda^2}{\xi}\right) r(\varphi)dr(\varphi)d\varphi, \tag{4.48}$$

where,

$$\xi = N_o(4\pi)^2(h_p^2 + d_i^2 + r(\varphi)^2 + 2r(\varphi)d_i\cos(\varphi)). \tag{4.49}$$

Unfortunately, there is no closed form expression for (4.48). This was verified after trying Chebyshev–Gauss quadrature approximation and using Mathematica, a software tool for mathematical analysis. However, it can be evaluated using numerical methods.

Consequently, the average area spectral efficiency $\tilde{S_{\eta_f}^i}$ in cell i is then expressed as

$$\tilde{S_{\eta_f}^i} = \frac{\tilde{C_{pu}^i}}{B_i A_i} = \frac{\eta_f^i}{A_i}. \tag{4.50}$$

where $\eta_f^i = \frac{\tilde{C_{pu}^i}}{B_i}$ is the spectral efficiency. Average area spectral efficiency $\tilde{S_{\eta_f}^i}$ highlights the effect of cell broadening on the capacity of a cell, which also skews the maximum power point of a beam away from the centre of the beam, by also considering the cell area. This effect becomes increasingly severe as the HAP service area increases. The degree of service area extension may be based on a given minimum average area spectral efficiency required which can be a design parameter. The widely used capacity density metric can be obtained by multiplying the average area spectral efficiency by the bandwidth of the cell.

4.5. Performance Evaluation

A simulation is set up to evaluate the performance of the proposed algorithm with the parameters given in Table 4.1. Using the antenna profile in (4.1) and 1600 antenna elements [163],

Table 4.1: Details of the simulation parameters.

Parameters	**Simulation Values**
HAP height h_p	20 km
HAP transmit power P_i^t	33 dBm
Channel bandwidth B	20 MHz
Noise figure N_f	5 dB
Receiver noise floor N^*	-95 dBm
Frequency f	2.1 GHz
Service area radius R	60 km
Angle subtended ρ	3.5°
Overlap ratio ε	0.1
User density λ	2 users/km^2
Receive antenna gain G_i^r	1.5 dB
Number of Antenna elements MN	1600

* The receiver noise floor is evaluated in dBm using $10log_{10}(kTB) + N_f + 30$, where k = 1.38*10^{-23} J/K is Boltzmann constant, T = 290 K is the assumed receiver temperature, $B = 20$ MHz is the assumed channel bandwidth, N_f = 5 dB is the noise floor based on 3GPP TR 36.942 [164].

heuristically, it is determined that the edge-of-cell subtended angle $\rho = 3.5°$, resulting in an SPPC of approximately 2.5 km diameter. This is used in Algorithm 4.1 with the resulting cell centre coordinates used as boresights in the antenna system. The CNR and CINR of all users in the service area are evaluated using (4.4) and (4.8) respectively. The results are compared with other alternative cell placement schemes including the state-of-the-art in [6].

4.5.1. Beam-pointing performance

In the equidistant scheme, cells are pointed such that neighbouring cells are approximately 2.5 km apart. Figure 4.11 shows the user CNR contour of the equidistant scheme, with antenna boresights at the centres of the cells. The contour highlights the significant overlap between neighbouring cells resulting from beam broadening. The severe overlap gives rise to high ICI, which worsens at the edges of cells and service area. Therefore, considering beam broadening and overlap is important. Figure 4.12 shows the user CNR contour with cells pointed such that neighbouring antenna boresights are 7° apart, which is the equiangular scheme. Similarly, this scheme results in severe overlap between cells because it does not explicitly consider beam broadening. The severe overlap and poor CNR, especially towards the edges of cells and coverage area, make these schemes challenging in practical systems due to significant ICI.

Unlike the equidistant and equiangular schemes, Algorithm 4.1 produces properly structured cells with better overlap control and CNR performance as shown in Figure 4.13, due to the direct consideration of beam broadening. The algorithm's CINR performance is also compared with the equidistant and equiangular schemes in addition to schemes proposed in [6]. The

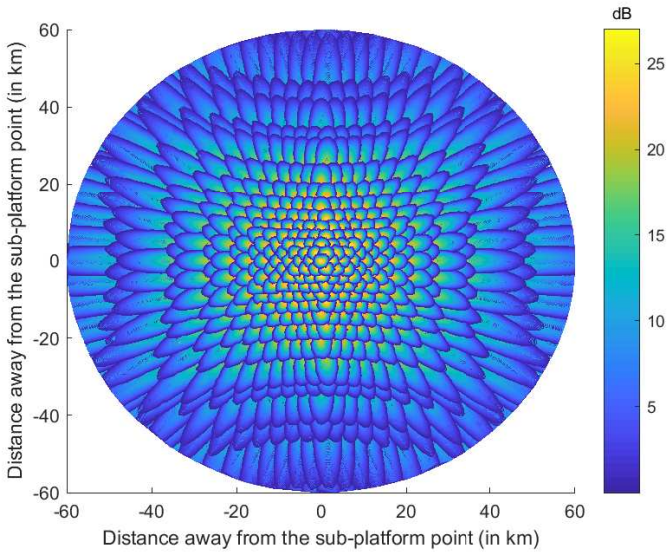

Figure 4.11: CNR contour within cells of the equidistant scheme.

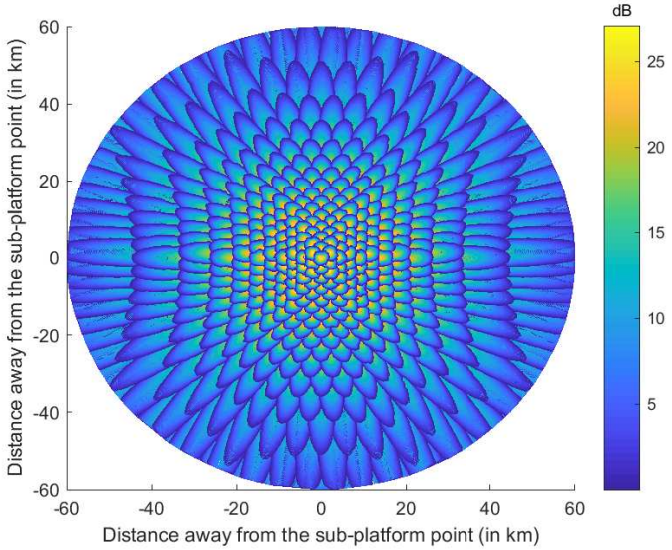

Figure 4.12: CNR contour within cells of the equiangular scheme.

empirical CDFs of user CINR distribution obtained with the different cell-pointing schemes are given in Figure 4.14 and Figure 4.15. Over 90% of the users using the proposed scheme achieve a CINR greater than 0 dB compared with the the equiangular and equidistant schemes with less than 50% of the users achieving above 0 dB CINR as shown in Figure 4.14. The scheme proposed in this chapter results in a CINR improvement of 7–15 dB. Furthermore, random and regular pointing of cells are presented in [6], as well as the proposed state-of-the-art clustering of users using k-means clustering with cells pointed at the centroid of the clusters. The CDFs of user CINR distributions of these schemes are compared with that of the enhanced and validated scheme proposed. Using the schemes in [6], Figure 4.15 shows that

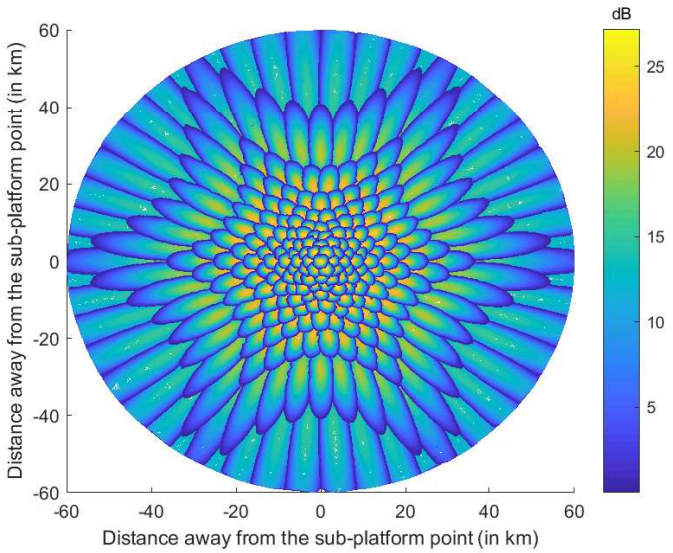

Figure 4.13: CNR contour within cells using the proposed scheme.

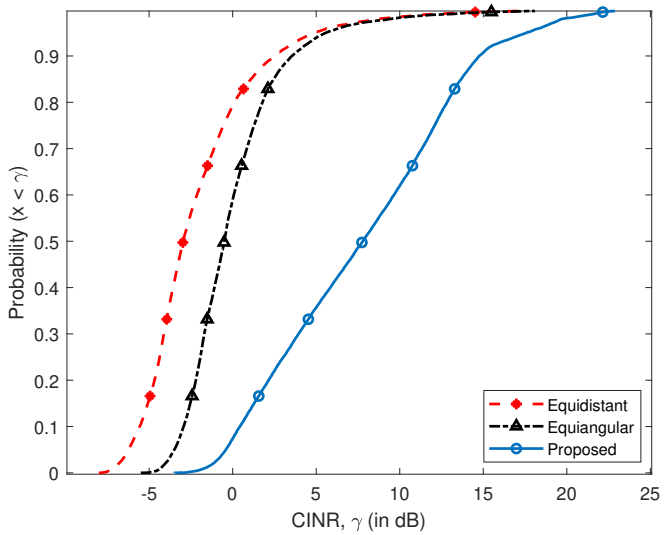

Figure 4.14: CINR distribution of the proposed scheme with equidistant and equiangular schemes.

less than 60% of the users achieve over 0 dB CINR compared with the over 90% obtainable using the proposed scheme with the additional CINR improvement ranging from 5–10 dB.

The effect of overlap ratio ε on both user allocation probability and 95$^{\text{th}}$ percentile user throughput is shown in Figure 4.16. On the one hand, user allocation probability is evaluated as the ratio of the number of users allocated to a cell to the total number of users. On the other hand, the 95$^{\text{th}}$ percentile is obtained by calculating the throughput of all users in the HAP system using (4.10). The results show that the user allocation probability increases with increasing overlap ratio up to a point beyond which it starts decreasing. This is expected as increasing overlap plugs coverage holes until there are no longer holes within the system coverage. In-

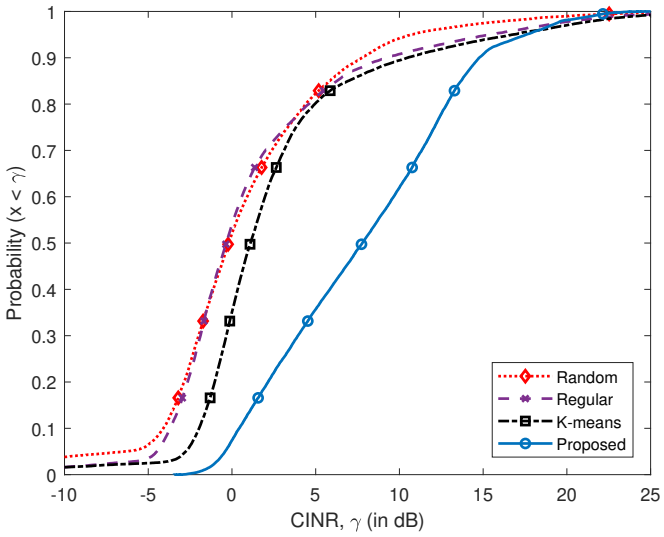

Figure 4.15: CINR distribution of the proposed scheme and schemes in [6].

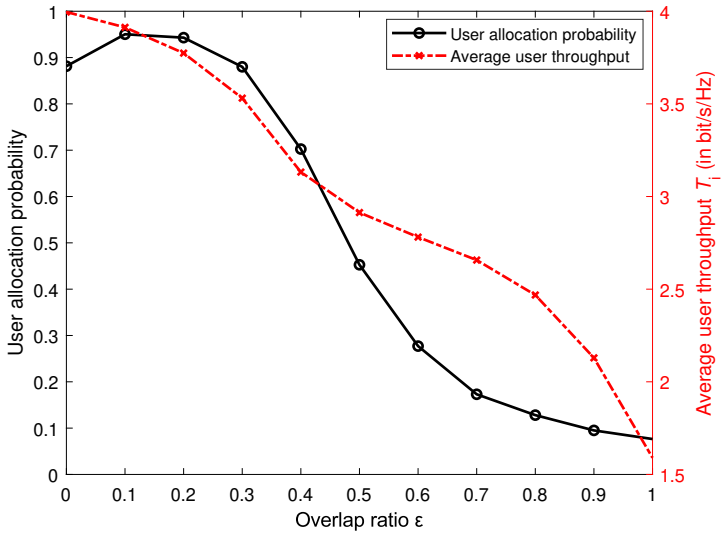

Figure 4.16: Overlap ratio vs. user allocation probability and 95th percentile user throughput.

troducing more overlap at this point starts creating coverage holes, resulting in a decrease in the user allocation probability. Furthermore, the 95th percentile user throughput expectedly decreases with increasing overlap as a result of the increasing ICI, which affects user throughputs. The best value for ε can be derived as an optimisation problem such that both user allocation probability and user throughput are maximised.

Figure 4.17 presents the CDFs of the achievable throughput per user for the different cell-pointing schemes, showing the probability that a user achieves a throughput greater than T. It highlights improvements of between 40%–70% by the proposed scheme compared with the other schemes, with more than 80% of users achieving throughput greater than 1 bit/s/Hz using

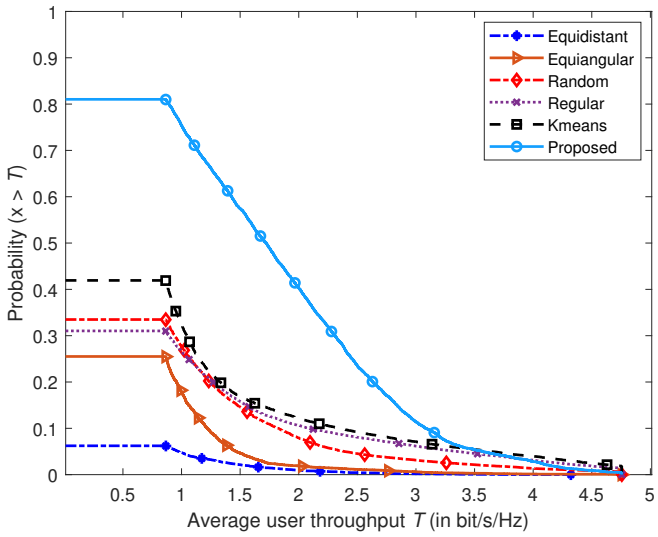

Figure 4.17: User throughput distribution of different cell-pointing schemes.

the proposed scheme compared with about 40% for the state-of-the-art. The improvement is profound because the other algorithms were not developed for extended coverage scenarios, therefore, they do not consider beam broadening, which worsens ICI and results in the poor throughput performance.

4.5.2. Capacity performance

Having evaluated the performance of the pointing algorithm, the capacity performance of the derived models is evaluated in this section. The results of the capacity evaluation facilitate the estimation of the theoretical limits of coverage extension of the HAP extended coverage system based on the desired minimum quality of service at the edges of cells and overall coverage area. To evaluate the average spectral efficiency of the system, a cell is pointed at increasing distances away from the SPP up to the edge of the extended coverage area.

After pointing the cell, area, average spectral efficiency, and average area spectral efficiency are evaluated. The upper and lower limits of the average area spectral efficiency are obtained. The lower limit is obtained by assuming that all users in the cell have a CNR equal to that of a user at the boundary of the cell (i.e. 9dB corresponding to the edge-of-cell CNR). This is then used in the Shannon equation (i.e. $log_2(1+CNR_{min})$) to compute the lower limit average area spectral efficiency by dividing with the cell area. For the upper limit, it is assumed that all users have a CNR equal to the peak CNR in the CNR distribution, which is the CNR at boresight. This is used in the Shannon equation (i.e. $log_2(1 + CNR_{max})$) to obtain the maximum achievable average spectral efficiency, which is divided by the cell area to obtain the average area spectral

efficiency. The lower and upper limit average area spectral efficiency give the bounds of each pointed cell. A different set of area spectral efficiency is also obtained by evaluating (4.48) using numerical integration. The parameters in Table 4.1 and a 30 dB transmit gain [71] at boresight are used to simplify the integration and validate the derived expression (4.48). These average area spectral efficiency values (i.e. lower limit, upper limit, and integral) are plotted against the distance of the cells from the SPP as shown in Figure 4.18. Note that the values represent the best case scenario because they are obtained without considering interference by assuming that the system is only noise limited. In practical systems, the achieved average area spectral efficiency is expected to be lower when interference is taken into account. The level of interference depends on the number of cells formed and the distance between the cells in addition to the antenna beam profile.

Figure 4.18 shows the average area spectral efficiency of cells pointing at increasing distances from the SPP. Notably, the integral average area spectral efficiency values are close to the maximum because peak transmit gain is used to simplify the (4.48) as mentioned above. The average area spectral efficiency reduces with increasing distance from the SPP due to two main factors increasing path loss and cell area. The number of interfering cells, their proximity to the cell under consideration and the HAP antenna beam profile are some other factors. The average area spectral efficiency starts tailing-off as the cell centre distance from the SPP approaches 60 km. The value at 60 km is approximately 0.05 bit/s/Hz/km^2. If users are allocated a resource block (RB) group each with 750 kHz bandwidth for instance and the area of a cell pointing at the 60 km distance using (4.37) is evaluated to be approximately 125 km^2, the best case

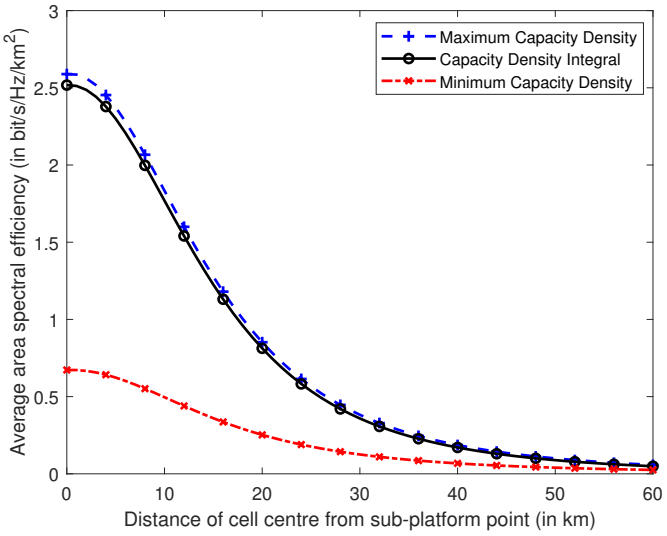

Figure 4.18: Average area spectral efficiency against distance of cell centre

achievable capacity of a user in this cell is approximately 4.6 Mbps. Practically, signals at cell edges will be considerably degraded due to ICI. System designers can therefore workout how wide their HAP service area can cover based on the number of cells to be deployed and the required capacity of the edge-of-coverage users. The average area spectral efficiency values obtained from evaluating (4.50), as shown in Figure 4.18, are close to the values obtained when all users in a cell have a signal gain that is equal to the boresight gain. The closeness is expected as there is a small difference between the boresight gain and the gain at the edge of cell due to small angle subtended ($\rho = 3.5^{\circ}$) by the cell centre and edge at the HAP.

4.6. Conclusion

In this chapter, a beam-pointing algorithm to deliver contiguous cellular coverage from a HAP over an extended service area of at least 60 km radius was proposed and validated. The algorithm enables the extension of the achievable HAPs coverage area radius state-of-the-art, which is under 30 km. Additionally, expressions for estimating the area, average spectral efficiency, and average and area spectral efficiencies of a HAP cell pointed at any given location within the HAP service area were derived. Using a uniform planar antenna array, the performance of the proposed algorithm was studied over a 60 km radius service area, which is considerably larger than the area of 30 km radius that much of the HAP related literature thus far focus on, including the state-of-the-art. It was shown that users in a HAP system using a cellular architecture obtained by using the proposed algorithm achieves CINR values between 5–15 dB better than the other schemes with better control of beam overlap. More than 90% of users achieve CINR greater than zero as against 45–70% in the compared schemes, which highlights its significant performance improvement.

In addition, theoretical models for estimating the bounds of HAP coverage extension given the operational parameters were derived. These include models for HAP cell footprint area and average area spectral efficiency. Using the derived models, the average capacity, spectral efficiency, and average area spectral efficiency for cells pointed at increasing distances up to 60 km were estimated. The results show that all three decrease with increasing SPP distance and that the average area spectral efficiency reduces significantly at extended areas due to the increase in cell size and path loss. For instance, the estimated average area spectral efficiency drops from 6 bits/s/Hz/km^2 at the centre to about 0.05 bits/s/Hz/km^2 at the edge of a 60 km radius service area. On coverage extension, the derived expressions could be used to estimate bounds of the achievable HAP coverage extension based on the required minimum QoS.

Chapter 5. Coverage and Capacity Enhancement with Two-tier Architecture

Contents

5.1. Motivation

The HAP beam-pointing algorithm and approach proposed in Chapter 4 describes how conventional contiguous cellular coverage for service provisioning using a HAP with multi-element planar antenna array can be delivered over an extended service area. However, extended service area cellular service provisioning is severely affected by inter-cell interference (ICI) and beam broadening especially at low elevation angles, which results in disproportionately sized projections of beams on the ground. The ICI also leads to the creation of coverage holes and poor cellular performance especially at the edges of cells and the extended service area. In addition, due to the effect of beam broadening, cells closer to the edge of the service area are multiple times bigger in size than cells closer to the sub-platform point (SPP), resulting in potentially more users sharing the limited spectrum resources in edge cells compared with cells closer to the SPP given a reuse factor of 1. This occurs despite the reducing signal strength with increasing distance away from the centre of coverage. These issues affect the overall per-

formance of the extended coverage HAP cellular system in general and the quality of service (QoS) of users especially at the edge of the extended service area in particular.

An appropriate HAP cellular architecture that specifically considers the peculiar features of the HAP beams in the extended coverage scenario can mitigate the challenges introduced by beam broadening and ICI. This chapter proposes a novel tier-based cellular architecture, designed to significantly mitigate the impact of ICI and beam broadening on the HAP extended coverage system in order to improve the user and overall network performance. In this thesis, a tier is defined as a collection of cells forming a contiguous cellular structure as in Chapter 4, which provides coverage over the entire HAP extended service area. This chapter also discusses how multiple beams can be deployed practically to form a tier-based architecture using a multi-beam antenna (MBA) or multi-beam phased-array antenna (MBPAA). The proposed architecture, which is designed to form homogeneous tiers of contiguous cells with users associating with the best cell and tier based on given metrics, aims to minimise edge-of-cell user connectivity and eliminate coverage holes by significantly enhancing the CINR of users. In studies on heterogeneous cellular network (HetNet) architectures [165, 166], the focus is on having a larger cell overlaying a number of small cells with each cell delivered from a different infrastructure, which is potentially prohibitively expensive and complicated for a HAP cellular network. The architecture proposed in this chapter comprises of a structure of contiguous cells with the tiers, formed from the same HAP infrastructure, having similar cellular characteristics and structure. It is about enhancing coverage and capacity using a tier-based architecture.

5.2. The Two-tier HAP Architecture

The fundamental challenge with the typical one-tier HAP cellular architecture for service provisioning over an extended area proposed in Chapter 4 is the poor edge-of-cell performance, which worsens with increasing distance away from the SPP as the beams broaden in size. Consequently, a two-tier cellular architecture is proposed where edge users in one tier potentially become centre users in the other tier. The architecture is comprised of alternative contiguous structures of cells of similar coverage, radio and antenna beam characteristics. Users are moved between tiers based on either the user or network QoS. The availability of alternative cells and tiers with orthogonal channel conditions that users can be associated with enhances their robustness.

Figure 5.1 illustrates the proposed architecture using two tiers of three contiguous cells with similar antenna beam characteristics. A one-tier architecture would typically be made up of

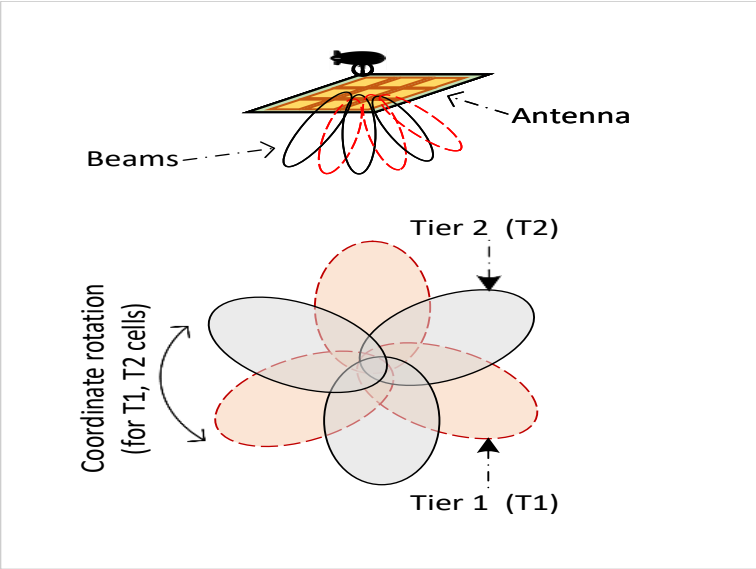

Figure 5.1: Example footprint of the proposed two-tier architecture highlighting the contiguous Tier 1 and Tier 2.

only the footprint of one of the tiers, however, significantly better overall performance can be achieved by using the proposed two-tier architecture. Figure 5.2 further highlights the potential benefits of the proposed architecture in terms of the robustness and flexibility of the system. Imagine the second tier as an offset in angle of the first tier, if the beam boresight (i.e. axis of maximum radiated power and gain) of the beamforming antenna is at the centre of the cells, the gain at the edge of the cells is expectedly less than the boresight value depending on the size of the beam and the number of the beamforming antenna elements. Consequently, it is expected that the signal quality of users at the edge of a cell is significantly less compared with those at the centre. Thus, as illustrated in Figure 5.2 using a linear arrangement of two cells per tier, the proposed two-tier architecture seeks to enhance the CNR and CINR of some users

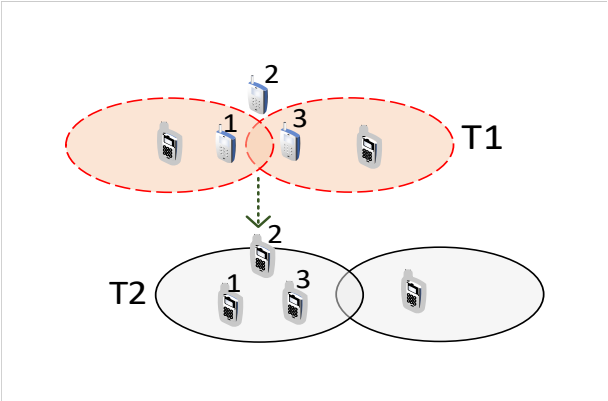

Figure 5.2: Network robustness and flexibility of the proposed two-tier architecture highlighting how the robustness of HAP network and users can be enhanced. The network associates the user to any of the two tiers that maximises the user experience.

by providing alternative options for edge users to enhance their received signal quality. For instance, UE 1 and 2 in Figure 5.2 moved from the edge of cells in Tier 1 (T1) to the centre of a cell in Tier 2 (T2), enhancing their signal quality. Some users such as UE 3, which ordinarily would have no coverage in a one tier network based on T1, are provided coverage by T2.

5.2.1. One-tier vs. Two-tier Architecture

Due to ICI and beam broadening resulting in disproportionately sized cells, the one-tier cellular architecture for the HAP extended coverage system in Chapter 4 can potentially experience coverage holes and unbalanced loading of users and traffic in cells across the network. However, with the proposed HAP two-tier architecture, users have the flexibility of associating to the tier that provides better signal quality or satisfies some given conditions, ensuring that the trunking efficiency in each cell remains sufficient to deliver adequate performance, while controlling the loading between cells. Additionally, it minimises cell edge user connectivity by significantly enhancing user signal quality, thereby mitigating the effect of ICI and mitigating coverage holes. In terms of complexity, the two-tier architecture implementation is expectedly more complex than a typical one-tier architecture.

5.2.2. Delivering a Tier-based Architecture

This section presents an intuitive discussion and practical considerations on how a tier-based HAP architecture can be delivered by leveraging the concept of forming multiple simultaneous beams from an array antenna aperture, which has been thoroughly discussed in the literature [167–170]. Irrespective of how multiple beams are delivered, there is need for proper isolation between the beams and tiers to ensure orthogonality and minimise interference, which is achieved by ensuring a low beam crossover level.

MBAs can be classified into passive MBA (PMBA), active digital MBA (DMBA), and multi-beam PAA (MBPAA) [170]. PMBA requires multiple cascaded beamforming networks (BFNs) such as Butler or Blass matrices, which doubles with the proposed two-tier architecture, to deliver beams required for HAP coverage over an extended service area. This results in a complicated antenna system layout and increased insertion loss, which can be mitigated by increasing the aperture, although it is limited by the aerial platform capabilities. The beams required for the two-tier architecture could however be divided into two different apertures, each delivering the beams needed for one tier, and the beams from the apertures could further be isolated with different polarisations [167]. Such solutions will result in an antenna system with reduced complexity, crossover level, and power consumption inherent in BFNs. The disadvantage is the

resultant large form factor due to the use of large numbers of couplers, phase shifters, feeders and lack of flexibility of BFNs as the number of beams and their pointing direction must be predefined prior to manufacture [170].

On the other hand, MBPAA is a more flexible antenna system since they are electronically scanned and have smaller form factor with no couplers compared with PMBA. Similarly, DMBA has even more flexibility and smaller form factor with no couplers or phase shifters. In both antenna systems, the number of beams and their directions for the two-tier architecture could be provided as inputs. Compared with PMBA, MBPAA and DMBA designs are more complicated and experience higher power consumption. Consequently, realistic and practical implementation of the two-tier architecture is possible with some of these antenna systems, however, size, weight, and power (SWaP) constraints should be considered among other factors given the limitations of the platforms as discussed in Chapter 3.

5.3. System Implementation

This section presents an implementation of the proposed two-tier architecture including the system model, performance metrics, tier formation, and user association to both a cell and tier. Fundamentally, the HAP forms multiple beams, which are used to create cellular footprints, at specific locations and elevation angles on the ground. Commonly, the resulting cellular footprints produce a single tier cellular architecture covering an area of around 30 km radius as in [6], which fails to fully exploit HAP benefits. Consequently, the approach proposed in Chapter 4 at least doubles the HAP coverage area radius to 60 km. However, the resulting architecture in Chapter 4 still suffers from ICI leading to poor signal quality especially for edge users and disproportionate cell sizes arising as result of beam broadening at low elevation angle, which leads to uneven loading in the network. Among other things, the implementation also investigates how orthogonality between the tiers can be achieved.

5.3.1. System Model and Performance Metrics

Generally, the same system model, user distribution, and some performance metrics such as CNR Γ_i, CINR γ_i, user and system throughput T_i discussed in Section 4.2 of Chapter 4 are also adopted in this chapter. However, a more complex user association procedure, described later in Section 5.3.3 of this chapter is implemented. In addition, some other performance metrics such as user allocation probability, user satisfaction probability, and Jain's fairness index are defined and used here for further performance evaluation of the proposed architecture.

User allocation probability $\frac{i_{RB}}{|\mathcal{I}|}$ is the ratio of the number of users allocated at least one spectrum resource block (RB) i_{RB} to the total number of users in the system $|\mathcal{I}|$. On the other hand, user satisfaction probability is the ratio of the number of users that achieved at least a particular satisfaction throughput T_{th} to the total number of users (i.e. $\frac{i_{T_{th}}}{|\mathcal{I}|}$). Jain's fairness index J is given as [171, 172]

$$J(\tilde{T}_1, \tilde{T}_2, \cdots, \tilde{T}_I) = \frac{(\sum_{i=1}^{|\mathcal{I}|} \tilde{T}_i)^2}{|\mathcal{I}| 1 \sum_{i=1}^{|\mathcal{I}|} \tilde{T}_i{}^2}, \tag{5.1}$$

where $\tilde{T}_i$ is the mean throughput of user i. $J(\tilde{T}_1, \tilde{T}_2, \cdots, \tilde{T}_I)$ is maximised when all users receive equal throughput.

5.3.2. Tier Formation

The formation of the proposed two-tier HAP cellular architecture for extended coverage shown in Figure 5.3 requires the contiguous cell centre coordinates obtained from Algorithm 4.1, the cell-pointing algorithm proposed in Chapter 4. The algorithm generates a set of coordinates used by the HAP antenna system for beam-pointing, to create a single tier of contiguous HAP cell structure across an extended service area.

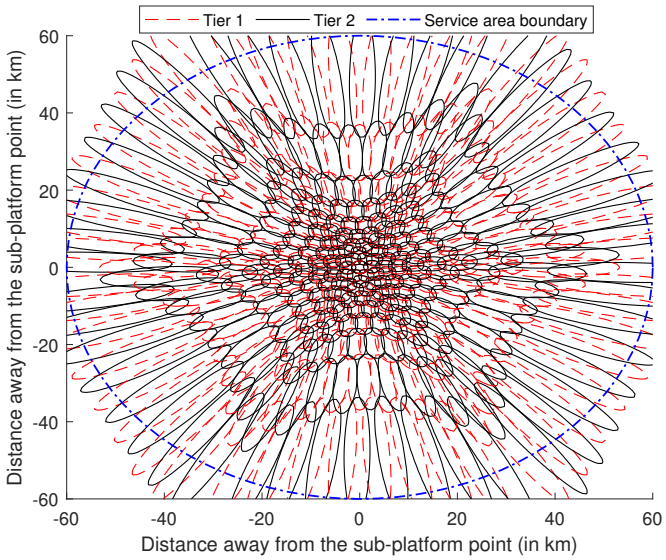

Figure 5.3: The two-tier architecture cellular footprint.

To then form the two-tier cellular architecture, the set of coordinates used in forming the single-tier architecture for tier 1 (T1) is rotated by an offset angle ξ to obtain a new set of coordinates for beam-pointing that creates a different contiguous HAP cell structure for T2 using

$$c^{'} = \begin{bmatrix} cos(\xi) & sin(\xi) \\ -sin(\xi) & cos(\xi) \end{bmatrix} c, \tag{5.2}$$

where $c^{'} \subseteq \mathcal{C}^2$ and $c \subseteq \mathcal{C}^1$ are column vectors denoting the rotated x- and y- axes coordinates for beamforming to form T1 and T2 respectively as given in Algorithm 5.1. The offset angle ξ can be defined by network operators based on their network requirements. In this thesis, the objective is to improve coverage and capacity by heuristically enhancing both user CNR and CINR, in addition to loading balancing between the tiers. The resulting two sets of coordinates, each forming one tier, are used for cell formation. In Figure 5.3, notice that the cells in one tier result from the rotation of the other tier's cells such that the edges of cells in one tier corresponds to the centre of a similar cells in the second tier. This is achieved by setting $\xi = \rho$, where ρ is the angle subtended at the HAP by the centre and edge of the broadside cell, which is approximately equal for all cells. The detailed tier formation algorithm is presented in Algorithm 5.1 with complexity analysis below. The two-tier cell centre coordinates obtained from Algorithm 5.1 are used by the antenna systems, which forms and points beams at these coordinates to produce the cellular architecture shown in Figure 5.3.

Algorithm 5.1: Tier Formation Algorithm

1: Run Algorithm 4.1 to obtain a obtain a set of tier 1 (T1) cell coordinates $\mathcal{C}^1$.
2: Obtain the broadside cell subtended angle ρ.
3: Set ξ : Tier offset angle $\xi \mid 0^\circ < \xi < 2\rho$.
4: Initialise the set of new cell coordinates $\mathcal{C}^2 = \emptyset$ for tier 2 (T2).
5: **for** each $c \subseteq \mathcal{C}^1 \mid c \in \mathbb{R}^2,\ c \neq \emptyset$ **do**
6: Evaluate $c_n := c$ rotated by angle ξ using (5.2).
7: Update $\mathcal{C}^2 := \mathcal{C}^2 \cup c_n$.
8: **end for**
9: Obtain set of all cell coordinates $\mathcal{T}^{1,2} := \mathcal{C}^1 \uplus \mathcal{C}^2$.
10: Collect $\mathcal{T}^{1,2}$, $\mathcal{C}^1$, $\mathcal{C}^2$: Coordinates used for beam pointing to form T1 and T2.

Complexity of Tier Formation Algorithm

Interestingly, Algorithm 5.1 has the same asymptotic time complexity as our cell-pointing algorithm earlier proposed in Chapter 4. The complexity is $\mathcal{O}(c)$, where c denotes the numbers of cells within the HAP service area. Therefore, the greater the number of cells or the wider the service are, the more complex the system becomes.

5.3.3. User Association

Each user in the network associates with a cell and a corresponding tier that maximises its CNR and CINR respectively. Typically, user i's Γ_i and γ_i, evaluated using (4.4) and (4.8) in

Chapter 4, greedily determine which cell and tier respectively the user associates with. User i associates with a cell $\iff \Gamma_i \geq \Gamma_{thr}$, where Γ_{thr} is the minimum CNR threshold required for user-cell association. In the association process, user i's Γ_i in all the covering cells in both tiers are initially obtained. User i then temporarily associates with the two cells, one in each tier, providing the highest Γ_i. Then, the user's γ_i in both cells are evaluated with the user permanently associating with the tier and and its corresponding cell providing the highest γ_i. If user i is located within a region where two or more cells in the same tier overlap, it associates with the cell that maximises its received power P_i^r. Invariably, users in an overlap region can detect all of the overlapping cells. Algorithm 5.2, which is implementable in a central controller at the HAP, either in a full base station configuration or on the ground, performs the user association. The controller can then forward association messages containing the cell and tier information to users for connection.

Algorithm 5.2: User Association Algorithm

1: Declare $\mathcal{T}^{1,2}$, $\mathcal{I}$, T: Tier information, associating user information, and number of tiers.
2: Set Γ_{thr}: Minimum CNR threshold.
3: Initialise the association information matrix $B_{i,t,c}$.
4: **for** each user $i \in \mathcal{I}$ **do**
5: **for** each tier $t \in T$ **do**
6: **for** each cell $c \in \mathcal{C}^t \mid \mathcal{C}^t \subseteq \mathcal{T}^{1,2}$ **do**
7: Compute Γ_i using (4.4).
8: **if** $\Gamma_i \geq \Gamma_{thr}$ **then**
9: Record Γ_i and the corresponding cell c, i.e. $A_i(c) := \Gamma_i$.
10: **end if**
11: **end for**
12: Obtain the cell c_i that maximises Γ_i from Line 9, i.e. $c_i = \arg\max_c A_i(c)$.
13: Compute the γ_i in cell c_i using (4.8).
14: Record the user's c_i and γ_i in tier t, i.e. $S_i(t) = [c_i, \gamma_i]$.
15: **end for**
16: Obtain the tier t and the corresponding cell c_i that maximises the user's γ_i from $S_i(t)$.
17: Update $B_{i,t,c}$ with user i's t, c_i, and γ_i.
18: **end for**
19: Collect $B_{i,t,c}$: User association information matrix used for user association.

Complexity of User Association Algorithm

The complexity of Algorithm 5.2 is $\mathcal{O}(itc)$, where i, t, c are the total number of cells, users, and tiers respectively. To reduce the complexity of Algorithm 5.2 in practical HAP system deployment, instead of evaluating the perceived Γ_i from all the cells for user i, the evaluation will be carried out for only the $|\mathcal{C}|$ cells providing signals with the highest quality for the UE based on a reference signal received power (RSRP). Consequently, Line 6 can be for each $c \in \mathcal{C} \mid \mathcal{C} \subseteq \mathcal{C}^t$, where $|\mathcal{C}|$ is controlled by appropriately setting the required minimum RSRP

depending on the network operator's requirement or standard for the specific technology such as LTE, 5G, or 6G. For instance, the minimum RSRP for LTE based on 3GPP standard is -140 dBm [173], which will limit $|\mathcal{C}|$ to just a few cells. Therefore, the reduced complexity becomes $\Omega(itc_s)$, where $c_s << c$ is the number of cells in the subset of the total number of cells in the service area covering user i. This is similar in both the one-tier and two-tier cases.

5.3.4. Resource Allocation

In a HAP system based on the proposed two-tier architecture, three resource allocation decisions must be made starting with the decision on how the system resources are shared between the tiers and then the cells. In LTE and beyond systems, implementing the proposed architecture would involve the sharing of RBs between the tiers. Subsequently, each tier then decides how to allocate resources between its cells, and eventually each cell in a tier assigns its allocated resources to users accordingly.

Full spectrum reuse (FR) and partial spectrum reuse (PR) schemes are considered for the sharing of RBs at each level of abstraction in the two-tier architecture. The FR scheme allows each cell in each tier to fully reuse all available RBs with a cell allocating RBs to its associated users based only on the local RB usage in the cell. Conversely, the system RBs in the PR scheme are partitioned between the two tiers with each cell fully reusing only its tier's allocated RBs. Defining mathematically, let $R_b = \{1, 2, 3, \cdots, N_{R_b}\}$ denote the whole set of system RBs, where N_{R_b} is the total number of the overall system RBs. In the FR scheme, each cell in each tier t is assigned a set $\tilde{R_b^t} = R_b$. In the PR scheme, $\tilde{R_b^1}$ and $\tilde{R_b^2}$ are the sets of RBs assigned to each cell in tiers 1 and 2 respectively such that $\tilde{R_b^1}, \tilde{R_b^2} \subseteq R_b$ and $\tilde{R_b^1} \cap \tilde{R_b^2} = \emptyset$. The system RBs are shared between the tiers such that each tier is assigned $\left\lfloor \frac{i_t}{I} N_{R_b} \right\rceil$ RBs, where i_t is the number of users associated with the tier, I is total number of users in the network, and $\lfloor . \rceil$ represents rounding to the nearest integer.

Irrespective of the allocation scheme, the HAP system expectedly will suffer from intra-tier ICI. Additionally, the FR scheme also suffers from considerable inter-tier interference (ITI) due to the significant overlap of the main lobes of the beams that form the two tiers. Therefore, the ITI, which significantly degrades the system performance, should be mitigated to enhance performance [174] and fully exploit the benefits of the two-tier architecture. The PR scheme eliminates ITI by ensuring orthogonality between the tiers, however, it still suffers from ICI. An appropriate scheme should mitigate both ICI and ITI to enhance system performance. Here, the possibility of using ICIC for interference mitigation in the HAP two-tier network is investigated. ICIC is discussed in more detail in Section 2.4.2.

Unlike conventional ICIC in terrestrial cellular networks, a modified ICIC where a cell is divided into edge and centre regions is implemented here as shown in Figure 5.4. Considering the steep roll-off of power inherent in HAP cells as highlighted in Section 2.4.2 of Chapter 2, neighbouring cells can reuse RBs in the cell centre region. A cell is divided into centre and edge regions based on a region partitioning CINR threshold γ_{thr}^{rp}, which can be varied as appropriate. User i is classed as being in the edge region if $\gamma_i \leq \gamma_{thr}^{rp}$, otherwise, it is classed as a centre region user. An RB in use at the edge region of a cell, as reported in ICIC signalling exchanges, is not allocated to an edge user in the neighbouring cell. However, cells can assign any available RB to a user in the centre region irrespective of the RB usage in the neighbouring cell. With the ICIC scheme, neighbouring cells avoid RBs being used in the edge region of their neighbouring cell. This ICIC implementation is illustrated in Figure 5.4 with three HAP cell, with the shaded areas highlighting the centre regions of the cells where RBs are fully reused. The size of the shaded region depends on the CINR distribution and γ_{thr}^{rp}. Considering the two UEs at the edges of cells 1 and 2, both cells can not allocated the same resource to these UEs for use simultaneously. However, cell 2 can allocate any RB to the UE in the centre region irrespective of the allocations in cell 1.

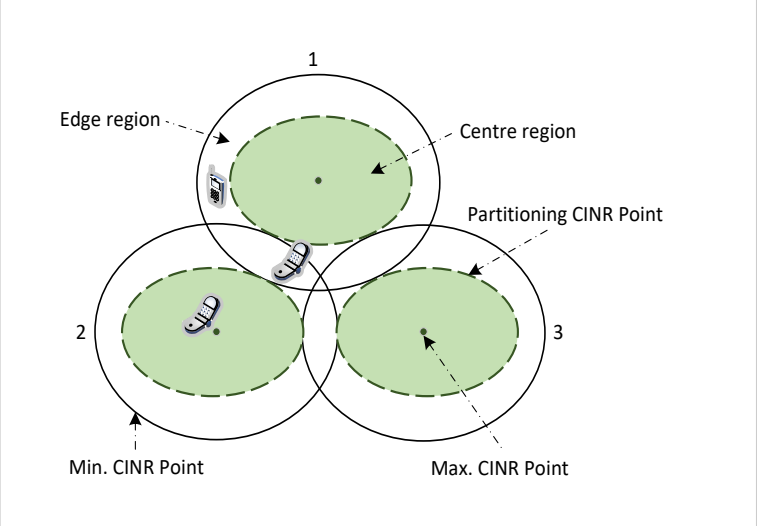

Figure 5.4: HAP inter-cell interference coordination implementation.

5.4. Performance Evaluation

A snapshot-based simulation was set-up to evaluate the performance of the proposed two-tier cellular architecture. The default simulation parameters are given in Table 4.1 of Chapter 4 in addition to those given in Table 5.1. As in previous chapters, the results of the simulation are evaluated from the cumulative outcomes of 50 independent snapshots unless otherwise stated.

Table 5.1: Key simulation parameters.

Key parameters	**Simulation Values**
Tier offset angle ξ	3.5°
System bandwidth	20 MHz
Maximum active users per cell	25
Number of resource blocks	100
Cell association CNR threshold Γ_{thr}	9 dB

As discussed earlier, the two-tier architecture is formed using Algorithm 5.1, which involves the cell-pointing algorithm proposed in Chapter 4. Users are randomly distributed over the service area of 60 km radius, and Algorithm 5.2 is used for associating the users with both a cell and tier that maximises their CNR and CINR respectively. Each cell accommodates a maximum of 25 simultaneously active users and frequency spectrum resources are shared between users in a cell at RB level. Considering the PR scheme, the number of RBs allocated to a tier is proportional to the number of its associated users as described in Section 5.3.4. However, irrespective of the scheme, the RBs allocated to each tier are fully reused by all the cells in the tier. Users requesting an RB in the HAP system are assumed to be scheduled by a centralised network controller, which is located at the HAP, parsing through the schedule one user at a time. The best available RB based on the achievable CINR is then allocated to each user by its home cell. The schedule is parsed multiple times by the controller to enable home cells to allocate extra RBs if available to active users, but with the controller prioritizing the users not allocated a RB in the previous parse. For a user to be allocated a RB, it must satisfy the minimum CINR of 1.8 dB [153], otherwise, the user is blocked.

Furthermore, the possibility of enhancing the FR and PR schemes by standard ICIC adapted for HAPs, as described in Section 5.3.4, is investigated. The coordination between HAP cells required for ICIC can be implemented in a network controller due to the possible co-location of the network entities in the HAP. It is assumed that cells exchange their RB usage through the use of regular messages sent frequently, every 20 ms for instance, which is sufficient for cells to have up-to-date knowledge of the RB usage in the neighbouring cell [175]. This frequent exchange of messages may not even be required in a centralised implementation of the HAP network where the HAP has a global view of the entire network. In addition, the possibility of exploiting other improved ICIC technologies such as the enhanced ICIC (eICIC) [176] and FeICIC [177] with its interference cancellation feature for even more enhanced performance is also discussed later.

5.4.1. Comparing one-tier and two-tier performance

This section presents the coverage, throughput, and fairness performance of the one-tier and two-tier architectures in comparison. Following user associations and RBs allocation, γ_i and T_i are evaluated for each user using (4.8) and (4.10) in Chapter 4 respectively.

Figures 5.5 and 5.6 show the surface plots of γ_i in the one-tier and proposed two-tier architectures, where cells in the one-tier architecture fully reuse the available spectrum while the spectrum is partitioned for orthogonality using PR between tiers in the two-tier architecture, with cells fully reusing the spectrum allocated to their associated tier. Both architectures have the same coverage area as indicated by the solid circle. In the two-tier architecture, γ_i is evaluated

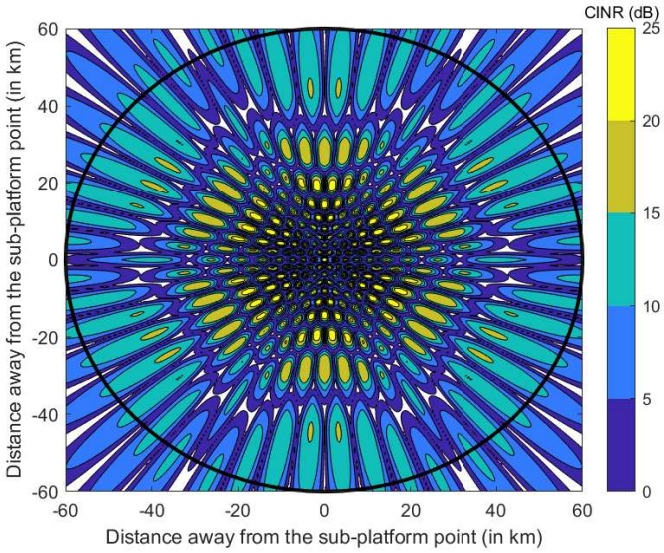

Figure 5.5: CINR footprint of a One-tier architecture.

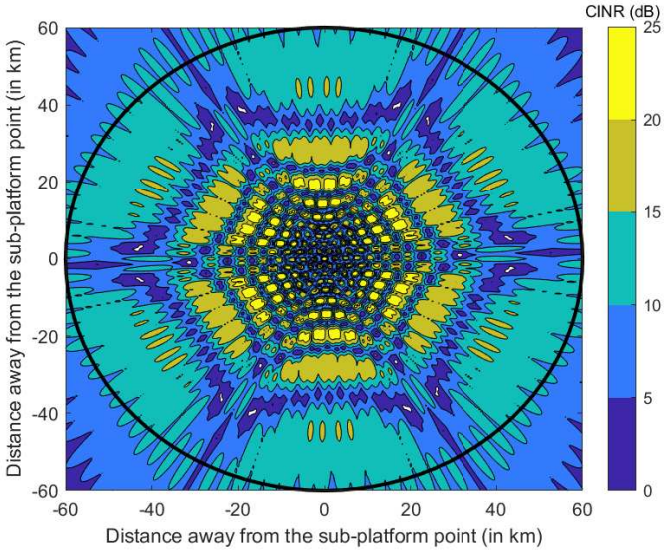

Figure 5.6: CINR footprint of the proposed Two-tier architecture.

for partial reuse without considering ITI due to the orthogonality between the tiers. Figure 5.5 shows that some significant areas within the HAP extended service area based on one-tier architecture have poor signal quality with the existence of considerable coverage holes resulting from the effects of ICI and signal quality degradation especially at cell edges. On the other hand, the two-tier architecture provides near-ubiquitous coverage with reduced edge-of-cell connectivity as shown in Figure 5.6 and confirmed in Figure 5.7, which shows the proportion of edge users to the total number of users in the one-tier and two-tier architectures for different region partitioning thresholds. Notice in Figure 5.6 that with two tiers, sufficient signal quality is received even at extended distances from the SPP, and identifying the edges of cells is not as straightforward as in the one-tier scenario. Clearly, the received signal quality is significantly increased, resulting in minimised cell edge connectivity and improved fairness.

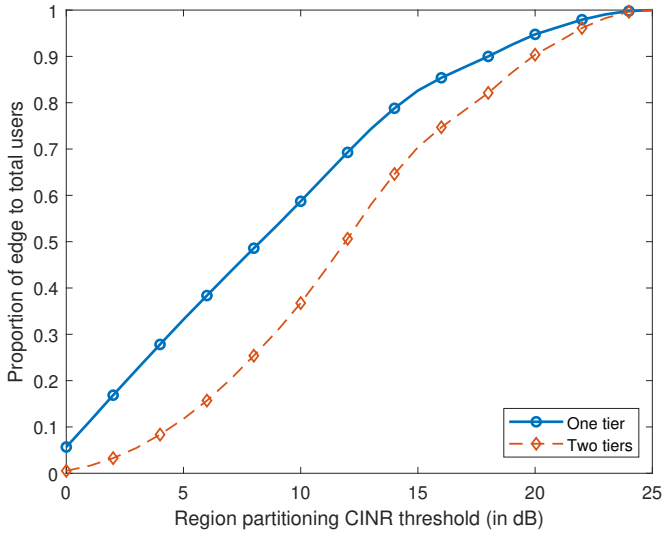

Figure 5.7: Proportion of edge to total users.

The achievable user throughput of the one-tier and proposed two-tier architectures with the FR and PR allocation schemes is shown in Figure 5.8. For the bottom 10% of the users highlighted in the zoomed-in segment, which is mainly made up of users further away from the centre of cells, the two-tier FR experiences up to 27% higher user throughput compared with the other schemes for the same probability. The percentage improvement, calculated using the points between the bidirectional arrow in the inset, is as a result of both the higher multiplexing gain resulting from each cell's ability to choose and allocate an RB for a user from the whole set of system RBs, in addition to the received signal strength improvements with the two-tier architecture. Unfortunately, this comes at the expense of most of the users suffering from considerable ITI, which degrades their signal quality resulting in throughput reduction. On the other hand, the two-tier PR completely outperforms the one-tier FR scheme and performs

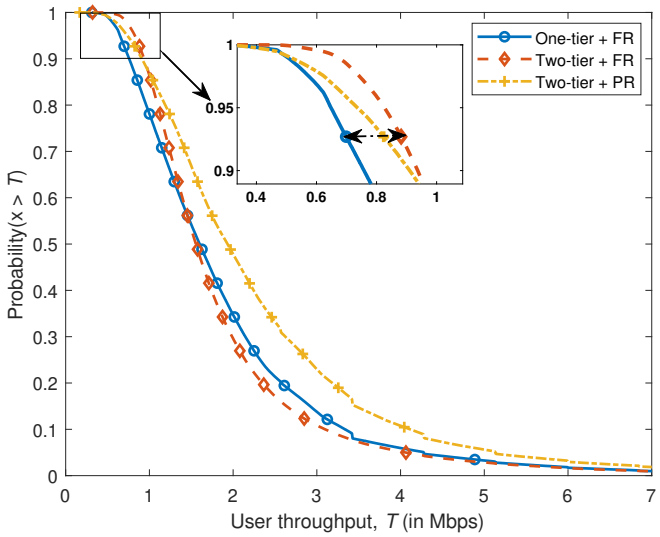

Figure 5.8: User throughput distribution.

significantly better than the two-tier FR scheme for the most part. Beyond the slightly poorer performance of the two-tier PR for the bottom 10% of the users within the inset, which is as a result of lower multiplexing gain and available RB choices, it notably outperforms the two-tier FR for the remaining 90% of the users many of which were edge users in the one-tier case.

Figures 5.9 and 5.10 show the 5th and 95th percentile user throughput with varying user densities across the extended service area. In Figure 5.9, the two-tier FR provides over 30% higher throughput at given user densities compared with one-tier FR for the 5th percentile users that are closer to the edge-of-cell, due to the higher multiplexing gain and improved received signal strength. For the edge users, ICI minimises the number of usable RBs, therefore, having a

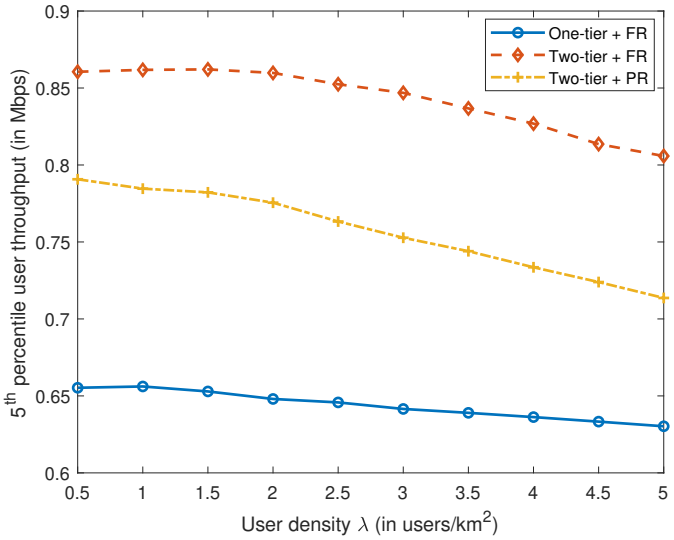

Figure 5.9: 5th percentile user throughput.

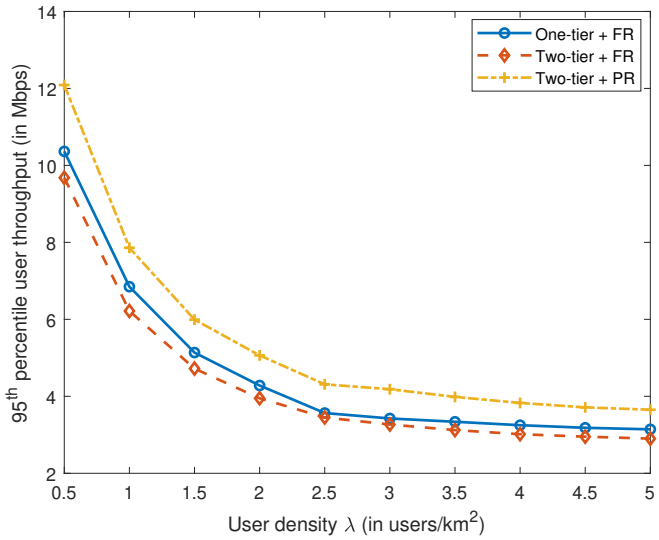

Figure 5.10: 95th percentile user throughput.

bigger set of RBs gives these users more options and higher multiplexing gain, which results in better performance. For the 95th percentile users that are closer to the centres of the cells, two-tier PR offers up to 25% higher throughput compared with two-tier FR and one-tier FR as shown in Figure 5.10. In the two-tier architecture, users at the centre of cells are minimally affected by ICI due to the steep roll-off of the HAP antenna radiation pattern, however, the resulting ITI degrades their throughput. The two-tier PR performs better because it orthogonally splits the available RBs between the two tiers, thereby effectively eliminating ITI. Unfortunately, as shown in Figure 5.10, one-tier FR slightly performs better than two-tier FR by about 7% again due to ITI. However, this should be offset against the over 30% improvement of the 5th percentile user throughput as shown in Figure 5.9. In the figures, user throughput drops with increasing user density, which happens because more users share the available RBs leading to a drop in throughput.

Figure 5.11 shows Jain's user throughput fairness index for varying user densities. At low densities, the two-tier PR results in a higher throughput fairness index because users have sufficiently enough RBs with no ITI, however, with increasing user density, the increased probability that RBs are shared less evenly between the two tiers results in a lower throughput fairness index compared with two-tier FR. This is worsened by the increasing number of edge users having fewer options of RBs with low interference to choose from. The similarity in user throughput fairness between the two-tier PR and one-tier FR at high user densities is due to the trade-off between interference and available RBs. Users in the two-tier PR experience lower interference but also have limited options of RB availability, however, those in one-tier

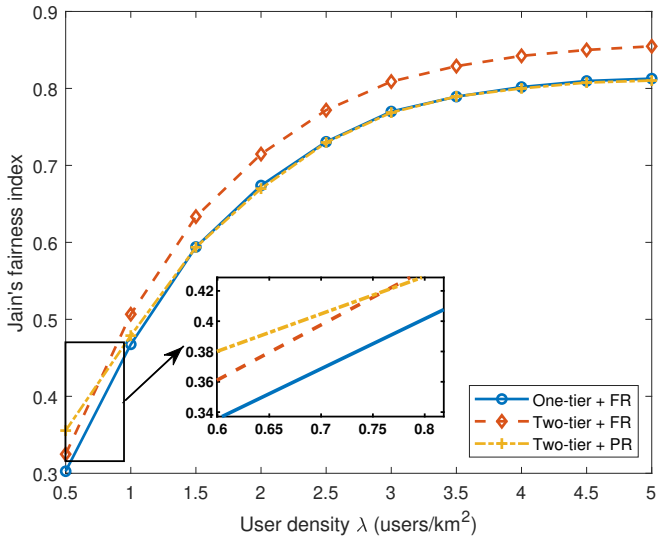

Figure 5.11: Fairness index of user throughputs.

FR experience higher interference that is counterbalanced by the potential availability of full system RBs. On the other hand, with two-tier FR, all users have an opportunity of being allocated RBs from the whole available set in addition to their enhanced received signal quality due to interference minimisation by the two-tier architecture, which results in a fairer network in comparison with the other two schemes.

Furthermore, energy efficiency is another relevant metric to discuss. The two-tier architecture improves CINR by more effectively exploiting the antenna beam pointing without increasing the transmit power. Therefore, the two-tier architecture, although requiring more power and signalling due to the increased number of cells, will expectedly have a marginally higher energy efficiency compared to the one-tier architecture. It is important to note that segments of the HAP network will be powered by renewable energy sources, e.g. HAPs with solar power, and ground-based infrastructure can be more conveniently sited close to renewable power sources. Therefore, energy efficiency is less important than with terrestrial systems, where base station towers have to be located to help maximise coverage and capacity. These are more likely to be reliant on fossil fuels to generate electricity.

5.4.2. The effect of ICIC

Having established in Section 5.4.1 that the proposed architecture with either FR or PR performs considerably better than a typical one-tier architecture, the effect of ICIC with the proposed architecture is therefore investigated in this section.

The total system throughput of the one-tier and two-tier architectures using FR, PR, and ICIC

with varying γ_{thr}^{rp} is shown in Fig. 5.12. At low γ_{thr}^{rp}, ICIC and FR have similar performance as the majority of the users are classified as centre region users, which implies that cells reuse the available RBs irrespective of the usages in other neighbouring cells of the other tier. For the two-tier FR with ICIC, the total throughput increases with increasing region partitioning CINR threshold as more edge region users emerge. Consequently, the number of users in the centre region where RBs can be reused are reduced. Thus, adjacent cells between the tiers avoid reusing RBs in the edge region with coordination. Similarly, for one-tier FR with ICIC, the total throughput increases with the CINR threshold up to a point beyond which a further increase in the CINR threshold decreases the total throughput. This is the ideal centre-edge partitioning boundary. Clearly, the ideal boundary threshold for the one-tier architecture is smaller at about 10 dB than the 25 dB boundary for the two-tier architecture. This is because there are considerably more edge users in the one-tier architecture for the same threshold, which results in the one-tier architecture peaking earlier. The region partitioning CINR threshold γ_{thr}^{rp} stops at 25 dB because majority of the users have CINR of less than 25 dB as indicated in Figure 5.7, and also, this work considers truncated Shannon bound [153] with a cap on the required CINR for maximum capacity. Despite the improvement with ICIC compared with FR, the two-tier PR still outperforms all the other schemes due to the elimination of ITI, which degrades the signal quality of users in the ICIC and FR schemes. The ITI is worsened by the contiguous nature of the architecture, which results in considerable overlap between cells. Since the total system throughput is dominated by the throughput of the centre users, the two-tier FR architecture is further affected by ITI and performs marginally poorer than the one-tier FR for this metric. Note that the constant throughput with FR and PR is a result of the fixed reuse pattern,

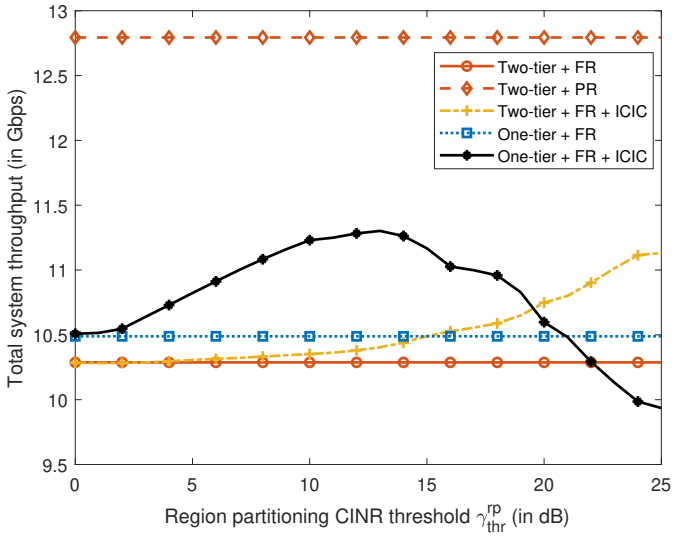

Figure 5.12: Total system throughput.

which is unaffected by the CINR threshold.

Figures 5.13, 5.14 and 5.15 show the effect of ICIC on the centre and edge user performance. Figure 5.13 shows that the edge user throughput performance is similar for PR and ICIC with FR marginally better due to higher multiplexing gain, until the breakaway point of PR at about $\gamma_{thr}^{rp} = 10$ dB and ICIC at $\gamma_{thr}^{rp} = 20$ dB. Considering the steep roll-off of power in a HAP cell, edge user performance increases with the elimination of as many interferers as possible. Since the ICIC scheme as implemented only avoids the highest interferer, edge users still experience significant interference from other cells both within and outside it's tier, which limits their performance. These edge region users at low γ_{thr}^{rp} mainly benefit from having a wide range of RBs

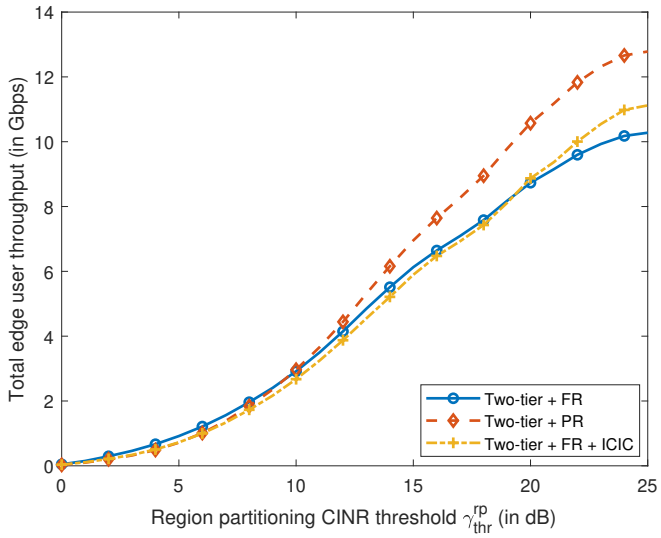

Figure 5.13: Total edge user throughput.

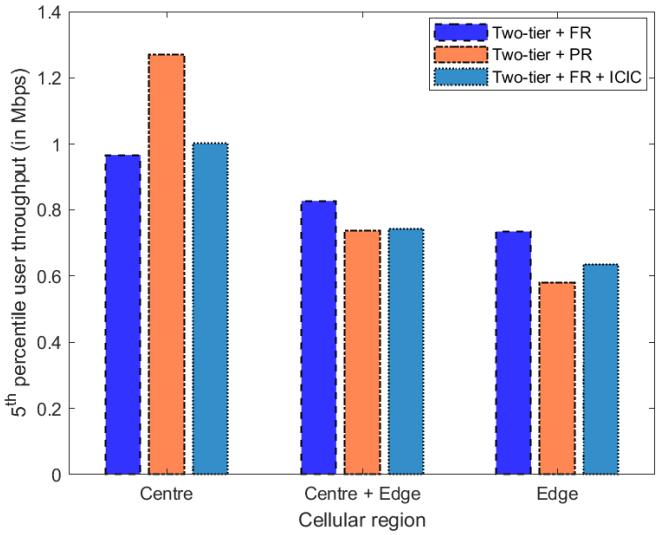

Figure 5.14: The 5th percentile user throughputs with 10 dB region partitioning boundary.

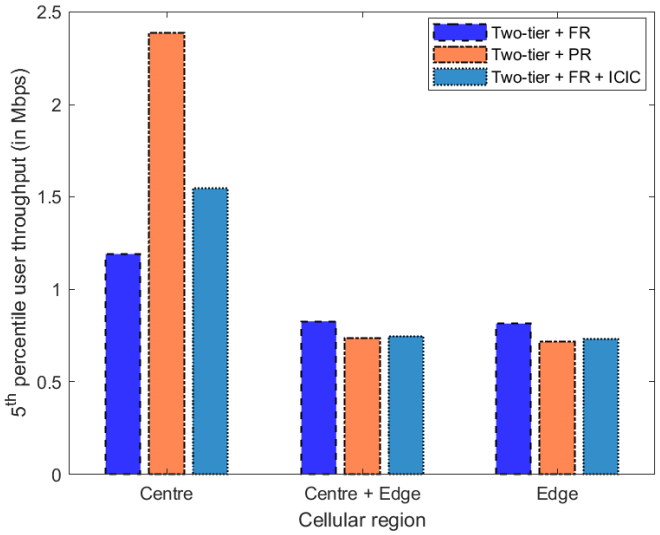

Figure 5.15: The 5th percentile user throughputs with 20 dB region partitioning boundary.

available as offered by FR. Moving towards the centre of cells with increasing γ_{thr}^{rp}, user performance is improved considerably by eliminating the highest interferer. Figures 5.14 and 5.15 show that as γ_{thr}^{rp} increases, the now fewer users in the centre region experience throughputs that are significantly higher than the edge region users because of the considerable difference in their received signal quality and the minimal ICI. However, the combined throughput drops because more users are classed as edge users and forced to share RBs. Some of these users that could ordinarily reuse RBs even with ICI due to the steep-roll off of HAP received power are forced to share RBs with other poorer edge users. It is therefore important to appropriately set γ_{thr}^{rp} to better dimension the centre and edge boundary to mitigate against the performance drop. Irrespective of the cellular region, using the two-tier PR results in superior 95th percentile user throughput as indicated in Figures 5.16 and 5.17.

Although using ICIC with the proposed two-tier architecture performs better than FR as shown in Figure 5.12–5.15, it is still limited by ITI, which is substantial considering the architecture and the significant overlap between cells of the two tiers. However, eICIC proposed in 3GPP Release 10 provides a time-domain multiplexing approach where signals of neighbouring cells in either the same or different tiers using same frequency resource are made orthogonal in the time domain. The benefits of eICIC, which significantly improves the performance of a heterogeneous network by up to 50% or over compared with the standard ICIC [176], are fully derived when there is a macro and micro cell architecture and an unequal loading between the tiers. This is not exactly the case with the proposed architecture due to the uniformity of cells in the two tiers, which could result in eICIC defaulting to the standard ICIC on equal loading

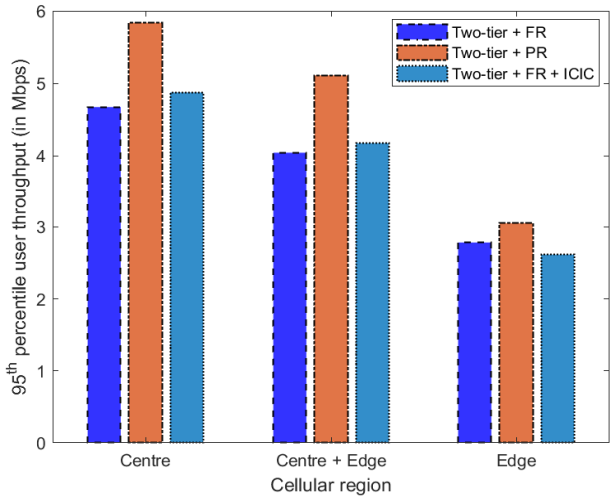

Figure 5.16: The 95th percentile user throughputs with 10 dB region partitioning boundary.

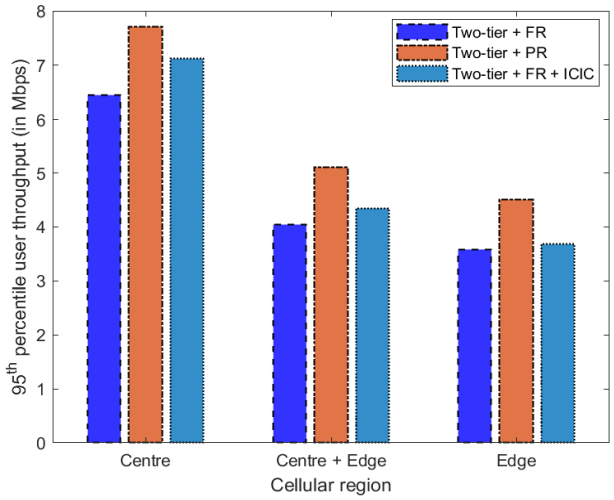

Figure 5.17: The 95th percentile user throughputs with 20 dB region partitioning boundary.

between the tiers. The FeICIC introduced in the 3GPP Release 11 provides an additional feature that can be more straightforwardly exploited by the proposed two-tier architecture. The interference cancellation feature of an FeICIC-enabled UE eliminates dominant interference possibly from a cell in different tier by using the cell-specific reference signal, thereby improving performance even well beyond the offerings of eICIC. Since interference cancellation happens at the receiver, potentially minimal increase in complexity at the HAP is expected.

5.5. Conclusion

This chapter has proposed a new two-tier cellular architecture over an extended service area. The architecture overcomes edge-of-cell problems associated with conventional cellular archi-

tectures, which have up until now significantly limited the system capacity. It is made up of two tiers with similar cellular structure and characteristics but with an offset in angle between the centres of the cells in the two tiers. Two algorithms were proposed with one implementing the two-tier architecture and the other performing user association. The performance of the two-tier architecture with full spectrum reuse (FR) and schemes implementing spectrum partitioning resource allocation (PR) were compared with a typical one-tier architecture.

Simulation results show that the HAP two-tier architecture significantly outperforms a typical one-tier architecture irrespective of the resource allocation scheme used. The architecture also significantly mitigates the impact of ICI and beam broadening on the performance of the HAP extended coverage system. Considering the two-tier architecture, the full reuse resource allocation scheme performs better than spectrum partitioning in terms of throughput, for the bottom 5–10% of users that are mostly at the edge-of-cells due to the higher multiplexing gain. However, spectrum partitioning outperforms full reuse for the remaining 90–95% of users because of the orthogonality between the tiers, which eliminates inter-tier interference (ITI). On the other hand, both schemes suffer from inter-cell interference, however, whereas inter-cell interference coordination improves the performance of the full reuse scheme by about 10% with a region partitioning CINR threshold of 20 dB, it is still outperformed by spectrum partitioning by over 12% due to ITI. Thus, ITI is the main constraint of the proposed two-tier architecture. Further benefits of the proposed architecture include minimised edge-of-cell connectivity and enhanced throughput fairness, evaluated here using Jain's fairness index.

Chapter 6. Enabling Ubiquitous 6G Communication Using High Altitude Platforms

Contents

6.1. Motivation

Some of the challenges of the HAP coverage extension and capacity enhancement approaches proposed in this thesis are addressed in Chapter 5 with the focus on solving the issues arising due to ICI and disproportionate cell sizes. However, the performance of the proposed technique and system architecture with non-uniform user densities across the network is yet to be assessed. With the increasing proliferation of devices in an unplanned and ad-hoc manner expected in next generation (NextGen) communication systems such as 5G and beyond (B5G), explicitly considering and optimising the HAP network for non-uniform user distribution is key to achieving a practical and realistic HAP system for state-of-the-art communication and beyond. Additionally, there are important aspects such as network flexibility, adaptability, and

reconfigurability, which are key requirements of B5G to 6G networks [178], that also need consideration. Therefore, the purpose of this chapter is to provide futuristic insights on how NextGen (i.e. 6G) communication can be provided using HAPs based on the proposed tier-based architecture proposed in Chapter 5 and an enhanced version of the architecture proposed in this chapter. These architectures together with the insights discussed can help satisfy the coverage ubiquity, capacity enhancement, network robustness, adaptability, reconfigurability, security and trustworthiness requirements of NextGen networks.

This chapter presents insights on how HAPs can facilitate the visions of 6G such as ubiquitous extended coverage and connectivity for UE, IoT and low-power devices, flexible and advanced backhauling, in addition to other innovative use cases such as multi-access edge computing (MEC) and the potential for far-field wireless power transfer. Considering that a crucial requirement of 6G is the flexibility to adapt and reconfigure the system in the presence of changing network conditions [178–180], a modified version of the two-tier architecture in Chapter 5, which provides optimised connectivity with dynamic cellular architecture over an extended service area, is proposed in addition to the two-tier architecture for non-uniform user and traffic densities. The proposed architecture aims to optimise the number and placement of antenna beams to better serve user hotspots, while also guaranteeing ubiquitous coverage over the entire service area for other non-clustered and sparsely distributed users. With the modified architecture, the regular two tiers provide a baseline ubiquitous coverage while an additional ad-hoc tier delivers beams of different sizes and placements for hotspots. Unlike the underlying two tiers, the layout of the ad-hoc tier is optimised dynamically for the hotspots, which can be achieved with the multi-element phased array antennas in the HAP. This work focuses on rural areas where HAPs have comparative advantages and are likely to be deployed first, although, the schemes can be implemented in other environments.

6.2. 6G HAP Scenarios

Among other things, it is envisioned that 6G networks will bring about limitless connectivity for robust, reliable, and resilient ubiquitous communication, which is crucial given the expected proliferation of devices including smaller and simpler low-power devices with applications in body area networks, internet of things (IoT), etc. [179, 181]. 6G aims to accurately localise and provide real-time connectivity to the projected 10 devices/m^2, which will involve trillions of devices globally [178]. Localisation is relevant as it would enable HAPs to quickly identify hotspots and to dynamically reconfigure and redistribute resources to effectively serve

the hotspots. Accomplishing the vision of 6G, as currently being discussed in white papers and other publications such as [178–183], means satisfying the requirements and overcoming both technical and non-technical challenges. For instance, achieving global ubiquitous coverage is a significant challenge that will be prohibitively costly for terrestrial networks alone. Satellite networks can complement but will fundamentally fall short of the stringent low latency and high capacity density requirements due to distance. Additionally, with the expected extreme densification and the eventual adoption of terrestrial cell-free architectures characterised by ad-hoc radio deployments, managing interference and coordination becomes more challenging especially for terrestrial networks, therefore, an appropriate system could be needed for management [180]. Given these limitations of the conventional systems, HAPs are essential due to their capability for extended ubiquitous coverage at reduced cost, high availability, easy and fast deployment as well as dynamic and incremental deployment, network programmability and reconfigurability [8, 14, 179]. They have the capability to continuously simplify and optimise the radio access network (RAN) architecture for enhanced performance. Since the user and traffic loading within a coverage area dynamically change over time, HAPs can form antenna beams of appropriate shapes and sizes to adequately serve the changing environment cost-effectively while maintaining ubiquitous coverage. In combination, these characteristics of HAPs will help in creating a sustainable and seamless 6G system.

6.2.1. Architecture

The commonest HAP system architectures are bent-pipe (transparent) and regenerative (full base station (BS)) architectures. In the bent-pipe architecture, only a transparent transponder is deployed in the HAP. Using reconfigurable intelligent surfaces (RIS), the transponder can be simplified further than a standard repeater or relay station. While a bent-pipe architecture results in reduced payload weight, and power consumption, it decreases flexibility and increases latency as processes that could have been handled at the HAP gets transmitted back and forth between a full BS on the ground. Alternatively, a full BS is deployed in regenerative architecture, which can take different configurations involving aggregation and disaggregation of network elements. In 5G, the baseband unit of a BS is disaggregated into distributed unit (DU) and centralised unit (CU). One configuration of the regenerative architecture involves the deployment of both the DU and radio unit (RU) at the HAP, while the CU could be deployed either on the ground as shown in Figure 6.1, in a separate HAP or satellite. An advantage of this approach is that some processing is carried out closer to the users, which reduces latency and backhaul costs. Otherwise, only the RU could be deployed in the HAP while both DU and

CU are deployed elsewhere. While this simplifies the payload at the HAP, there is a lack of flexibility at the HAP and increased latency. In rural areas, a full BS architecture is the most ideal as the long distance to core network will result in higher latency for transparent architecture. Beyond these, there could be a need for an aggregation node for backhauling remote BSs in rural areas. HAPs can be exploited to cost-effectively aggregate backhaul traffic from remote terrestrial BSs for transmission to the core network.

Additional layers are added in the rural 6G HAP architecture shown in Figure 6.1 to account for the possible direction of NextGen network architectures. For instance, 6G architecture is going towards software defined networking (SDN) where a controller in the core network manages the entire network [184]. SDN allows for easier network management, enhances scalability, and reduces both capital and operational costs [185]. In a rural 6G HAP, SDN will allow network operators to easily manage remote network elements through the HAP backhaul link to the core network. Furthermore, MEC [186] forms another layer in the architecture that can be deployed in the HAP. Its application in rural areas can provide computing and/or storage services to remotely operated factories requiring ultra-low latency in line with Industry 4.0, computation offloading for connected autonomous vehicles (CAV) travelling along rural segments of a high-way needing route computation, IoT devices for smart farming, wearables for health monitoring, etc. The detailed discussion of these technologies is beyond the scope of this thesis. There are many use cases for these architectures as discussed in Section 6.2.2.

These architectures are flexible, and depending on use case or scenario, some functionalities or network elements might be added or removed. The scale of hardware deployments is expected to reduce with the level of softwarisation anticipated in 6G through network function virtualisation and software-defined radios. The choice of a particular architecture depends on factors such as use case, costs, reconfigurability, etc. Part of the reconfigurability involves determining how and where to point antenna beams to better serve the changing usage conditions.

6.2.2. Use Cases

This section briefly presents a few of the most relevant rural 6G HAP use cases as highlighted in Figure 6.1, which are made possible with the availability of many spectrum bands for 6G.

Wide Coverage

One of the key requirements of 6G is achieving global ubiquitous communications, which can be challenging and costly with only terrestrial or satellite systems. HAPs, which can provide ubiquitous coverage over an area of at least 60 km radius as shown in Chapters 4 and

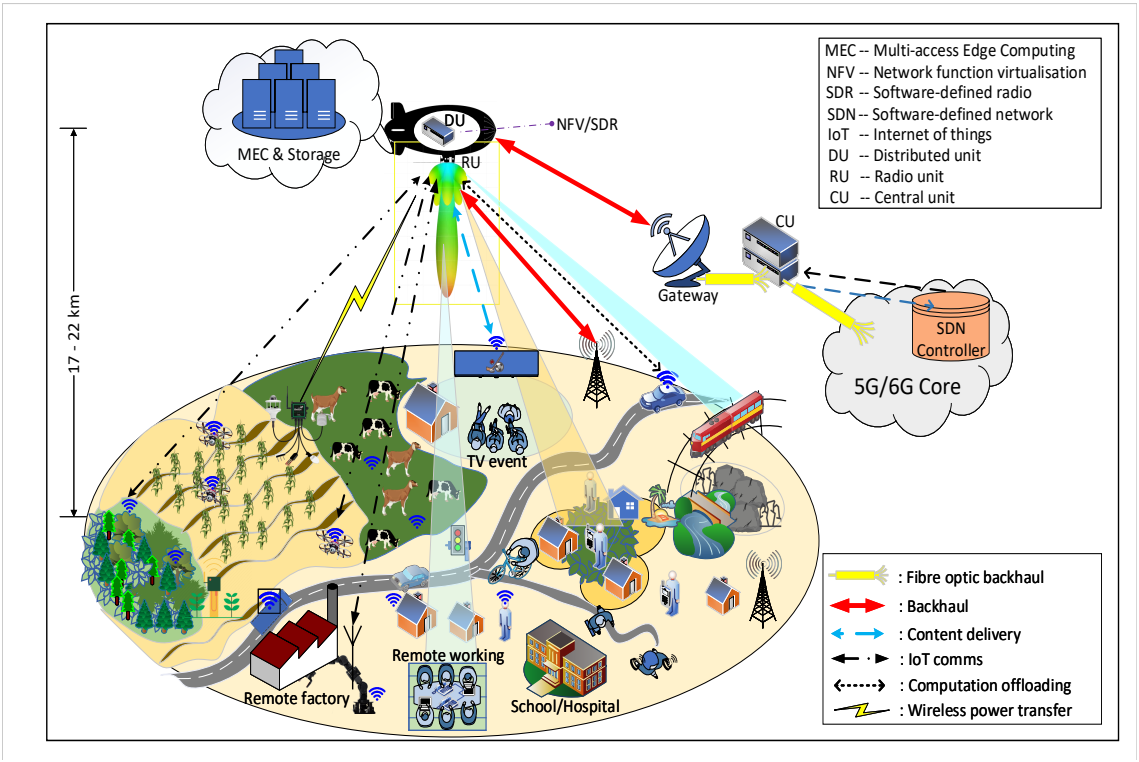

Figure 6.1: Architecture and use cases for rural HAP network of the future.

5, is a cost-effective complement. It can provide connectivity for various applications ranging from the conventional mobile UE communication, devices for precision agriculture and smart forestry, CAVs, remote working, etc. This implies coverage for potentially over 11 billion devices considering a 60 km coverage area with the expected 10 devices/m^2 [178]. As a result of the growing interest in IoT communications, 3GPP is studying the feasibility of adapting IoT networks to support non-terrestrial networks (NTN) such as HAPs in its release 17 [80].

Ultra-reliable High Capacity

Tactile internet is expected to gain traction in 6G. While it is expected that it would be used mostly in urban areas, it definitely also has applications such as remote surgery and factory automation in rural areas. With Industry 4.0, companies might decide to site their factories in rural or remote areas, and operate these factories remotely due to economies of scale or regulations. Similarly, doctors can remotely perform surgeries on patients for faster access to medical care in rural areas. These need ultra-reliable and enhanced mobile broadband with high capacity communication, which can be delivered from HAPs by using optimised beams enabled by beamforming possibly through massive MIMO antennas [183] or cooperative beamforming between multiple HAP constellations. The HAP constellation management can be distributed and autonomous through an artificial intelligence empowered wireless network [182].

Enhanced Backhauling

Backhauling remote BSs terrestrially is typically prohibitively costly. Alternatively, remote BSs can be backhauled through a HAP ideally using mm-Wave radio or optical communications directly to the ground or through satellites [187]. Both communication media can be deployed in a hybrid configuration to provide redundancy. With cost being an important factor in rural area deployments, integrated access and backhaul could even be used. Beam deployment/adjustment optimisation or even dynamic HAP placement can be used to compensate potential delay issues that may arise due to the long link length [188]. An added advantage of backhauling through a HAP is that it can serve as a natural aggregation point for many cells where traffic from those cells can be aggregated and multiplexed for onward transmission.

Computing and Storage

MEC takes information technology and cloud-computing capability of the network closer to the potential users. For the massive IoT devices, MEC in a HAP can provide resources for computing, data offloading, and storage [189]. For instance, in rural areas, devices for precision agriculture can link up with the HAP MEC to either offload data or perform computation. Similarly, near-user computation could be critical for remote factories due to the low-latency requirements. Also, CAVs would need fast processing and storage for its vast amount of data [190], and this also includes vehicles driving through rural areas.

Wireless Power Transfer

The drive towards multi-vendor open, disaggregated, softwarised, and energy efficient networks in 6G, HAPs could potentially have significantly more energy available than needed. One way to utilise the excess energy, even though it seems a bit far-fetched at the moment, is for wireless power transfer using radio signals [191]. With the proliferation of low-power devices in many hard to reach and remote areas where access for battery change is limited or lacking, far-field wireless power transfer [192] from the HAP could be exploited. The main challenge is the significant propagation losses due to distance.

These use cases, in addition to others such as content delivery and broadcasting, will enable digital inclusiveness in rural areas to mitigate the rural-urban digital divide. Section 6.3 focuses on how HAPs can facilitate 6G communications using tier-based architectures, which includes the two-tier architecture in Chapter 5 and a modified three-tier architecture that extends the functionalities of the two-tier with optimisation and reconfigurability proposed in this chapter.

6.3. Tier-based Architectures for 6G HAPs

Some of the main requirements of 6G communication include ubiquitous coverage and connectivity, high capacity, network reconfigurability and adaptability with dynamic network conditions in a cost-effective way. The one-tier architecture resulting from the beam-pointing algorithm proposed in Chapter 4, which considerably extends the achievable coverage from a HAP beyond the state-of-the-art, does not establish coverage ubiquity with some areas of insufficient coverage especially at the edge of the extended service area arising due to beam broadening and ICI. Compellingly, the proposed two-tier architecture in Chapter 5 significantly improves coverage and enhances user signal quality, which results in better spectral efficiency. It can be used as a baseline for 6G HAP communication. However, the architecture as presented was not optimised for a non-uniform user distribution and did not directly consider reconfigurability and adaptability with changing conditions such as user hotspots and clusters, something this chapter partly aims to address. Therefore, an extended three-tier architecture, which aims to address the issue of network flexibility while using the two-tier architecture for baseline coverage, is proposed in this chapter. In addition to the two-tier baseline, an ad-hoc tier is created to dynamically optimise the RAN architecture with the aim of effectively serving user clusters and hotspots as well as other users insufficiently covered by the two tiers. This chapter focuses on the tier-based architectures in general and how they can be deployed to provide 6G communication using a HAP. This includes some comparative results between the tier-based architectures and a typical one-tier architecture. Section 6.3.1 presents insights on some further potential applications of the HAP tier-based architecture for 6G communications.

6.3.1. Further Insights on Tier-based 6G HAPs

The tier-based architectures have a number of other possibilities that can be exploited to address particular network requirements, although, they are open for further research. For instance, in applications where ultra-high capacity is required, multiple HAPs could be used to create a tier for the capacity-intensive UEs through cooperative beamforming. The placement of the HAPs and constellation management can be optimised to minimise the cumulative link length, reduce latency, and significantly enhance capacity for the specific applications or UEs. Alternatively, an ad-hoc tier from a HAP implemented with mm-Wave can be dedicated to the ultra-high capacity applications while the other tiers provide coverage and sufficient capacity for other applications using other frequency spectra. With carrier aggregation capabilities of NextGen UEs, they can also simultaneously connect to more than one tier for increased capac-

ity. Furthermore, tiers can be formed incrementally based on the changing network conditions such as varying loading to ensure energy efficiency, which is another important metric for 6G.

Furthermore, the architectures can also potentially be used for addressing the issue of security and trustworthiness in 6G. For instance, users, traffic, or vendors can be separated between the tiers based on the required security level. This ensures that those requiring high levels of security and trust can be abstracted from the rest. The tiers can also be implemented in line with the concepts of network slicing, which facilitates scalability, availability, and resource optimisation, with slices (i.e. tiers in this case) forming a logical or physical network configured based on service requirements [180, 193]. The tier-based architectures are in line with the goal of network slicing to build flexible network on top of a common physical infrastructure such as a HAPs in this case. The HAP could also dynamically create or remove slices (tiers) to avoid underutilisation or over-utilisation of resources [194]. Specific use cases such as IoTs for farming and remote factory, network types, and service class for instance could be allocated to different network slices (tiers). Clearly, the tier-based architectures possess so much potentials, however, there is still the question of practicability addressed in the following section.

6.3.2. Practical Considerations

The discussions in Section 5.2.2 of Chapter 5 still apply here. In addition, the tier-based architectures fundamentally rely on the capabilities of producing a large number of antenna beams from a multi-beam-capable antenna [169, 170, 195], which is a very mature field of research. The tiers could have different dedicated smaller multi-beam antennas or created using a massive antenna array such as in [70]. With 6G expected to use multiple spectra including sub-6 GHz and mm-Wave, ultra-wideband or multi-band antenna arrays can also be considered [183]. The key point here is that the proposed tier-based architectures are practically achievable even with the current antenna system state-of-the-art, more so with the advancements such as using programmable meta-surfaces, unique geometrical designs, and vertical stacking as expected in the 6G era and proposed in [183, 196].

As stated in Chapter 5, notwithstanding how the multiple beams and tiers are delivered, proper isolation between the beams and tiers is required for orthogonality and to minimise interference. This is important for enhanced performance of the tier-based architectures as shown in Chapter 5. Interestingly, isolation is highlighted as the main challenge of network slicing in [180], which is relevant here especially if one relates network slices to the tiers or a sub-set of the tiers. It can be achieved by approaches such as ensuring different antenna polarization and intelligent sharing of spectrum resources between the beams and tiers.

6.3.3. The Three-tier Architecture Overview

The proposed three-tier architecture comprises the two-tier architecture in Chapter 5 created from the contiguous one-tier architecture resulting from the beam-pointing technique in Chapter 4. The two tiers are formed such that one tier is the offset of the other. They are typically kept regular in form and predictable to provide coverage ubiquity throughout the HAP service area. Additionally, a third ad-hoc tier, which is dynamic and follows the changing user and network conditions, is formed to provide adequate cells mainly for user clusters, hotspots, and any other users under-served in the two tiers. The HAP needs to evaluate the number of beams required to provide adequate coverage and capacity to the clusters and how the pointing of the beams can be optimised. Unlike the first two tiers, the ad-hoc tier is dependent on localising and clustering users in higher user density areas within the service area. While localisation is beyond the scope of this thesis, it is expected to be a fundamental part of 6G networks characterised by the coexistence of communication, localisation, and sensing services [197]. The proposed three-tier architecture enhances the robustness of the network by providing users with alternative orthogonal tiers.

Figures 6.2 and 6.3 pictorially illustrate the proposed architecture. Figure 6.2 shows an example cellular footprint of the tiers using a few cells. The three-tier architecture delivers better coverage and capacity especially for users under-served by the other tiers and it provides supporting beams for hotspots. Figure 6.3 disaggregates the tiers to further highlight their benefits. Notice that Tier 2 (T2) is an offset of Tier 1 (T1) as in Chapter 5, while Tier 3 (T3) is ad-hoc in

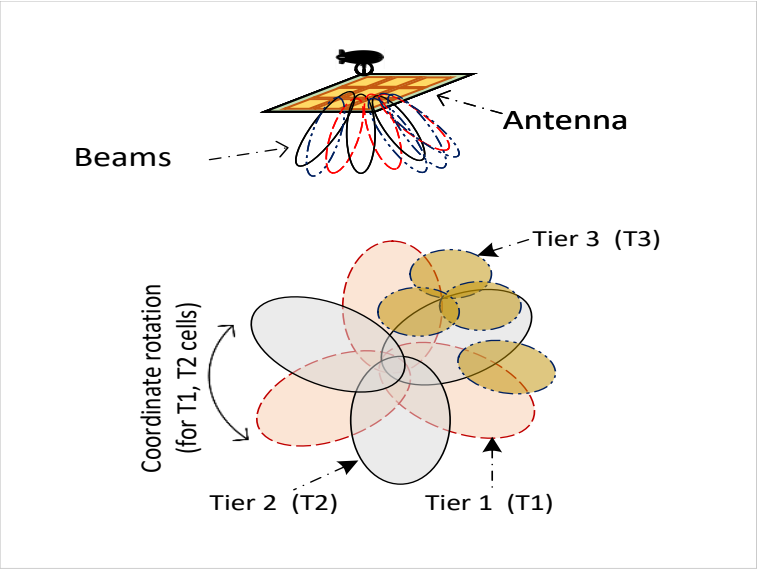

Figure 6.2: Example footprint of the three-tier architecture highlighting three tiers of contiguous tier 1 (T1) and tier 2 (T2) and non-contiguous tier 3 (T3) cells.

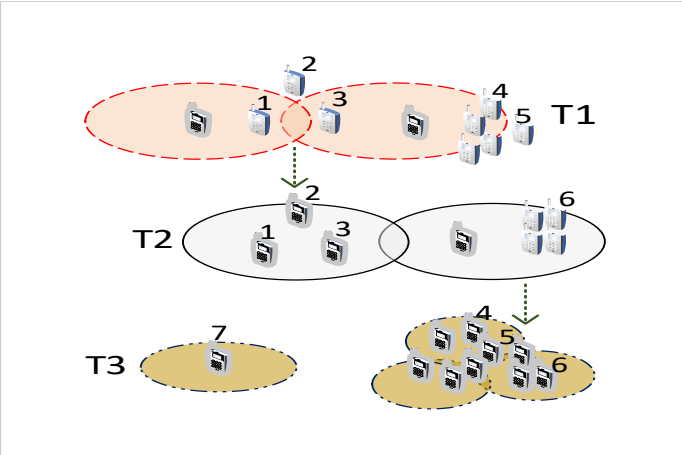

Figure 6.3: Network and user robustness of the three-tier architecture highlighting how the robustness of HAP network and users can be enhanced by the architecture.

placement targeting hotspots with the tiers proving alternative connectivity options for users to improve their receive signal quality. The darker UEs in the figure are associated with the covering cell and tier while the lighter ones move to the other tiers. Considering T1 and T2, a cell edge in one tier corresponds to a cell centre in the other tier, hence, users that would ordinary be at the cell edge or uncovered (e.g. UEs 1–3 in T1) are handed over to the centre of another cell in the other tier to enhance their signal. Also, T3 further helps in load balancing, which is also relevant due to the disproportionate cell sizes across the extended service area, by detecting hotspots and providing additional beams for more capacity. For instance, the hotspots that UEs 4–6 belong to are served in T3 for improved performance and load balancing. This flexibility and robustness further highlights the benefits of the tier-based architecture.

6.4. System Implementation

This section presents an implementation of the proposed three-tier architecture including the system model, performance metrics, tier formation, and beam-pointing optimisation. This is similar to the two-tier architecture in Section 5.3 of Chapter 5, however, the work here extends that by also considering non-uniform cluster-based user distribution and dynamic HAP RAN. The aim is to understand how the proposed tier-based architectures establish coverage ubiquity, capacity enhancement, and dynamically changing RAN to facilitate 6G HAP.

6.4.1. System Model and Performance Metrics

Mostly the same system model and performance metrics presented in Chapters 4 and 5 are adopted. As an extension in this chapter, a non-uniform cluster-based user distribution model

is implemented as described later. Furthermore, coverage probability, which is defined as the probability that a user associates to a cell in any of the tiers, is also introduced as an additional performance metric.

User Distribution

Considering a non-uniform multi-user cellular HAP network, users are distributed over the extended service area following two stages - non-clustering and clustering. In the non-clustering stage, a set of users $\mathcal{U}^a$ are generated by the Poisson point process (PPP) as described in Section 4.2.1 of Chapter 4.

In the clustering stage, K user clusters are distributed on top of the random user distribution resulting from the PPP in the non-clustering stage. K is a fixed integer variable and the position of the clusters is random. Similarly, the mean number of users $\overline{p}_k$ in each cluster $k = 1, 2, 3, \cdots, K$ is a random variable following a discrete uniform distribution over a range $[p_{min}, p_{max}]$. Assuming that the radius of the clusters are randomly distributed over a range $[r_{min}, r_{max}]$, the position of users in each cluster follow a bivariate normal distribution with mean $\mu = \overline{p}_k$ and standard deviation σ_k as a function of the cluster radius given as

$$\sigma_k = (r_{max} - r_{min})X + r_{min}, \tag{6.1}$$

where $X \in [0, 1]$ is a uniformly distributed random variable.

At the end of the clustering stage, a set of the clustered users $\mathcal{U}^b$ is obtained. Then, all the users distributed over the HAP service area is contained in set $\mathcal{I} = \mathcal{U}^a \cup \mathcal{U}^b$. It is assumed that the locations of the users in $\mathcal{I}$ is known to the HAP and were acquired through localisation techniques with good accuracy. An example of the resulting user distribution with $\lambda = 0.5$ user/km^2, $A = \pi 60^2$ km^2, $K = 5$, cluster radius range of $[2,\ 5]$ km, and cluster population range of $[100,\ 500]$ is shown in Figure 6.4 with the clusters encircled.

6.4.2. Tier Formation

This section details how the three-tier architecture is formed by starting with the underlying two-tier architecture and then building the third tier on top of the two tiers.

Tier 1 and Tier 2 Formation

In order to form the three-tier architecture, the regular tiers 1 and 2 are initially started with. Forming T1 requires the contiguous cell centre coordinates provided by Algorithm 4.1 in Chap-

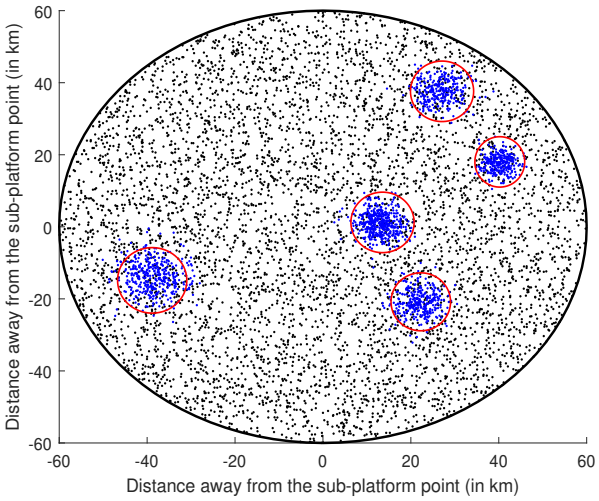

Figure 6.4: Example user distribution.

ter 4, which generates a set of coordinates $\mathcal{C}^1$ used by a HAP antenna system for beamforming to create a single tier of contiguous HAP cell structure across an extended service area making up T1. These coordinates are then used as the foundation for the formation of T2 by coordinate rotation to obtain another set of coordinates $\mathcal{C}^2$ just as described in Section 5.3.2.

The set of coordinates $\mathcal{T}^{1,2} = \mathcal{C}^1 \cup \mathcal{C}^2$ for the two tiers are used for cell and tier formation resulting in the cellular architecture shown in Figure 5.3 in Chapter 5. More discussion on T1 and T2 formation is presented in Section 5.3.2, which also includes Algorithm 5.1 used for the implementation. After the formation of T1 and T2, the HAP users are temporarily associated to the appropriate cells and tiers, to obtain the information needed to form T3 and complete the three-tier architecture.

Tier 3 Formation

Forming tier 3 builds on the existing two-tier architecture, using information obtained from the architecture that will facilitate the determination of hotspots or areas of interest for instance. Practically, T3 can be formed on top of the existing two tiers when hotspots or areas of interest appear within the operating network. Designers can specify modalities for forming T3 or even removing it, given the network requirements or specifications.

With T1 and T2 formed, users in the service area temporarily associate with an appropriate cell and the corresponding tier using Algorithm 5.2 as described in Section 5.3.3. Based on the interim association information, and in line with the concepts of radio frequency-based clustering proposed in [1], the HAP obtains the number of users associated to each cell in both T1 and T2. Then, using Algorithm 6.1, all the cells with the number of associated users

Algorithm 6.1: Tier 3 User Identification Algorithm

1: Run Algorithm 5.2 to obtain the two-tier user association matrix $B_{i,t,c}$.
2: Set I^A_{max}: Maximum number of associated users per cell.
3: Obtain a set of cells $\mathcal{L}$ from $B_{i,t,c}$ with $I^A > I^A_{max}$.
4: **if** $\mathcal{L} \neq \emptyset$ **then**
5: Initialise a new set of users $\mathcal{N} = \emptyset$.
6: **for** each cell $l \in \mathcal{L}$ **do**
7: Update $\mathcal{N}$ with the worst $I^A - I^A_{max}$ users based on γ_i from $B_{i,t,c}$.
8: **end for**
9: **end if**
10: Update $B_{i,t,c}$ by dissociating the users in $\mathcal{N}$.
11: Update $\mathcal{N}$ by including users without cell and tier association.
12: Collect $B_{i,t,c}, \mathcal{N}$: association matrix and set of users for T3.

$I^A > I^A_{max}$ are identified, where I^A_{max} is the maximum number of associated users that can be supported by a cell. For these cells covering the hotspot areas, the worst $I^A - I^A_{max}$ users in terms of γ_i representing the excesses are identified and contained in a set $\mathcal{N}$. Also included in $\mathcal{N}$ are users without cell and tier association in T1 and/or T2.

The ad-hoc T3 is formed based on a set of users $\mathcal{N}$ obtained from Algorithm 6.1. Due to the resulting irregular distribution of these users, the pointing of beams from the HAPs is optimised to provide better coverage and to enhance capacity. Beams are formed and their location optimised using SA [198, 199] for users in $\mathcal{N}$ as in Lines 12–31 of Algorithm 6.2. Prior to the optimisation, the achieved CINR for all users γ_i^1 based on an initial pointing of beams are evaluated. The choice of SA is due to its low complexity and suitability in a large search space such as the extended service area. Algorithm 6.2 performs T3 formation and beam optimisation. It takes the T3 user set $\mathcal{N}$ from Algorithm 6.1, initial annealing temperature T_Θ, and decay constant d_k as inputs. The starting beam pointing coordinates for tier 3 can be initialised with $\mathcal{C}^1$ or $\mathcal{C}^2$ or $\mathcal{T}^{1,2}$ from Algorithm 5.1 as $\mathcal{T}^3$. Using $\mathcal{T}^3$, beams are formed and users initially associate to cells maximising their Γ_i before SA starts. While the annealing temperature rounded to the lowest integer is > 0, the x- and y- axes coordinates of a randomly selected beam pointing coordinate are temporarily adjusted by normally distributed random variables with zero mean and standard deviation of 1. The adjustment is temporarily recorded in $\mathcal{T}^3$. The achievable γ_i^2 of all users with the optimised beams are evaluated and if the sum of $\gamma_i^2 > \gamma_i^1$, $\mathcal{T}^3$ is updated to permanently accept the adjustment. Otherwise, the adjustment is accepted with a given probability. The temperature is then decayed as $T_\Theta \leftarrow d_k T_\Theta$. The algorithm stops when $\lfloor T_\Theta \rceil \not> 0$ and outputs the optimised T3 optimised beam coordinate and user set for the formation of T3 in $\mathcal{T}^3$.

Algorithm 6.2: Tier 3 Formation Algorithm

1: Declare $\mathcal{N}$: A set of potential T3 associated users.
2: Set T_Θ, d_k: Annealing temperature and temperature decay constant.
3: Initialise T3 the beam pointing coordinates $\mathcal{T}^3$.
4: Form cells based on the initialised $\mathcal{T}^3$.
5: **for** each user $i \in \mathcal{N}$ **do**
6: **for** each cell $c_i \in \mathcal{T}^3$ **do**
7: Compute Γ_i using (4.4).
8: **end for**
9: Associate users to the cells maximising their Γ_i as in Algorithm 5.2 with $T = 1$.
10: Compute γ_i^1 in the associated cell using (4.8).
11: **end for**
12: **while** $\lfloor T_\Theta \rceil > 0$ **do**
13: Select a coordinate randomly (x, y) from $\mathcal{T}^3$.
14: Modify (x, y) to $(x + \Delta_x, y + \Delta_y)$ where $\Delta_x, \Delta_y \sim \mathcal{N}(0, 1)$.
15: Update $\mathcal{T}^3$ with $(x, y) -> (x + \Delta_x, y + \Delta_y)$ temporarily.
16: **for** each user $i \in \mathcal{N}$ **do**
17: **for** each cell $c_i \in \mathcal{T}^3$ **do**
18: Compute Γ_i using (4.4).
19: **end for**
20: Associate users to the cells maximising their Γ_i as in Algorithm 5.2.
21: Compute γ_i^2 in cell c_i using (4.8).
22: **end for**
23: Evaluate $\sum_{i=1}^{|\mathcal{N}|} \gamma_i^1$ and $\sum_{i=1}^{|\mathcal{N}|} \gamma_i^2$.
24: **if** $\sum_{i=1}^{|\mathcal{N}|} \gamma_i^2 - \sum_{i=1}^{|\mathcal{N}|} \gamma_i^1 > 0$ **then**
25: Update $\mathcal{T}^3$ permanently as per Line 15.
26: Set $\gamma_i^1 = \gamma_i^2$.
27: **else**
28: Update $\mathcal{T}^3$ permanently as per Line 15 with $e^{\frac{\sum_{i=1}^{|\mathcal{N}|} \gamma_i^2 - \sum_{i=1}^{|\mathcal{N}|} \gamma_i^1}{T_\Theta}}$ probability.
29: **end if**
30: Update annealing temperature $T_\Theta = d_k T_\Theta$.
31: **end while**
32: Collect $\mathcal{T}^3$, $\mathcal{N}$: T3 optimised beam pointing coordinate and and set of users.

The asymptotic time complexity of Algorithm 6.1 is the same as Algorithm 5.2 in Chapter 5, which is $\mathcal{O}(itc)$. However, Algorithm 6.2 has a complexity of $\mathcal{O}(sic)$, where i, t, c are the total number of users, tiers, and cells respectively, and s is the annealing temperature that determines the termination of the SA segment of the algorithm.

6.4.3. User Association

Each user i accesses the HAP tiered network by associating to the appropriate cell and tier. As highlighted earlier, a possible practical implementation can involve a central controller, running the algorithms and establishing the association of users in stages. Firstly, after the identification of a set potential T3 users $\mathcal{N}$ by Algorithm 6.1, the controller uses the association

information from the updated association matrix $B_{i,t,c}$, which is also an output of the algorithm, to complete T1 and T2 user connections with the rest of the users connected after T3 formation. After forming T3, user association is carried out by running Algorithm 5.2 as given in Section 6.4.3. However, in this case, the inputs are set as $T = 1$, $C = \mathcal{T}^3$, and $I = \mathcal{N}$. Once the association matrix is obtained for T3, the controller then completes the T3 user connection. In a real network deployment, this is achievable by the controller exchanging control messages with the users requesting access to establish their connection to the network.

6.4.4. Resource Allocation

Having shown in Chapter 5 that partitioning the spectrum orthogonally between the tiers to eliminate ITI results in a better performance, partial spectrum reuse (PR) schemes are considered for the sharing of RBs between the three tiers in the proposed three-tier architecture. Defining mathematically, let $R_b = \{1, 2, 3, \cdots, N_{R_b}\}$ denote the whole set of system RBs, where N_{R_b} is the total number of the overall system RBs. With PR, $\tilde{R_b^1}$, $\tilde{R_b^2}$ and $\tilde{R_b^3}$ are the sets of RBs assigned to tiers 1, 2 and 3 respectively such that $\tilde{R_b^1}, \tilde{R_b^2}, \tilde{R_b^3} \subseteq R_b$ and $\tilde{R_b^1} \cap \tilde{R_b^2} \cap \tilde{R_b^3} = \emptyset$. Also, the pairwise intersections of all three subsets of R_b are null, i.e. $\tilde{R_b^1} \cap \tilde{R_b^2} = \emptyset$, $\tilde{R_b^2} \cap \tilde{R_b^3} = \emptyset$, and $\tilde{R_b^1} \cap \tilde{R_b^3} = \emptyset$. The total system RBs contained in set R_b are shared between the tiers such that each tier is assigned $\left\lfloor \frac{i_t}{I} N_{R_b} \right\rceil$ RBs, where i_t is the number of users associated with tier t, I is total number of users in the network, and $\lfloor . \rceil$ represents rounding to the nearest integer. Once RBs have been allocated to the tiers, each cell in tier t can fully reuse all available RBs in the tier (i.e. $\tilde{R_b^t}$) with a frequency reuse of 1.

6.5. Performance Evaluation

This evaluation aims to investigate the initial performance and capability of both the two-tier and three-tier architectures to enable 6G communication from HAPs within the context of coverage ubiquity, capacity enhancements, and network adaptability. More detailed studies and evaluations of the tier-based architectures for 6G HAPs are needed though.

As in the previous chapters, a snapshot-based simulation was set-up for this performance evaluation. The default simulation parameters are given in Table 4.1 of Chapter 4 in addition to those in Table 6.1 below. The parameters are defined considering a typical rural area with clusters assumed to be village centres where residents return to from their various daily activities in areas dispersed within the HAP extended service area. Users are distributed over the service area of 60 km radius as described in Section 6.4.3, and tiers are formed using Algorithm

Table 6.1: Key simulation parameters.

Key parameters	**Simulation Values**
User density λ	0.5 users/km^2
Cluster user population range $[p_{min},\ p_{max}]$	[100, 500]
Cluster radius range $[r_{min},\ r_{max}]$	[2, 5] km
Number of clusters K	5
Annealing temperature T_Θ	5000
Temperature decay constant d_k	0.95
Maximum cell associated users I^A_{max}	25

4.1 in Chapter 4, Algorithm 5.1 in Chapter 5, Algorithms 6.1 and 6.2 in this chapter. Then, user association to cell and tier is carried out using Algorithm 5.2. The coverage and capacity performance of the system is then evaluated using the metrics described in Section 6.4.1. In line with the other chapters, the results of the simulation presented are evaluated from the cumulative outcomes of 50 independent snapshots.

6.5.1. Coverage Performance

Figures 6.5 and 6.6 show the system's coverage probability as defined in Section 6.4.1. The tier-based architectures significantly outperform a typical one-tier architecture by up to 25%. Furthermore, the tier-based architectures both achieve over 99.5% coverage probability, which is greater that the 99% required for 6G networks [200]. From the zoomed insets in the figures, notice that the three-tier architecture further improves the coverage probability compared with the two-tier architecture albeit marginally. A more significant improvement is experienced by the under-served users in the hotspots as highlighted later. Achieving an improvement towards

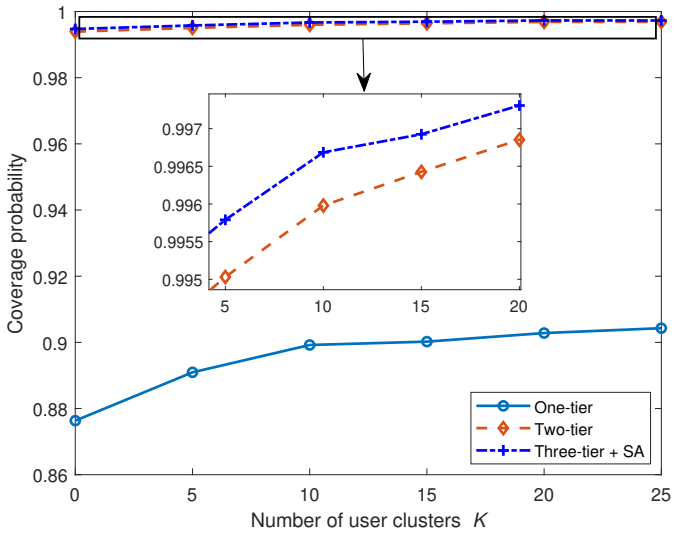

Figure 6.5: The HAP network coverage probability with varying number of user clusters.

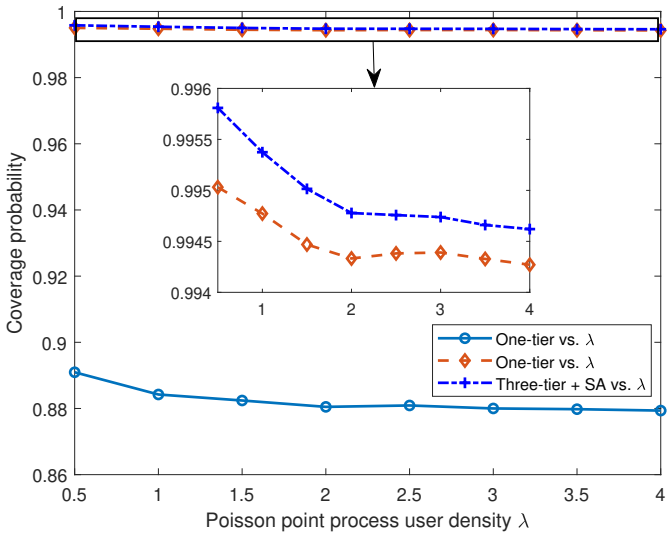

Figure 6.6: The HAP network coverage probability with varying number of user clusters and densities.

100% coverage probability by the three-tier architecture is possible with appropriate parameterisation of the simulated SA algorithm, development of better optimisation, and/or an improved user association algorithm. These are subject to further work. Noticeably, Figures 6.5 and 6.6 highlight contrasting trends where coverage probability increases with increasing number of user clusters K but decreases with increasing PPP user density λ. This is because with increasing K, there is higher probability of cluster users being in the areas with coverage given the cluster size and the minimal areas without coverage as seen in Figure 5.6, thereby increasing coverage probability. On the other hand, coverage probability drops with increasing λ because as λ increases in the homogeneous PPP, the probability that more users can be located within the uncovered area relative to similar area within the entire HAP service area increase. This is particularly so at low levels of λ.

The CDFs of user CINR are presented in Figures 6.7 and 6.8. Figure 6.7 presents the CINR CDF of some specific users not served with one-tier and two-tier architectures. It shows that the three-tier architecture significantly improves these users' γ_i by more than 2 dB compared with both one-tier and two-tier architectures. For these users, using only the two-tier architecture offers little improvement compared with using just the one-tier architecture. However, the ad-hoc tier's flexibility and adaptability in the three-tier architecture, which are important elements of 6G, offer considerable improvement as highlighted by the results in the following section. It is critical that users are effectively served irrespective of their location within the service area, which is what the three-tier architecture facilitates. Note that the improvement in performance of the under-served users does not affect the overall system CINR performance as

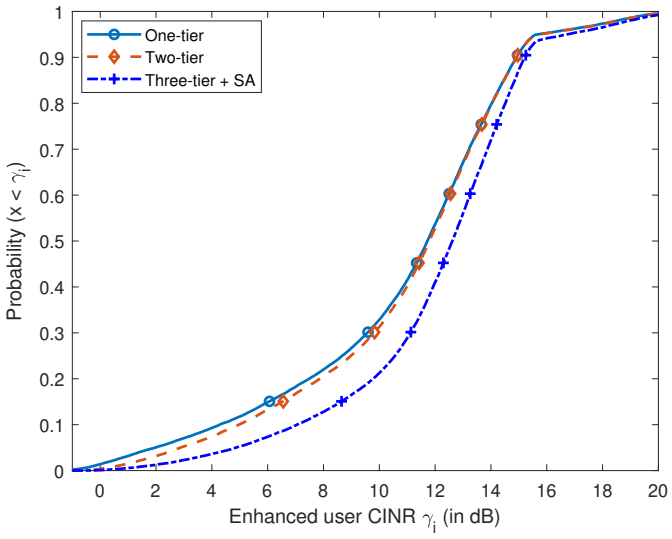

Figure 6.7: CINR distribution of the under-served enhanced by the ad-hoc tier.

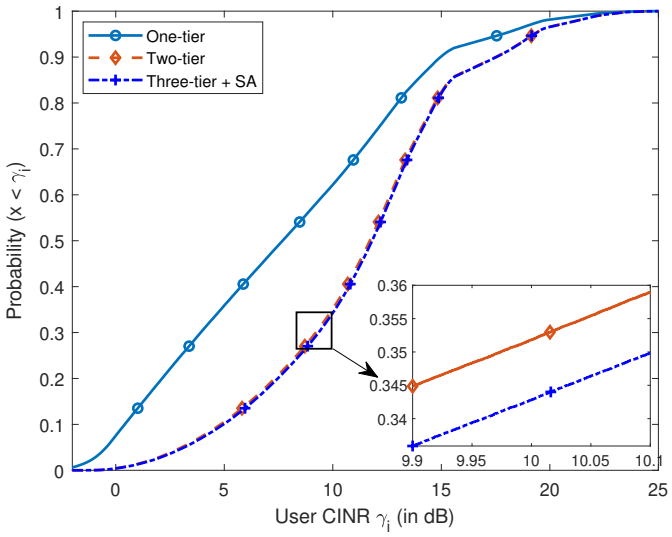

Figure 6.8: CINR distribution of all users in the network.

indicated in Figure 6.8, which shows the CDF of all user CINR γ_i in the system, highlighting further marginal improvement achieved by the three-tier architecture. Clearly, the tier-based architectures significantly improves user signal by as much as 5 dB compared with the one-tier architecture and are suitable for 6G systems at least in rural areas.

6.5.2. Capacity Performance

On capacity enhancements, Figure 6.9 shows that the sum capacity of the under-served users are considerably improved by the three-tier architecture by up to 0.2 kbps/Hz compared with the one-tier and two-tier architectures. This improvement arises from the adaptability of the

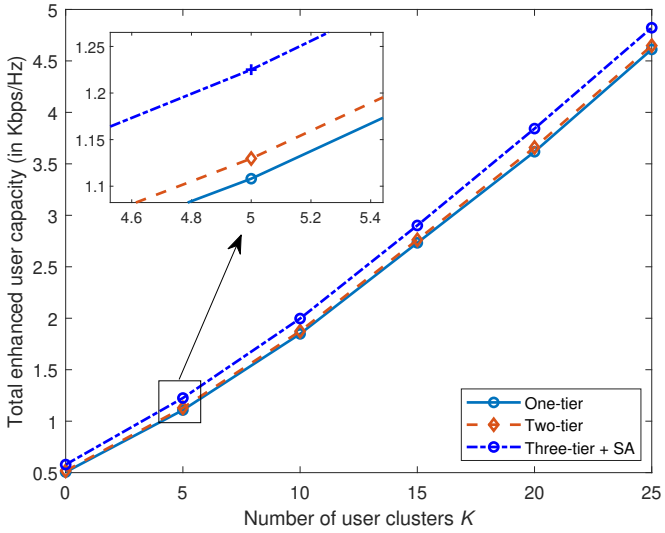

Figure 6.9: Total capacity of the under-served users enhanced by the ad-hoc tier.

ad-hoc tier to serve users that can not be sufficiently served using the one-tier or two-tier architectures. Again, this does not affect the overall system capacity. Figure 6.10 shows the total system capacity evaluated using equation (4.10) in Chapter 4. These figures highlight the capacity gains of the tier-based architectures compared with the one-tier architecture and that the gain increases more rapidly with increasing number of clusters.

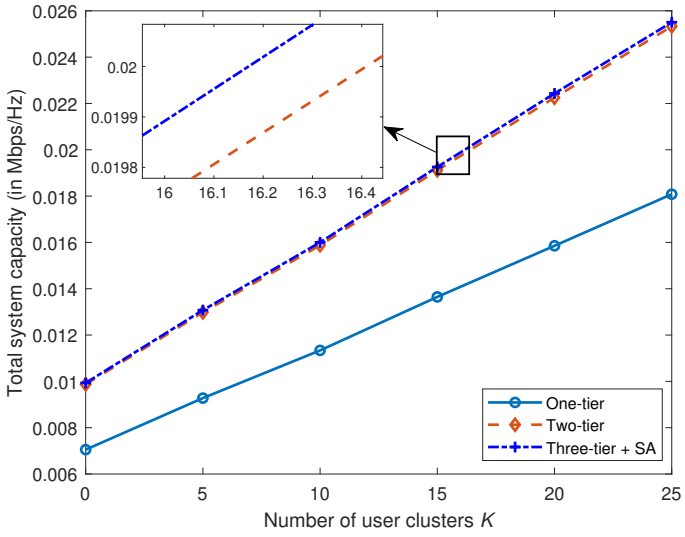

Figure 6.10: Total system user capacity.

Furthermore, Figure 6.11 presents the complementary capacity CDF of the users enhanced by the ad-hoc tier, which shows up to 8% improvement in per-user capacity. For instance, the probability that an under-served user has a capacity greater than 1.5 bps/Hz increases from about 73% to 81%. On the other hand, Figure 6.12 shows that the tier-based architectures have

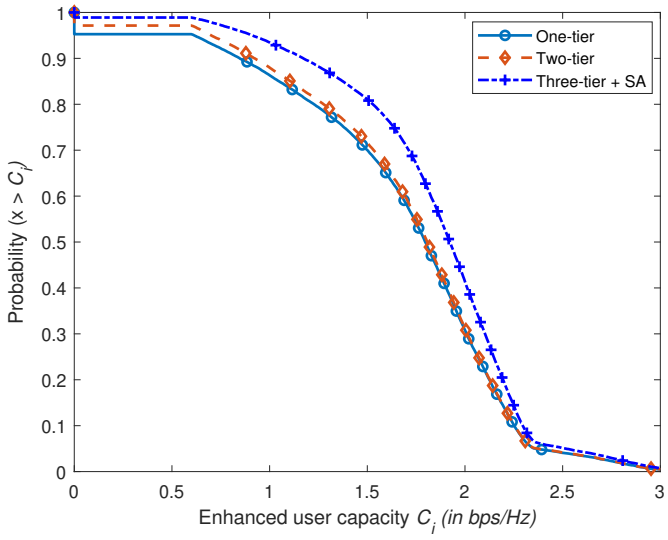

Figure 6.11: Capacity distribution of the under-served users enhanced by the ad-hoc tier.

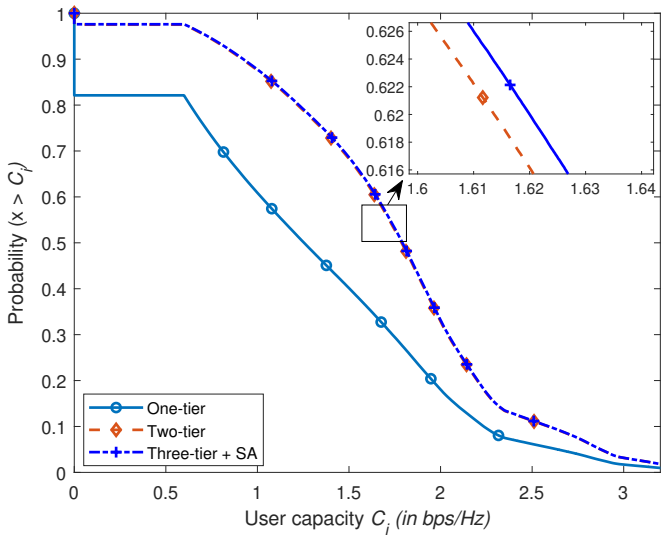

Figure 6.12: Capacity distribution of all the users.

a significant impact on the overall system capacity compared with the one-tier architecture. The probability of a user having a capacity greater than 0.5 bps/Hz is increased to approximately 98% from about 82% by the tier-based architectures, a 16% improvement.

Clearly, these results highlight the significant performance improvements by the tier-based architecture, which comprises of the two-tier and three-tier architectures, in comparison with the one-tier architecture, even with non-uniform user densities. Remarkably, the tier-based architectures achieve more than the 99% coverage probability required for 6G networks. However, in further comparison with the two-tier architecture, the three-tier architecture offers some extra but marginal performance enhancements. The marginality is possibly as a result of the

limitations of SA algorithm used for the beam pointing optimisation. As a metaheuristic algorithm, SA can only approximate the global optimum, therefore, the accuracy of the approximation is low in a large search space such as the extended HAP service area. SA is also very parametrised, therefore, obtaining the optimum parameters would require a parameter sweep over a large space and multiple dimensions, which may be computationally complex. Therefore, the marginality in the ad-hoc tier performance improvement could be as a result of inadequate parametrisation as implemented. More work is needed to understand how SA or other optimisation algorithms can be parametrised appropriately and used in this scenario, and how close they can get to a global optimum. A more appropriate parametrisation and implementation of simulated annealing, or using other algorithms for beam placement optimisation and/or user association can potentially result in a more considerable improvements overall. However, note that in some cases, a significant improvement of the overall system performance may not be the objective. The optimisation objective could be to ensure that a particular group of users are better served by the ad-hoc tier. Based on this, there is a trade-off between enhancing the performance of this set of users at the cost of the overall system performance.

Furthermore, the ad-hoc nature of the third tier that optimises the RAN based on network conditions is a significant advantage over the other architectures since making the RAN dynamic is one of the visions of 6G. For instance, deploying a given number of cells with different cell sizes at different locations based on traffic and the radio environment is greatly desirable. It also ensures that cells are not deployed in places where they are not needed, thereby increasing the energy efficiency of the network, which is desired by MNOs especially given than RANs consumes significantly more energy compared with other parts of the overall mobile network [137]. Given all other benefits of HAPs such as cost-effectiveness, ease of deployment, and flexibility, the HAP tier-based architecture offers a good potential for delivering cost-effective 6G communications especially in rural areas with minimal or no coverage. In fact, a tessellation of coverage from multiple HAPs delivering the proposed architecture could provide regional or even nationwide coverage. More work on the potential of the tier-based architecture for 6G could be undertaken. For instance, it would be good to investigate the energy efficiency of the architectures, which is one of the key metrics of NextGen communication systems. Furthermore, the possibility of users connecting to all tiers simultaneously with carrier aggregation to enhance capacity is also another extension of the work that will facilitate the delivery of NextGen communication using the tier-based architecture. However, as discussed earlier, better algorithms for beam placement optimisation and/or user association would result in a more considerable improvements overall.

6.6. Conclusion

This chapter has discussed the possibility of using HAP tier-based cellular architectures, which includes the two-tier and the modified three-tier architectures, to enable 6G communication from a HAP. Possible system architectures including bent-pipe and regenerative were briefly discussed. Some use cases within the context of rural area deployment were also briefly presented. Furthermore, a three-tier architecture, which is an extension of the two-tier architecture was proposed. The proposed architecture introduces an additional ad-hoc tier that optimises beam placement for hotspot and cluster coverage. The performance of the proposed tier-based architectures are evaluated with non-uniform user distribution and compared with a typical one-tier architecture.

The simulation results show that the tier-based architectures achieve a coverage probability of over 99.5%, which is greater than the 99% requirement for 6G networks. Additionally, it is shown that the coverage probability increases with increasing number of clusters, but decreases with increasing user density of random distribution following Poisson point process. In comparison with the single tier, the tier-based architectures are shown to improve user CINR by as much as 5 dB, while the three-tier architecture further improves CINR of some underserved users by as much as 2 dB. Additionally, the total system capacity of the tier-based architectures perform better than the one-tier architecture at a more rapid rate with increasing number of user clusters. Further comparing the two-tier and three-tier architectures, the latter offers some additional improvement which could be further enhanced through the use of a better beam placement optimisation algorithm and tier 3 user selection and association. Finally, more work is needed to understand how the proposed tier-based architectures can facilitate next generation communication systems.

Chapter 7. Conclusions and Further Work

Contents

7.1. Conclusions

HAPs have the potential to provide cost effective rural and remote coverage, with careful design of their antenna and energy management systems. Their operation allows direct connection to the user equipment in the same way as a terrestrial system, unlike satellite based connectivity, potentially allowing much more flexible coverage than terrestrial systems. The work in this thesis has demonstrated how extended contiguous coverage can be delivered from a HAP over an area of at least 60 km radius, using a multi-element planar phased array antenna with significant flexibility and reconfigurability. It aimed to maximise the utility achievable from a HAP especially in areas with minimal or no wireless coverage such as remote or rural areas. This has allowed for the development of adaptable tier-based architectures that can facilitate HAPs to credibly deliver ubiquitous coverage and enhanced capacity over the extended coverage area for the first time. Interestingly, it is a step towards achieving the global coverage ubiquity and enhanced capacity vision for 6G communication systems. Furthermore, a

feasibility analysis and energy management of a solar-powered HAP for long duration wireless communications mission at a particular latitude and over a given area has been presented. It has shown that the feasibility is mostly dependent on the HAP's payload capability and wingspan, which directly influences the amount of energy that can be harvested for operation at a particular location.

Generally, the research in this thesis has been guided by the hypothesis stated in Section 1.1 of Chapter 1. In order to prove the hypothesis, the work presented here considered both the theory and practicality of extending wireless coverage and enhancing capacity from a HAP beyond the state-of-the-art, given the constraints and limitations of HAPs highlighted above. Firstly in Chapter 3, the feasibility of some state-of-the-art example HAPs for long mission duration wireless communications over an extended service are was evaluated, with different models of energy collection and consumption of the platforms proposed. Interestingly, the evaluation has shown how the feasibility varies depending on the operation configuration, latitude of operation and platform characteristics such as wingspan and payload carrying capability. Next in Chapter 4, an algorithm which allows for an extended contiguous coverage from a HAP was proposed as well as a theoretical analysis on predicting the limits of coverage extension, with derived expressions for estimating the achievable capacity and spectral efficiency of any cell pointed within the HAP service area. The algorithm enables a HAP to significantly extend its coverage area beyond the state-of-the-art. Then in Chapter 5, a tier-based HAP cellular architecture, which ensures ubiquitous coverage over the extended service area was proposed. The proposed architecture mitigates against inter-cell interference especially at the edges of the HAP cells and extended service area. Finally in Chapter 6, insights on the possibility of enabling 6G communication over a wide area using tier-based architectures were presented. This includes how the architecture can satisfy some of the requirements of B5G networks such as coverage ubiquity, capacity enhancement, and robustness under changing network conditions. A modified three-tier architecture, which extends the earlier two-tier architecture by introducing an extra ad-hoc tier for adaptable coverage and capacity was also proposed.

As a whole, the focus in this thesis has been predominantly based on rural area application scenarios, which is where the comparative advantages of HAPs are more pronounced. However, all algorithms, techniques and approaches developed here can be directly extended to other environment types with minimal modifications if required. A more detailed chapter-by-chapter discussion of the original contributions of this thesis towards the extension of coverage, enhancement of capacity, development of adaptable and robust HAP cellular architectures is given in the following section.

7.2. Original Contributions

7.2.1. Energy management and feasibility study for extended coverage

Chapter 3 shows how the design and operation of solar-powered aircraft might affect the feasibility of providing wireless communications over varying areas with respect to power and weight constraints of the aircraft. This involves showing the necessary calculations required to highlight the methodology of using HAPs, the lack of which in a wider sense has led to the failure of some HAP projects. While significant work on the design of HAPs is ongoing in industry, there is a lack of details of the specific engineering involved in open literature as much of that is proprietary, an issue this chapter attempts to fix. Furthermore, different mathematical models of energy harvesting and consumption for a solar-powered aircraft-based HAP suitable for wireless communication missions are developed. Bringing together the necessary calculations for the first time in the public domain, the feasibility of the current state-of-the-art HAPs for wireless communications service delivery over an extended coverage area and at different geographical locations is analysed using three example platforms. It is shown that with current HAPs, with wingspans ranging between 25–35 m, wireless communications can be provided for a duration of 15–24 hours per day depending on the platform, latitude of operation, operation configuration, and radius of coverage. However, continuous operation for a whole day with current platforms with smaller wingspans especially during the days of the year with limited daylight is infeasible. To achieve that, the use of airships or aircraft with alternative energy sources are necessary. It is also highlighted that large-wingspan aircraft of at least 60 m provides sufficient gains in payload carrying capability since doubling wingspan increases this by up to a factor of 6, thereby enhancing feasibility. Interestingly, Hawk30 from HAPSMobile Inc. and the hydrogen-powered HAP from Stratospheric Platform Limited both fit this with 60 m and 78 m wingspans respectively. This work has now been published in the MDPI Electronics Journal.

7.2.2. Contiguous coverage extension algorithm and theoretical analysis

Chapter 4 proposes a recursive beam-pointing algorithm, which forms cells over an extended coverage area using multiple beams from the HAP. The algorithm provides a set of beam-forming coordinates used to create a contiguous cellular architecture that provides coverage over an area of 60 km radius, which is at least double the 30 km prevalent in the literature. It compensates for beam broadening by appropriately adjusting neighbouring cell placements, providing flexibility on cell size variations and the level of overlap needed between the cells,

and minimising the level of interference arising due to the overlap of neighbouring antenna beams main lobes. The chapter also provides a theoretical analysis, which is based on some derived mathematical expressions, on estimating the achievable limits of coverage extension. One of the derived closed-form expressions is for estimating the area covered by the footprint of a HAP beam on the ground when pointed at any given distance away from the sub-platform point. This is useful for estimating the number of cells required to provide adequate coverage over a given service area. Other expressions for the estimating average per-user capacity, average spectral efficiency, and average area spectral efficiency of a cell given its distance from the sub-platform point are derived. The proposed algorithm and theoretical expressions derived in this chapter are relevant in designing and assessing the performance of a HAP system over an extended service area. Part of this work was presented in IEEE Vehicular Technology Conference in 2019. An extended version of the conference paper consisting of the work in this chapter has been submitted to the IEEE Transactions on Aerospace and Electronic Systems.

7.2.3. Two-tier architecture for coverage extension and capacity enhancement

Up to now, the state-of-the-art in HAP architectures has focused on a one-tier architecture. The work in Chapter 5 shows how a two-tier architecture can mitigate the effects of ICI and beam broadening on cellular service provisioning from HAPs, while ensuring ubiquitous coverage over an extended area. The proposed architecture, which forms two tiers of cells over an extended area with users having the flexibility of connecting to the best cell and tier based on a given metric, significantly improves coverage and capacity by minimising edge-of-cell connectivity and coverage holes. It significantly enhances user CINR, mitigates the effect of ICI, and increases fairness in the system. The chapter proposes an algorithm that implements the two-tier architecture by forming two independent tiers of contiguous cells with one tier as an offset of the other in angle. The proposed algorithm gives users the choice of associating to the best cell from the two tiers while also facilitating even loading between the tiers. Another algorithm for user association with the cell and tier that maximises its CNR and CINR is also proposed. Finally, the chapter investigates different resource and interference management schemes, showing how full spectrum reuse (i.e reuse factor of 1), spectrum partitioning, and ICIC perform with the proposed two-tier scheme. Here, ensuring orthogonality between the tiers by partitioning the spectrum results in better performance. Monte Carlo simulation results show that the proposed two-tier architecture significantly outperforms the typical one-tier architecture developed in Chapter 4. This work is currently under revision for resubmission in IEEE Transactions on Mobile Computing.

7.2.4. Enabling ubiquitous 6G communication with tier-based architectures

To date, the main focus in the literature has been on 5G or earlier generations. Chapter 6 of this thesis presents futuristic insights on how HAPs can be vital for delivering 6G NTN, and central is the proposed tier-based architectures. This includes the two-tier architecture proposed in Chapter 5 and a modified dynamic three-tier architecture that allows variable capacity densities to be delivered. The tier-based cellular architectures can be used to satisfy the coverage ubiquity, adaptability, flexibility, and reconfigurability requirements of 6G, where alternative solutions such as satellite and terrestrial systems will be insufficient or too costly for both operators and users. The modified three-tier architecture is based on two tiers of contiguous cells in the structure and an additional tier of ad-hoc cellular structure. It achieves both ubiquitous coverage with over 99% coverage probability and dynamic beam pointing to appropriately serving changing users or traffic conditions in the network. For the dynamic beam pointing, a beam pointing location optimisation based on a modified simulated annealing algorithm with structured initialisation is proposed. Furthermore, the work in this chapter also present insights on a possible 6G HAP network implementation and deployment including use cases, applications, and architecture. Discussions on how the proposed architectures can address some of the issues of future networks such as security and trustworthiness, adaptability, and its correlation with network slicing are also presented. Unlike Chapters 4 and 5, which focus on a uniform user distribution, results here show how the tier-based architectures can be used to enhance coverage and capacity under variable user densities. This work will later form a publication to be submitted to IEEE Transactions on Mobile Computing.

7.3. Recommendations for Further Work

Some recommendations for further work on the areas explored in this thesis are given in this section. They predominantly involve possible extension of the proposed techniques and approaches and their applicability to a wider range of scenarios both within and the big picture beyond the scope of this work.

7.3.1. Mobility and Handover

Logically, in extending the work presented in this thesis, the natural next step will be to consider user mobility. Given mobile users, it will be interesting to understand how the system reacts and the best way to service them. This would involve investigating whether mobile users are best served using beam tracking and steering or through frequent handovers. The ultra-low

latency requirement could play a role in deciding the better approach. It would be interesting to see how handover between cells in Chapter 4, cells and tiers in Chapters 5–6 could be implemented. Machine learning and artificial intelligence could form a vital part by providing mobility prediction that could aid both beam steering and handover. Also, further work could be carried out on how the coverage and capacity of the system changes if the HAP is not positioned at the centre of the service area. These considerations would all result in significant extensions of the work this thesis, which would be interesting to observe.

7.3.2. Advanced Beamforming Techniques

The architectures presented throughout this thesis are based solely on conventional beamforming from a planar phased array antenna to form multiple cells as highlighted in Chapter 4. However, with the push for cell-free architectures and increased spectral efficiency in future communication systems, different beamforming techniques could be required. Spatially disjointing users and simultaneously serving them would lead to increased capacity and spectral efficiency [21]. Therefore, multi-user MIMO and cell-free massive MIMO techniques such as zero-forcing beamforming and minimum mean square error could provide alternative ways of exploiting multi-element phased array antennas. It would be useful to understand the effect of such techniques on the performance and complexity of a HAP extended wireless communication system as proposed in this work.

7.3.3. Advanced Antenna Designs

Instead of the 2D planar phased array antennas considered here, the performance of proposed architectures with unique geometrical antenna designs and vertical stacking [183] could be investigated. This is particularly relevant in the extended coverage scenario where planar arrays have very small elevation angles at the edge. 3D-based antenna designs such as hemispherical array antennas or a combination of antennas with different shapes, sizes, placement, and orientation can potentially offer improvements in the system performance as well as increased elevation angle at edge of coverage.

7.3.4. Dynamically Varying HAP Placement

An important parameter in a HAP wireless communication system is the link length, which directly influences both the QoS and quality of experience of a user. Fundamentally, shorter HAP links result in lower latency and faster speed. The work presented in this thesis is based on the assumption that the HAP is quasi-stationary. On the other hand, dynamically varying

the placement of the HAP to reduce either individual or combined link lengths is worth investigating. For instance, considering the clusters in Chapter 6, the HAP could be positioned such that it is effectively closer to the clusters to improve their coverage and capacity. On the other hand, such systems could be limited by the computational complexity involved in continuously optimising the HAP placement. Additionally, the increased energy consumption resulting from continuous optimisation and movement of the platform can potentially be limiting considering that the HAP is fundamentally energy constrained. There is scope for investigating the changing platform placement scenario, including its effect on energy efficiency, which is a key metric in next generation networks.

7.3.5. Multiple HAP Constellations

With increasing requirement of ubiquitous coverage, ultra-high capacity and spectral efficiency in NextGen systems, which challenges what can be delivered from a single HAP, the deployment of multiple HAPs forming constellations as being done with satellites is expected. Understanding how the architectures proposed in Chapters 4–6 could be delivered from multiple HAPs through cooperative beamforming using distributed massive MIMO to significantly enhance capacity would be interesting. With respect to coverage, multiple HAPs can achieve full ubiquitous regional coverage, which is in line with the 100% coverage vision of B5G networks. This includes the capability of delivering up to a countrywide coverage using HAPs. However, to achieve these, a number of issues such as constellation design and inter-HAP coordination need to be addressed. Designing new HAP constellations or evaluating how existing designs can be applied with the architectures is subject to further work. Similarly, the coordination between the HAPs in a constellation is an important area of research. HAPs can coordinate in a distributed manner or there could be a central orchestrator in charge of coordination and management functions. Machine learning and artificial intelligence techniques would be relevant here. Unlike in previous studies where multiple HAPs are considered within the context of regional coverage [201], capacity [202], or diversity [51], the regional coverage aspects could be satisfied by the individual footprints of the HAPs in one tier, while the capacity and diversity aspects could be satisfied in another tier through cooperative beamforming by the HAPs in the constellation.

7.3.6. Optimised Tier-based Architecture

The tier-based architectures in Chapters 5–6 have been shown to improve coverage, capacity, and robustness of the network. However, both the tier formation and user association can be

optimised for better performance. To further the work in this thesis, a two-tier based architecture with one fixed tier and one ad-hoc tier is worth investigating. Tiers could be formed and removed based on the changing environment or usage conditions with measurable effect on network performance and energy efficiency. This could follow the ideals of network slicing where slices are created for specific purposes in the network. It would be worth investigating the impact of such approach to the system. Additionally, based on the requirement for each tier, the number, shape, placement of beams can be continually optimised for either coverage, capacity, or both. On the other hand, the association of users to tiers could be improved. For instance, rather than permanently associating any user to any particular tier, users could dynamically change association based on the condition of the radio environment and the experienced QoS. Alternatively, users could associate to all available tiers simultaneously, following the multi-connectivity desire in 6G, using technique such as carrier aggregation. Understanding how these improvements impact on the overall system performance could be interesting.

7.3.7. Integrating HAP in Vertical Heterogeneous Networks

One of the visions for next generation systems is having a vertical integrated network of both terrestrial and non-terrestrial networks. Based on the conditions, user connections can be switched between either. Understanding how HAPs will be integrated, including the required signalling and protocols, would be important. Furthermore, such integration emphasises the issue of cross-interference. Typically, there would be need to manage interference between the different segments. On that note, radio environment maps could be exploited for managing both interference within and between segments. The use of computer vision, machine learning or artificial intelligence techniques could be useful.

Acronyms

3GPP 3rd Generation Partnership Project

5G 5th Generation

6G 6th Generation

B5G Beyond 5th Generation

BS Base Station

CAV Connected Autonomous Vehicle

CINR Carrier-to-Interference-plus-Noise Ratio

CNR Carrier-to-Noise Ratio

CoMP Coordinated Multipoint

CU Centralised Unit

DU Distributed Unit

eICIC Enhanced Inter-Cell Interference Coordination

FeICIC Further Enhanced Inter-Cell Interference Coordination

FR Full Spectrum Reuse

FSO Free Space Optical

HAP High Altitude Platform

HAPS High Altitude Platform Station

HetNet Heterogeneous Network

HFC Hydrogen Fuel Cell

ICI Inter-Cell Interference

ICIC Inter-Cell Interference Coordination

INR Interference-to-Noise Ratio

IoT Internet of Things

ITI Inter-Tier Interference

ITU-R International Telecommunication Union - Radiocommunications

LTE Long Term Evolution

LoS Line of Sight

MEC Multi-access Edge Computing

MIMO Multiple Input Multiple Output

mm-Wave Millimetre Wave

NextGen Next Generation

NFV Network Function Virtualisation

NTN Non-Terrestrial Network

PAA Phased Array Antenna

PPP Poisson Point Process

PR Partial Spectrum Reuse

QoS Quality of Service

RAN Radio Access Network

RB Resource Block

RIM Resource and Interference Management

RU Radio Unit

SA Simulated Annealing

SDN Software-Defined Networking

SPP Sub-Platform Point

SPPC Sub-Platform Point Cell

SWaP Size, Weight, and Power

T1, T2, T3 Tier 1, Tier 2, Tier 3

UAV Unmanned Aerial Vehicle

UE User Equipment

References

[1] M. D. Zakaria, "The coverage, capacity and coexistence of mixed high altitude platform and terrestrial segments," Ph.D. dissertation, University of York, Jul. 2019. [Online]. Available: https://etheses.whiterose.ac.uk/25596/

[2] F. Qamar, K. B. Dimyati, M. N. Hindia, K. A. B. Noordin, and A. M. Al-Samman, "A comprehensive review on coordinated multi-point operation for LTE-A," *Computer Netw.*, vol. 123, pp. 19–37, 2017.

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

[4] E. Matricciani, "Space communications with variable elevation angle faded by rain: Radio links to the sun–earthfirst Lagrangian point l_1," *Int. J. Satell. Commun. Network*, vol. 2016, no. 34, pp. 809–831, 2016.

[5] C. A. Balanis, *Antenna Theory: Analysis and Design*. John wiley & sons, 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 *IEEE MICC*, Nov. 2017, pp. 259–264.

[7] Stratospheric Platforms Limited. (2021) Welcome to the future of global communicatoins - fast, reliable 5G technology. Stratospheric Platforms Limited. Accessed on 29 Jul. 2021. [Online]. Available: https://www.stratosphericplatforms.com/

[8] F. A. d'Oliveira, F. C. Melo, and T. C. Devezas, "High-altitude platforms—present situation and technology trends," *J. Aerosp. Technol. Manag.*, vol. 8, no. 3, pp. 249–262, 2016.

[9] D. Grace, T. Jiang, S. Allsopp, L. Reynaud, and M. Mohorcic, "Aerial platform study," *Deliverable D2.3 of ABSOLUTE WP2*, 2013.

[10] T. H. Chauhan, S. Agarwal, S. Purohit, and A. Kumar, "Wireless communications from high altitude platforms," *Int. J. Emerg. Technol. Adv. Eng.*, vol. 3, no. 4, pp. 220–223, 2013.

[11] Airbus Defence and Space. Zerphyr persistence and flexibility. Accessed on 02 Nov. 2018. [Online]. Available: https://lf5422.com/wp-content/uploads/2018/08/0296_18_2_zephyr_datasheet_e_horizontal_a4.pdf

[12] BAE Systems. Phasa-35. Accessed on 10 Oct. 2018. [Online]. Available: http://prismaticltd.co.uk/products/phasa-35/

[13] Y. Li, H. A. Chan, and J. I. Agbinya, "Networking in rural environments: Benefits, feasibilities, and requirements," *African J. Inf. Commun. Technol.*, vol. 2, no. 2, pp. 50–55, 2006.

[14] J. Gonzalo, D. Lopez, D. Dominguez, A. Garcia, and A. Escapa, "On the capabilities and limitations of high altitude pseudo-satellites," *Prog. Aerosp. Sci.*, vol. 98, pp. 37–56, 2018.

[15] A. Mohammed, A. Mehmood, F.-N. Pavlidou, and M. Mohorcic, "The role of high-altitude platforms (HAPs) in the global wireless connectivity," *Proc. IEEE*, vol. 99, no. 11, pp. 1939–1953, 2011.

[16] G. Romeo, G. Frulla, and E. Cestino, "Design of a high-altitude long-endurance solar-powered unmanned air vehicle for multi-payload and operations," *Int. J. Aerosp. Eng.*, vol. 221, no. 2, pp. 199–216, 2007.

[17] B. Kranjec, S. Sladic, W. Giernacki, and N. Bulic, "PV system design and flight efficiency considerations for fixed-wing radio-controlled aircraft—a case study," *Energies*, vol. 11, no. 10, p. 2648, 2018.

[18] A. H. Ali, "Investigation of indoor wireless-N radio frequency signal strength," in *IEEE ICIEA*, Sep. 2011, pp. 200–203.

[19] G. M. Djuknic, J. Freidenfelds, and Y. Okunev, "Establishing wireless communications services via high-altitude aeronautical platforms: a concept whose time has come?" *IEEE Commun. Mag.*, vol. 35, no. 9, pp. 128–135, Sep. 1997.

[20] D. Grace, T. Tozer, and N. Daly, "Communications from high altitude platforms a complementary or disruptive technology?" *IET*, pp. 8–8(1), Jan. 2000.

[21] G. K. Kurt et al., "A vision and framework for the high altitude platform station (HAPS) networks of the future," *IEEE Commun. Surv. Tut.*, pp. 1–1, 2021.

[22] 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, 2016.

[23] Sceye. Stratospheric platforms to improve life on our planet. Sceye. Accessed on 23 Jan. 2020. [Online]. Available: https://www.sceye.com/

[24] D. Grace and M. Mohorcic, *Broadband communications via high altitude platforms*. John Wiley & Sons, 2011.

[25] NASA. (2014) NASA armstrong fact sheet: Proteus high-altitude aircraft. Accessed on 10 Nov. 2018. [Online]. Available: https://www.nasa.gov/centers/armstrong/news/FactSheets/FS-069-DFRC.html

[26] HAPS Alliance, "Bridging the digital divide with aviation in the stratosphere," pp. 1–25, Dec. 2021. [Online]. Available: https://hapsalliance.org/publications/

[27] S. Adurogboye. (2016) Nigeria:- NCAA issues safety guidelines for drone operators. Accessed on 30 Jun. 2018. [Online]. Available: https://www.suasnews.com/2016/05/43512/

[28] Law Library Of Congress, U.S. Global Legal Research Directorate, "Regulation of drones: United Kingdom," Law Library Of Congress, U.S.. Global Legal Research Directorate, Tech. Rep. 2016-013384, Apr. 2016.

[29] D. Yuniarti, "Regulatory challenges of broadband communication services from high altitude platforms (HAPs)," in *IEEE ICOIACT*, 2018, pp. 919–922.

[30] H. Liu and F. Tronchetti, "Regulating near-space activities: Using the precedent of the exclusive economic zone as a model?" *Ocean Develop. Int. Law*, vol. 50, no. 2-3, pp. 91–116, 2019.

[31] ITU, "Provisional Final Acts," in *World Radiocommunication Conference 2019*. ITU Publications, 2019.

[32] Airbus Defence and Space. Zerphyr persistence and flexibility. Airbus Defence and Space. Accessed on 2 Nov. 2018.

[33] BAE Systems. PHASA-35. Accessed on 10 Oct. 2018. [Online]. Available: http://prismaticltd.co.uk/products/phasa-35/

[34] Boeing. High altitude long endurance (HALE): Long mission duration for the ultimate persistent surveillance. Boeing. Accessed on 14 Dec. 2018. [Online]. Available: https://www.boeing.com/defense/autonomoussystems/hale/index.page#/overview

[35] A. Boscaleri, F. Castagnoli, P. Rissone, and M. Corti, "Stratobus: a multiuser platform system for making access to LDB flight easier and cheaper," *European Space Agency,(Special Publication) ESA SP*, vol. 671, pp. 209–213, 2008.

[36] AeroVironment. (2018) Aerovironment announces joint venture and solar high-altitude long-endurance unmanned aircraft system development program. Accessed on 13 Jun. 2019. [Online]. Available: https://www.avinc.com/resources/press-releases/view/solar-high-altitude-long-endurance-uas

[37] HAPS Mobile. (2019, Apr.) SoftBank Corp. develops aircraft that delivers telecommunications connectivity from the stratosphere. Accessed on 10 Feb. 2019. [Online]. Available: https://www.hapsmobile.com/en/news/press/2019/20190425_01/

[38] Loon LLC. Connect people everywhere. Loon LLC. Accessed on 21 Mar. 2018. [Online]. Available: https://loon.co/

[39] L. Nagpal and K. Samdani, "Project Loon: Innovating the connectivity worldwide," in *IEEE RTEICT*, May 2017, pp. 1778–1784.

[40] Y. Albagory and A. E. Abbas, "Smart cell design for high altitude platforms communication," *AEU Int. J. Electron. Commun.*, vol. 67, no. 9, pp. 780–786, 2013.

[41] M. D. Zakaria, D. Grace, P. D. Mitchell, and T. M. Shami, "User-centric JT-CoMP for high altitude platforms," in *IEEE SoftCOM*, 2018, pp. 1–6.

[42] J. A. Wickboldt, W. P. De Jesus, P. H. Isolani, C. B. Both, J. Rochol, and L. Z. Granville, "Software-defined networking: management requirements and challenges," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 278–285, 2015.

[43] S. Abdelwahab, B. Hamdaoui, M. Guizani, and T. Znati, "Network function virtualization in 5G," *IEEE Commun. Mag.*, vol. 54, no. 4, pp. 84–91, 2016.

[44] J. Ordonez-Lucena, P. Ameigeiras, D. Lopez, J. J. Ramos-Munoz, J. Lorca, and J. Folgueira, "Network slicing for 5G with SDN/NFV: Concepts, architectures, and challenges," *IEEE Commun. Mag.*, vol. 55, no. 5, pp. 80–87, 2017.

[45] S. Alfattani, W. Jaafar, Y. Hmamouche, H. Yanikomeroglu, A. Yongaçoglu, N. D. Đào, and P. Zhu, "Aerial platforms with reconfigurable smart surfaces for 5G and beyond," *IEEE Commun. Mag.*, vol. 59, no. 1, pp. 96–102, 2021.

[46] O. Khalifa, A. Ismail, R. Islam, O. Elshaikh, and Z. Elabdin, "High altitude platform for wireless communications and other services," in *IEEE ICCEE*, 2006, pp. 432–438.

[47] J. Lun et al., "TV white space broadband for rural communities using solar powered high altitude platform and terrestrial infrastructures," *White paper*, 2017.

[48] 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, 2005.

[49] A. Malinowski and R. Zieliński, "High altitude platform — future of infrastructure," *Int. J. Electron. Telecommun.*, vol. 56, no. 2, pp. 191–196, 2010.

[50] D. Grace, J. Thornton, G. Chen, G. P. White, and T. C. Tozer, "Improving the system capacity of broadband services using multiple high-altitude platforms," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 700–709, 2005.

[51] F. Dong, M. Li, X. Gong, H. Li, and F. Gao, "Diversity performance analysis on multiple HAP networks," *Sensors*, vol. 15, no. 7, pp. 15 398–15 418, 2015.

[52] S. C. Arum, D. Grace, P. D. Mitchell, and M. D. Zakaria, "Beam-pointing algorithm for contiguous high altitude platform cell formation for extended coverage," in *IEEE VTC*, Sep. 2019, pp. 1–5.

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

[54] B. El-Jabu and R. Steele, "Cellular communications using aerial platforms," *IEEE Trans. Veh. Technol.*, vol. 50, no. 3, pp. 686–700, May 2001.

[55] M. I. Dessouky, H. A. Sharshar, and Y. A. Albagory, "Design of high altitude platforms cellular communications," *Prog. Electromagn. Res.*, vol. 67, pp. 251–261, 2007.

[56] M. I. Dessouky and H. A. Sharshar and Y. A. Albagory, "Geometrical analysis of high-altitude platform's cellular footprint," *Prog. Electromagn. Res.*, vol. 67, pp. 263–274, 2007.

[57] M. D. Zakaria, D. Grace, P. D. Mitchell, T. M. Shami, and N. Morozs, "Exploiting user-centric joint transmission – coordinated multipoint with a high altitude platform system architecture," *IEEE Access*, vol. 7, pp. 38 957–38 972, 2019.

[58] Y. Albagory, "An efficient 2D-DOA estimation technique for high-altitude platforms mobile communications," *Wireless Pers. Commun.*, vol. 88, no. 3, pp. 429–448, 2016.

[59] M. J. Colella, J. N. Martin, and F. Akyildiz, "The HALO network," *IEEE Commun. Mag.*, vol. 38, no. 6, pp. 142–148, 2000.

[60] E. Falletti, M. Mondin, F. Dovis, and D. Grace, "Integration of a HAP within a terrestrial UMTS network: interference analysis and cell dimensioning," *Wireless Pers. Commun.*, vol. 24, no. 2, pp. 291–325, 2003.

[61] Y. Albagory, "Flat-top ring-shaped cell design for high-altitude platform communications," *Int. J. Comput. Netw. Inf. Secur.*, vol. 5, no. 7, p. 51, 2013.

[62] J. Holis, D. Grace, and P. Pechac, "Effect of antenna power roll-off on the performance of 3G cellular systems from high altitude platforms," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 46, no. 3, pp. 1468–1477, Jul. 2010.

[63] R. S. ITU-R, "Methodology for determining the power level for high altitude platform stations ground terminals to facilitate sharing with space station receivers in the bands 47.2-47.5 GHz and 47.9-48.2 GHz," ITU-R, Tech. Rep., Oct. 2007. [Online]. Available: https://www.itu.int/rec/R-REC-SF.1843-0-200710-I

[64] A. Hourani, K. Sithamparanathan, and S. Lardner, "Optimal LAP altitude for maximum coverage," *IEEE Wireless Commun. Let.*, no. 99, pp. 1–4, 2014.

[65] N. Goddemeier, K. Daniel, and C. Wietfeld, "Coverage evaluation of wireless networks for unmanned aerial systems," in *IEEE Globecom*. IEEE, 2010, pp. 1760–1765.

[66] A.-H. Akram et al., "Coverage and rate analysis of aerial base stations," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 52, no. 6, pp. 3077–3081, 2016.

[67] P. K. Sharma and D. I. Kim, "Coverage probability of 3D UAV networks with RWP mobility-based altitude control," in *2018 IEEE International Conference on Communications Workshops (ICC Workshops)*. IEEE, 2018, pp. 1–6.

[68] X. Shen, Z. Wei, and Z. Feng, "A novel algorithm of UAV-mounted base station placement and frequency allocation," in *Springer 5GWN*. Springer, 2017, pp. 182–193.

[69] S. Karapantazis and F. Pavlidou, "Broadband communications via high-altitude platforms: a survey," *IEEE Commun. Surv. Tut.*, vol. 7, no. 1, pp. 2–31, 2005.

[70] Cambridge Consultants. (2021) The world's largest commercial airborne antenna. Cambridge Consultants. Accessed on 8 Jul. 2021. [Online]. Available: https://www.cambridgeconsultants.com/case-studies/worlds-largest-commercial-airborne-antenna

[71] J. Thornton, D. Grace, M. H. Capstick, and T. C. Tozer, "Optimizing an array of antennas for cellular coverage from a high altitude platform," *EEE Trans. Wireless Commun.*, vol. 2, no. 3, pp. 484–492, May 2003.

[72] M. H. Capstick and D. Grace, "High altitude platform mm-Wave aperture antenna steering solutions," *Wireless Pers. Commun.*, vol. 32, no. 3-4, pp. 215–236, 2005.

[73] Y. Yang, R. Zong, X. Gao, and J. Cao, "Channel modeling for high-altitude platform: A review," in *IEEE ISPACS*. IEEE, 2010, pp. 1–4.

[74] M. Li, M. Yang, G. Lv, and Q. Guo, "Three-state Semi-Markov channel model for HAP-high speed train communication link," in *IEEE CHINACOM*. IEEE, 2011, pp. 279–283.

[75] B. Rouzbehani and A. Dana, "Statistical modelling of small-scale fading effects for high-altitude platform propagation channels," *J. Chinese Inst. Eng.*, vol. 37, no. 4, pp. 540–546, 2014.

[76] M. R. K. Aziz et al., "Channel estimation for LTE downlink in high altitude platforms (HAPs) systems," in *IEEE ICoICT*. IEEE, 2013, pp. 182–186.

[77] X. Liu, X. Zeng, C. Liu, W. Liu, and Y. Zhang, "The channel estimation and modeling in high altitude platform station wireless communication dynamic network," *Discrete Dyn. Nature Soc.*, vol. 2017, 2017.

[78] J. Holis and P. Pechac, "Elevation dependent shadowing model for mobile communications via high altitude platforms in built-up areas," *IEEE Trans. Ant. Propag.*, vol. 56, no. 4, pp. 1078–1084, 2008.

[79] W. Khawaja, I. Guvenc, D. W. Matolak, U.-C. Fiebig, and N. Schneckenberger, "A survey of air-to-ground propagation channel modeling for unmanned aerial vehicles," *IEEE Commun. Surv. Tut.*, 2019.

[80] X. Lin, S. Rommer, S. Euler, E. A. Yavuz, and R. S. Karlsson, "5G from space: An overview of 3GPP non-terrestrial networks," 2021.

[81] 3GPP, "Study on new radio (NR) to support non-terrestrial networks," Tech. Rep. 38.811, Oct. 2020. [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/38_series/38.811/

[82] T. Bogale and L. Le, "Massive MIMO and mmwave for 5G wireless hetnet: Potential benefits and challenges," *IEEE Veh. Technol. Mag.*, vol. 11, no. 1, pp. 64–75, 2016.

[83] E. I. Elsaidy, M. I. Dessouky, S. Khamis, and Y. Albagory, "Concentric circular antenna array synthesis using comprehensive learning particle swarm optimizer," *Prog. Electromag. Res. Let.*, vol. 29, pp. 1–13, 2012.

[84] V. Rabinovich and N. Alexandrov, "Typical array geometries and basic beam steering methods," in *Antenna Arrays and Automotive Applications*. Springer, 2013, pp. 23–54.

[85] M. Nofal, S. Aljahdali, and Y. Albagory, "Tapered beamforming for concentric ring arrays," *AEU Int. J. Electron. Commun.*, vol. 67, no. 1, pp. 58–63, 2013.

[86] Mathworks, "Hann (Hanning) window - MATLAB hann," accessed on 05 Dec. 2021. [Online]. Available: https://www.mathworks.com/help/signal/ref/hann.html

[87] S. H. Alsamhi and N. S. Rajput, "Efficient cooperative HAPS-terrestrial WiMAX system," in *WASET ICRAEECE*, 2014.

[88] S. H. Alsamhi and N. S. Rajput, "Neural network in a joint HAPS and terrestrial fixed broadband system," *Int. J. Technol. Explor. Learn.*, vol. 3, pp. 344–348, 2014.

[89] M. Mokayef, T. A. Rahman, R. Ngah, and M. Y. Ahmed, "Spectrum sharing model for coexistence between high altitude platform system and fixed services at 5.8 GHz," *Int. J. Multimedia Ubiquitous Eng.*, vol. 8, no. 5, pp. 265–275, 2013.

[90] B. Jabbari, "Fixed and dynamic channel assignment," *The Communications Handbook, 2nd edition, CRC Press LLC*, 2002.

[91] S. A. E. Ahmed and A. B. Nabi, "Dynamic versus static channel allocation scheme," *Int. J. Sci. Res.*, vol. 4, no. 6, pp. 2319–7064, 2013.

[92] M. P. Mishra and P. C. Saxena, "Survey of channel allocation algorithms research for cellular systems," *Int. J. Netw. Commun.*, vol. 2, no. 5, pp. 75–104, 2012.

[93] J. Del Ser, M. Matinmikko, S. Gil-López, and M. Mustonen, "Centralized and distributed spectrum channel assignment in cognitive wireless networks: a harmony search approach," *Appl. Soft Comput.*, vol. 12, no. 2, pp. 921–930, 2012.

[94] R. Mochaourab, B. Holfeld, and T. Wirth, "Distributed channel assignment in cognitive radio networks: Stable matching and walrasian equilibrium," *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3924–3936, 2015.

[95] Y. Li, H. C. B. Chan, P. Lam, and P. H. J. Chong, "Channel allocation method for multi-radio wireless mesh networks based on a genetic algorithm," *Lecture Notes Eng. Comput. Sci.*, 2016.

[96] A. Jella and S. L. Sabat, "Dynamic channel access of secondary users in a heterogeneous network using game theory," in *IEEE COMSNETS*. IEEE, 2018, pp. 425–428.

[97] F. D. Calabrese, L. Wang, E. Ghadimi, G. Peters, L. Hanzo, and P. Soldati, "Learning radio resource management in RANs: Framework, opportunities, and challenges," *IEEE Commun. Mag.*, vol. 56, no. 9, pp. 138–145, 2018.

[98] T. O. Olwal, K. Djouani, and A. M. Kurien, "A survey of resource management toward 5G radio access networks," *IEEE Commun. Surv. Tut.*, vol. 18, no. 3, pp. 1656–1686, 2016.

[99] S. Sesia, I. Toufik, and M. Baker, *LTE-the UMTS long term evolution: from theory to practice*. John Wiley & Sons, 2011.

[100] N. Morozs, "Accelerating reinforcement learning for dynamic spectrum access in cognitive wireless networks," Ph.D. dissertation, University of York, Sep. 2015. [Online]. Available: https://etheses.whiterose.ac.uk/11523/

[101] W. Mei and R. Zhang, "UAV-sensing-assisted cellular interference coordination: A cognitive radio approach," *IEEE Wireless Commun. Let.*, vol. 9, no. 6, pp. 799–803, 2020.

[102] N. Morozs, T. Clarke, and D. Grace, "Heuristically accelerated reinforcement learning for dynamic secondary spectrum sharing," *IEEE Access*, vol. 3, pp. 2771–2783, 2015.

[103] A. Galindo-Serrano, B. Sayrac, S. B. Jemaa, J. Riihijärvi, and P. Mähönen, "Harvesting MDT data: Radio environment maps for coverage analysis in cellular networks," in *IEEE CROWNCom*. IEEE, 2013, pp. 37–42.

[104] H. B. Yilmaz, T. Tugcu, F. Alagöz, and S. Bayhan, "Radio environment map as enabler for practical cognitive radio networks," *IEEE Commun. Mag.*, vol. 51, no. 12, pp. 162–169, 2013.

[105] V.-P. Chowdappa, C. Botella, J. J. Samper-Zapater, and R. J. Martinez, "Distributed radio map reconstruction for 5G automotive," *IEEE Trans. Intell. Transp. Syst.*, vol. 10, no. 2, pp. 36–49, 2018.

[106] P. He, N. Cheng, and S. Ni, "Improved LMS predictive link triggering for handover in HAPS communication system," in *IEEE WCSP*. IEEE, 2016, pp. 1–5.

[107] X. Wang, L. Li, and W. Zhou, "The effect of HAPS unstable movement on handover performance," in *IEEE WOCC*. IEEE, 2019, pp. 1–5.

[108] H. Kaushal and G. Kaddoum, "Optical communication in space: challenges and mitigation techniques," *IEEE Commun. Surv. Tut.*, vol. 19, no. 1, pp. 57–96, 2016.

[109] H. W. Philips, "Some design considerations for solar-powered aircraft," Jun 1980.

[110] J. P. Nelson, "High-altitude considerations for electrical power systems and components," *IEEE Trans. Ind. Appl.*, vol. IA-20, no. 2, pp. 407–412, 1984.

[111] G. S. Aglietti, S. Redi, A. R. Tatnall, and T. Markvart, "High altitude electrical power generation," *WSEAS Trans. Environ. Develop.*, vol. 4, no. 12, pp. 1067–1077, 2008.

[112] M. Brizon, "Solar energy generation model for high altitude long endurance platforms," 2015. [Online]. Available: https://pdfs.semanticscholar.org/f837/02e7ab679b192c9a237dc0c150b94bdea317.pdf

[113] R. Islam and K. Saraswat, "Limitation of optical enhancement in ultra-thin solar cells imposed by contact selectivity," *Sci. Rep.*, vol. 8, no. 8863, 2018.

[114] D. Yuniarti, "Regulatory challenges of broadband communication services from high altitude platforms (HAPs)," in *IEEE ICOIACT*, 2018, pp. 919–922.

[115] T. M. Lim, A. M. Cramer, J. E. Lumpp, and S. A. Rawashdeh, "A modular electrical power system architecture for small spacecraft," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 4, pp. 1832–1849, 2018.

[116] I. Dincer, "Hydrogen and fuel cell technologies for sustainable future," *Jordan J. Mech. Ind. Eng.*, vol. 2, no. 1, pp. 1–14, 2008.

[117] V. M. Sanchez, R. Barbosa, J. C. Cruz, and F. C. J. Hernandez, "Optimal sizing of a photovoltaic-hydrogen power system for HALE aircraft by means of particle swarm optimization," *Math. Problems Eng.*, vol. 2015, no. 183701, pp. 1–8, 2015.

[118] V. S. Dwivedi, P. Kumar, A. K. Ghosh, and G. M. Kamath, "Selection of size of battery for solar powered aircraft," *IFAC-PapersOnLine*, vol. 51, no. 29, pp. 424–430, 2018.

[119] B. Lee, S. Kwon, P. Park, and K. Kim, "Active power management system for an unmanned aerial vehicle powered by solar cells, a fuel cell, and batteries," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 4, pp. 3167–3177, 2014.

[120] X.-Z. Gao, Z.-X. Hou, Z. Guo, J.-X. Liu, and X.-Q. Chen, "Energy management strategy for solar-powered high-altitude long-endurance aircraft," *Energy Convers. Manage.*, vol. 70, pp. 20–30, 2013.

[121] E. Cestino, "Design of solar high altitude long endurance aircraft for multi payload & operations," *Aerosp. Sci. Technol.*, vol. 10, no. 6, pp. 541–550, 2006.

[122] J. R. Boucher, "Sunrise, the world's first solar-powered airplane," *J. Aircr.*, vol. 22, no. 10, pp. 840–846, 1985.

[123] A. Jenkins, "The sun's position in the sky," *Eur. J. of Phys.*, vol. 34, no. 3, pp. 633–652, 2013.

[124] L. Elterman, "Aerosol measurements in the troposphere and stratosphere," *Applied Optics*, vol. 5, no. 11, pp. 1769–1776, 1966.

[125] C. Honsberg and S. Bowden. Elevation angle. Accessed on 02 Feb. 2019. [Online]. Available: https://www.pveducation.org/pvcdrom/properties-of-sunlight/elevation-angle

[126] ITACA. Solar energy reaching the earth's surface. Accessed on 02 Feb. 2019. [Online]. Available: https://www.itacanet.org/the-sun-as-a-source-of-energy/part-2-solar-energy-reaching-the-earths-surface/

[127] X.-Z. Gao, Z.-X. Hou, Z. Guo, J.-X. Liu, and X.-Q. Chen, "Energy management strategy for solar-powered high-altitude long-endurance aircraft," *Energy convers. manage.*, vol. 70, pp. 20–30, 2013.

[128] T. Benson, "Lift formula," May 2021, accessed on 05 Dec. 2021. [Online]. Available: https://www.grc.nasa.gov/www/k-12/WindTunnel/Activities/lift_formula.html

[129] MIT, "Flight thrust, power, and energy relations," Feb. 2009, Lab 1 Lecture Notes, Accessed on 05 Dec. 2021. [Online]. Available: https://web.mit.edu/16.unified/www/SPRING/systems/Lab_Notes/airpower.pdf

[130] R. l. Liu, Z. j. Zhang, Y. f. Jiao, C. h. Yang, and W. j Zhang, "Study on flight performance of propeller-driven UAV," *J. Aersp. Eng.*, vol. 2019, no. 6282451, pp. 1–11, 2019.

[131] M. B. Guillem, P. Nurgeldy, and V. Ârpâd, "Performance analysis of hybrid electric and distributed propulsion system applied on a light aircraft," *Elsevier Energy*, vol. 214, no. 118823, pp. 1–15, 2021.

[132] Light Aircraft Association, "Electric propulsion systems," Tech. Rep. TL 3.28, May 2020. [Online]. Available: http://www.lightaircraftassociation.co.uk/engineering/technical_leaflets.html

[133] K. Katzis and D. Grace, "Inter-high altitude platform handoff for communications systems with directional antennas," *URSI Radio Sci. Bull. Special Iss. HAPs*, 2010.

[134] F. D. Rogers. (2018) Turning performance - sustained level turns. Accessed on 12 Nov. 2018. [Online]. Available: http://www.nar-associates.com/technical-flying/turning/turnprpa_wide_screen.pdf

[135] G. Auer et al., "How much energy is needed to run a wireless network?" *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 40–49, 2011.

[136] C. Desset et al., "Flexible power modeling of LTE base stations," in *IEEE WCNC*, 2012, pp. 2858–2862.

[137] A. Mathur. (2018, Sep.) 5G densification and network power efficiency. Nokia Networks. [Online]. Available: https://futurenetworks.ieee.org/images/files/pdf/SantaClaraTutorial2018/5-5G_EET_5G_Densification_Power_Efficiency_Apurv_Mathur_Nokia-FINAL-MATHUR-NOKIA-PDF-TO-SHARE-9-28-18.pdf

[138] E. Bertran and A. Sànchez-Cerdà, "On the tradeoff between electrical power consumption and flight performance in fixed-wing UAV autopilots," *IEEE Trans. Veh. Technol.*, vol. 65, no. 11, pp. 8832–8840, 2016.

[139] MicroLink Devices Inc. Microlink devices achieves certified 37.75% solar cell power conversion efficiency. Accessed on 10 Oct. 2018. [Online]. Available: http://prismaticltd.co.uk/products/phasa-35/

[140] Engineering ToolBox. (2003) U.S. standard atmosphere. Engineering ToolBox. Accessed on 2 Aug. 2021. [Online]. Available: https://www.engineeringtoolbox.com/standard-atmosphere-d_604.html

[141] B. Mattos, N. Sêcco, and E. F. Salles, "Optimal design of a high-altitude solar- powered unmanned airplane," *J. Aerosp. Technol. Manag.*, vol. 5, pp. 349–361, 2013.

[142] A. Alsahlani, L. J. Johnston, and P. A. Atcliffe, "Design of a high altitude long endurance flying-wing solar-powered unmanned air vehicle," *Prog. Flight Phys.*, vol. 9, pp. 3–24, 2017.

[143] A. Mendez, T. J. Leo, and M. A. Herreros, "Current state of technology of fuel cell power systems for autonomous underwater vehicles," *Energies*, vol. 7, no. 7, pp. 4676–4693, 2014.

[144] J. Wishart, "Fuel cells vs batteries in the automotive sector," 2014.

[145] Z. F. Pan, L. An, and C. Y. Wen, "Recent advances in fuel cells based propulsion systems for unmanned aerial vehicles," *Appl. Energy*, vol. 240, pp. 473–485, 2019.

[146] S. J. Eickhoff and J. M. Klein, "High power fuel cell system," *US Patent US20180277860A1*, 07 2019.

[147] K. Flittie and B. Curtin, "Pathfinder solar-powered aircraft flight performance," in *AIAA AFMC*, 1998, pp. 618–632.

[148] J. H. DelFrate, "Helios prototype vehicle mishap: Technical findings, recommendations, and lessons learned," Jan. 2007. [Online]. Available: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20070022260.pdf

[149] A. Checko, H. L. Christiansen, Y. Yan, L. Scolari, G. Kardaras, M. S. Berger, and L. Dittmann, "Cloud RAN for mobile networks—a technology overview," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 405–426, 2015.

[150] L. Perlman and M. Wechsler. (2019, Apr.) Mobile coverage and its impact on digital financial services. Accessed on 01 Dec. 2020. [Online]. Available: https://ssrn.com/abstract=3370669

[151] A. A. Oloyede, A. Shamsudeen, N. Faruk, L. A. Olawoyin, S. I. Popoola, and A. Abdulkarim, "Cost effective tri-band mobile phone jammer for hospitals applications," in *IEEE IREHI*, 2018.

[152] Z. Yang and A. Mohammed, "Deployment and capacity of mobile WiMAX from high altitude platform," in *IEEE VTC*, Sep. 2011, pp. 1–5.

[153] A. Burr, A. Papadogiannis, and T. Jiang, "MIMO truncated shannon bound for system level capacity evaluation of wireless networks," in *IEEE WCNCW*, Apr. 2012, pp. 268–272.

[154] Teltonika, "Mobile signal strength recommendations," accessed on 31 Dec. 2021. [Online]. Available: https://wiki.teltonika-networks.com/view/Mobile_Signal_Strength_Recommendations

[155] T. Tiong, "Adaptive transceivers for mobile communications," in *IEEE ICAIET*, 2014, pp. 275–279.

[156] M. Alouini and A. J. Goldsmith, "Area spectral efficiency of cellular mobile radio systems," *IEEE Trans. Veh. Technol.*, vol. 48, no. 4, pp. 1047–1066, Jul. 1999.

[157] S. Chatzinotas, M. A. Imran, and C. Tzaras, "On the capacity of variable density cellular systems under multicell decoding," *IEEE Wireless Commun. Lett.*, vol. 12, no. 7, pp. 496–498, Jul. 2008.

[158] T. Barrera, A. Hast, and E. Bengtsson, "Incremental spherical linear interpolation," in *SIGRAD Conf. Special Theme - Environ. Vis.*, vol. 013, no. 004, Nov. 2004, pp. 7–10.

[159] W. Xu and Y. K. Deng, "Investigation on electronic azimuth beam steering in the spaceborne SAR imaging modes," *J. Electromag. Waves Appl.*, vol. 25, no. 14-15, pp. 2076–2088, 2011.

[160] H.-C. Zeng, J. Chen, W. Yang, and H.-J. Zhang, "Impacts of azimuth antenna steering angle quantization on TOPS and sliding spotlight SAR image," in *IEEE IGARSS*, 2018, pp. 7813–7816.

[161] M. Taranetz, J. Colom Ikuno, and M. Rupp, "Capacity density optimization by fractional frequency partitioning," in *IEEE ASILOMAR*, Nov 2011, pp. 1398–1402.

[162] J. D. Lawrence, *A catalog of special plane curves*. Courier Corporation, 2013.

[163] H. Yang et al., "A 1600-element dual-frequency electronically reconfigurable reflectarray at X/Ku-band," *IEEE Trans. Antennas Propag.*, vol. 65, no. 6, pp. 3024–3032, 2017.

[164] 3GPP, "Evolved universal terrestrial radio access (E-UTRA); radio frequency (RF) system scenarios," Tech. Rep. 36.942, Jul. 2020. [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/36_series/36.942/

[165] Y. Xu, Y. Hu, Q. Chen, T. Song, and R. Lai, "Robust resource allocation for multi-tier cognitive heterogeneous networks," in *IEEE ICC*, 2017, pp. 1–6.

[166] S. Sekander, H. Tabassum, and E. Hossain, "Multi-tier drone architecture for 5G/B5G cellular networks: Challenges, trends, and prospects," *IEEE Commun. Mag.*, vol. 56, no. 3, pp. 96–103, 2018.

[167] P. Balling, "Multibeam antennas," *Encyclopedia of RF and Microw. Eng.*, 2005.

[168] Tran Cao Quyen and Trinh Anh Vu, "The application of a multibeam antenna for 4G," in *IEEE ATC*, 2012, pp. 227–230.

[169] J. B. L. Rao, R. Mital, D. P. Patel, M. G. Parent, and G. Tavik, "Low-cost multibeam phased array antenna for communications with GEO satellites," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 28, no. 6, pp. 32–37, 2013.

[170] W. Hong et al., "Multibeam antenna technologies for 5G wireless communications," *IEEE Trans. Antennas Propag.*, vol. 65, no. 12, pp. 6231–6249, 2017.

[171] R. K. Jain, D.-M. W. Chiu, and W. R. Hawe, "A quantitative measure of fairness and discrimination," Eastern Research Laboratory, Digital Equipment Corporation, Hudson, MA, Tech. Rep. TR-301, Sep. 1984.

[172] M. Yassin et al., "Survey of ICIC techniques in LTE networks under various mobile environment parameters," *Wireless Netw.*, vol. 23, no. 2, pp. 403–418, 2017.

[173] M. C. Batistatos, G. V. Tsoulos, D. A. Zarbouti, G. E. Athanasiadou, and S. K. Goudos, "LTE measurements for flying relays," in *IEEE MOCAST*, 2018, pp. 1–4.

[174] M. Ndong and T. Fujii, "Cross-tier interference management with a distributed antenna system for multi-tier cellular networks," *EURASIP J. Wireless Commun. Netw.*, vol. 2014, no. 1, p. 73, 2014.

[175] N. Morozs, T. Clarke, and D. Grace, "Distributed heuristically accelerated Q-learning for robust cognitive spectrum management in LTE cellular systems," *IEEE Trans. Mobile Comput.*, vol. 15, no. 4, pp. 817–825, 2016.

[176] K. I. Pedersen, Y. Wang, B. Soret, and F. Frederiksen, "eICIC functionality and performance for LTE HetNet co-channel deployments," in *EEE VTC*, Sep. 2012, pp. 1–5.

[177] W. Li, Y. Zhang, L. Huang, J. Cosmas, and Q. Ni, "A cell specific reference signal interference cancellation scheme for LTE cellular access systems," in *IEEE BMSB*, Jun. 2015, pp. 1–6.

[178] H. Tataria, M. Shafi, A. Molisch, M. Dohler, H. Sjöland, and F. Tufvesson, "6G wireless systems: Vision, requirements, challenges, insights, and opportunities," *IEEE Proc.*, vol. 109, no. 7, pp. 1166–1199, Mar. 2021.

[179] G. Wikström et al., "White paper: Ever-present intelligent communication – a research outlook towards 6G," Ericsson, Tech. Rep. GFTL-20:001402, Nov. 2020. [Online]. Available: https://www.ericsson.com/en/reports-and-papers/white-papers/a-research-outlook-towards-6g

[180] X. You et al., "Towards 6G wireless communication networks: vision, enabling technologies, and new paradigm shifts," *Sci. China Inf. Sci.*, vol. 64, no. 1, pp. 1869–1919, 2020.

[181] T. Nakamura, "5G evolution and 6G," in *IEEE Symp. VLSI Technol.*, Dec. 2020, pp. 1–5.

[182] K. B. Letaief, W. Chen, Y. Shi, J. Zhang, and Y.-J. A. Zhang, "The roadmap to 6G: AI empowered wireless networks," *IEEE Commun. Mag.*, vol. 57, no. 8, pp. 84–90, 2019.

[183] A. Pärssinen et al., "White paper: RF enabling 6G – opportunities and challenges from technology to spectrum," University of Oulu, Tech. Rep. 13, Apr. 2021.

[184] H. Farhady, H. Lee, and A. Nakao, "Software-defined networking: A survey," *Computer Netw.*, vol. 81, pp. 79–95, 2015.

[185] M. Karakus and A. Durresi, "Economic viability of software defined networking (SDN)," *Computer Netw.*, vol. 135, pp. 81–95, 2018.

[186] A. Filali, A. Abouaomar, S. Cherkaoui, A. Kobbane, and M. Guizani, "Multi-access edge computing: A survey," *IEEE Access*, vol. 8, pp. 197 017–197 046, 2020.

[187] Z. Jia, M. Sheng, J. Li, D. Zhou, and Z. Han, "Joint HAP access and LEO satellite backhaul in 6G: Matching game-based approaches," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 4, pp. 1147–1159, 2021.

[188] K. G. Karabulut et al., "A vision and framework for the high altitude platform station (HAPS) networks of the future," *IEEE Commun. Surv. Tut.*, vol. 23, no. 2, pp. 729–779, 2021.

[189] E. C. Strinati et al., "6G in the sky: On-demand intelligence at the edge of 3D networks," *ETRI J.*, vol. 42, no. 5, pp. 643–657, 2020.

[190] K. Mershad, H. Dahrouj, H. Sarieddeen, B. Shihada, T. Al-Naffouri, and M.-S. Alouini, "Cloud-enabled high-altitude platform systems: Challenges and opportunities," 2021.

[191] J. Qiu, D. Grace, G. Ding, M. D. Zakaria, and Q. Wu, "Air-ground heterogeneous networks for 5G and beyond via integrating high and low altitude platforms," *IEEE Wireless Commun.*, vol. 26, no. 6, pp. 140–148, December 2019.

[192] P. S. Yedavalli, T. Riihonen, X. Wang, and J. M. Rabaey, "Far-field RF wireless power transfer with blind adaptive beamforming for internet of things devices," *IEEE Access*, vol. 5, pp. 1743–1752, 2017.

[193] F. Rinaldi et al., "Non-terrestrial networks in 5G beyond: A survey," *IEEE Access*, vol. 8, pp. 165 178–165 200, 2020.

[194] G. K. Kurt and H. Yanikomeroglu, "Communication, computing, caching, and sensing for next-generation aerial delivery networks: Using a high-altitude platform station as an enabling technology," *IEEE Veh. Technol. Mag.*, pp. 2–11, 2021.

[195] T. C. Quyen and T. A. Vu, "The application of a multibeam antenna for 4G," in *IEEE ATC*, 2012, pp. 227–230.

[196] M. Giordani and M. Zorzi, "Non-terrestrial networks in the 6G era: Challenges and opportunities," *IEEE Netw.*, vol. 35, no. 2, pp. 244–251, 2021.

[197] C. de Lima et al., "White paper: 6G white paper on localization and sensing - 6G research visions," University of Oulu, Tech. Rep. 12, Jun. 2020. [Online]. Available: http://urn.fi/urn:isbn:9789526226743

[198] E. Yaacoub and Z. Dawy, "LTE radio network planning with hetnets: BS placement optimization using simulated annealing," in *IEEE MELECON*, Apr. 2014, pp. 327–333.

[199] E. Yaacoub and Z. Dawy, "LTE BS placement optimization using simulated annealing in the presence of femtocells," in *IEEE EWC*, May 2014, pp. 1–5.

[200] S. Chen, Y.-C. Liang, S. Sun, S. Kang, W. Cheng, and M. Peng, "Vision, requirements, and technology trend of 6G: How to tackle the challenges of system coverage, capacity, user data-rate and movement speed," *IEEE Wireless Commun.*, vol. 27, no. 2, pp. 218–228, 2020.

[201] F. Dong, Y. He, X. Gong, J. Wang, and H. Li, "A constellation design methodology based on QoS and user demand in high-altitude platform broadband networks," *IEEE Trans. Multimedia*, vol. 18, no. 12, pp. 2384–2397, 2016.

[202] F. Dong, Y. He, H. Nan, Z. Zhang, and J. Wang, "System capacity analysis on constellation of interconnected HAP networks," in *IEEE BDCC*, 2015, pp. 154–159.