New approaches for resource allocation in future communication networks using NOMA and UAVs

Marie-Joséphа Youssef

DOCTORAT BRETAGNE LOIRE / MATHSTIC

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

THESE DE DOCTORAT DE

L'ÉCOLE NATIONALE SUPERIEURE MINES-TELECOM ATLANTIQUE
BRETAGNE PAYS DE LA LOIRE - IMT ATLANTIQUE

ECOLE DOCTORALE N° 601
Mathématiques et Sciences et Technologies de l'Information et de la Communication
Spécialité : *Télécommunications*

Par

Marie-Joséphа YOUSSEF

Nouvelles approches pour l'allocation des ressources dans les réseaux de communication futurs utilisant NOMA et aidés par les drones

New approaches for resource allocation in future communication networks using NOMA and UAVs (in English)

Thèse présentée et soutenue à Brest, le 19 novembre 2020
Unité de recherche : Lab-STICC
Thèse N° : 2020IMTA0208

Rapporteurs avant soutenance :

Mylène PISCHELLA	Maître de conférences (HDR), Conservatoire National des Arts et Métiers (CNAM)
Jean-Marie GORCE	Professeur des universités INSA de Lyon

Composition du Jury :

Président :	Frédéric GUILLOUD	Professeur, IMT Atlantique
Examinateur :	Raphaël VISOZ	Chercheur, Orange Labs
Rapporteurs :	Mylène PISCHELLA	Maître de conférences (HDR), Conservatoire National des Arts et Métiers (CNAM)
	Jean-Marie GORCE	Professeur des universités INSA de Lyon
Directrices de thèse :	Catherine DOUILLARD	Professeur, IMT Atlantique
	Joumana FARAH	Professeur, Université Libanaise
Encadrant de thèse :	Charbel ABDEL NOUR	Maître de conférences, IMT Atlantique
Invité :	Venugopal VEERAVALLI	Professeur, Université d'Illinois à Urbana-Champaign

Sous le sceau de l'Université Bretagne Loire

IMT Atlantique

École Doctorale MATHSTIC

New Approaches for Resource Allocation in Future Communication Networks using NOMA and UAVs

---

Thèse de Doctorat

Spécialité : Télécommunications

Présentée par **Marie-Joséph a Youssef**

Département : Électronique

Laboratoire : Lab-STICC

Directeurs de thèse : **Catherine Douillard, Joumana Farah**

Encadrant : **Charbel Abdel Nour**

Abstract

New Approaches for Resource Allocation in Future Communication Networks using NOMA and UAVs

Marie-Josépha Youssef

Electronics Department, IMT Atlantique

The Internet of Things (IoT) is driving the evolution of future wireless communication networks. Indeed, IoT devices are expected to form the major portion of 5G communication networks and beyond with a foretasted number of billions of connected machine-type devices (MTDs). This growth in the number of connected MTDs will not only lead to the exponential increase in the required wireless capacity, but will also result in the emergence of new wireless use cases that greatly differ from conventional human users-oriented services. To ensure a seamless integration of MTDs into future communication networks, while continuously providing support for human users, new techniques must be leveraged to meet the quality of service (QoS) requirements of all users. To this effect, some of the key elements 5G communication networks and beyond are expected to rely on include novel spectrum access techniques, self-organized networks (SON) and unmanned aerial vehicles (UAV)-aided wireless communication networks.

In this regard, the main objective of this thesis is to provide novel resource allocation and network design solutions to optimize the use of available radio resources in next generation wireless communication networks. In particular, we focus on the use of non-orthogonal multiple access (NOMA) for radio access. By exploiting the power domain, NOMA allows multiple users to share the same orthogonal resource block, thus improving system performance in terms of spectral efficiency, achieved throughput and fairness. First, a distributed antenna system (DAS) that consists of users having heterogeneous mobile traffic requirements is studied. In particular, two user categories are considered: the first consisting of best-effort (BE) users seeking to maximize their achieved data rates, while preserving a high level of fairness between them; and the second consisting of users running real-time (RT) applications, hence aiming to receive a quantity of data bits before the expiration of a predefined latency limit. In this context, two novel solutions, leveraging the use of NOMA scheduling, are proposed. The first solution consists of a low-complexity greedy algorithm, while the second relies on matching theory to find the best antenna assignment and the best allocation of frequency resources to users. The aim of both proposed methods is to maximize the number of satisfied RT users, while optimizing, to the extent of the possible, the performance of BE users. The convergence and stability of the matching theory-based method are proved, and its computational complexity is found. Simulation results show that the proposed solutions greatly outperform conventional resource allocation algorithms. In fact, the low-complexity greedy algorithm outperforms the proportional fairness scheduler by up to 90% in terms of RT users satisfaction. Simulations show also that the matching theory-based method can yield up to 25% gain in the number of satisfied RT users when compared to a previously proposed method in the literature. Moreover, the matching theory-based method is able to achieve more than 90% of the performance of the optimal method based on exhaustive search, with a much lower computational complexity. Second, we shift our focus to uncoordinated

spectrum access, as grant-free communications and SONs are expected to make use of uncoordinated spectrum access to organize their transmissions. A multi-player multi-armed bandits (MP-MAB) framework is introduced to solve the uncoordinated spectrum access problem. For grant-free communications, varying channel rewards across users are considered. While previous studies on MP-MAB and uncoordinated spectrum access limited each user to choose one channel in each timeslot, in this thesis, each user is allowed to access multiple channels in a timeslot. This results in an MP-MAB problem with varying rewards across users and multiple plays. For SONs, both the uncoordinated spectrum access and the distributed power control problems are studied. NOMA is exploited to allow multiple access points (AP) to simultaneously access the same channel. Unlike previous work on the combination of NOMA and MP-MAB, non-zero rewards when multiple APs choose the same channel are assumed in this thesis. Varying channel rewards between APs are also considered, and each AP is allowed to access multiple channels. Hence, the uncoordinated spectrum access in the SON is modeled using the MP-MAB framework with varying reward across APs, non-zero rewards on collision and multiple plays. In its turn, the distributed power control is modeled using the MP-MAB framework with varying rewards across APs. A game-theoretic solution for these problems is proposed. Using theoretical derivations validated by numerical simulations, the regret of the proposed method is proved to be sub-linear. Third, we study the integration of UAVs as aerial BSs in wireless communication networks. In this context, we give particular attention to the wireless backhaul link between the UAV-BS and an MBS serving as gateway to the core network. In fact, this wireless backhaul link is necessary for the orderly functioning of UAV-aided wireless communication networks. To increase system spectral efficiency, an in-band wireless backhaul is considered, leading the access and the backhaul links to share the same frequency band. We also assume that the UAV is equipped with full-duplex (FD) capabilities, allowing it to receive and transmit data at the same time. To enhance system performance, NOMA is considered for spectrum access. The aim of this work is to optimize the deployment position of the UAV, in addition to finding the optimal subband assignment and power allocation values in the backhaul and the access links. In this regard, an optimization problem that minimizes the UAV-BS transmit power is solved. Using simulation results, we prove the superior performance of the proposed methods, when compared to previously proposed algorithms. In fact, the proposed method yields up to 85% performance gain in terms of achieved system throughput, when compared with a method that relies on out-band backhauling.

Keywords: Non-orthogonal multiple access, mixed traffic, uncoordinated spectrum access, self-organized networks, unmanned aerial vehicles, optimization, matching theory, multi-armed bandits.

Contents

List of Figures

List of Tables

List of Abbreviations

3G	Third Generation
4G	Fourth Generation
5G	Fifth Generation
A2G	Air-To-Ground
AP	Access Point
BE	Best-Effort
BI	Backhaul Interference
BS	Base Station
C-RAN	Cloud Radio Access Networks
CAS	Centralized Antenna System
CDF	Cumulative Distribution Function
CSI	Channel State Information
D2D	Device-To-Device
DA	Deferred Acceptance
DAS	Distributed Antenna System
EE	Energy Efficiency
eMBB	Enhanced Mobile Broadband
FD	Full-Duplex
FTPA	Fractional Transmit Power Allocation
HAP	High Altitude Platform
HD	Half-Duplex
i.i.d.	Independent And Identically Distributed
IBFD	In-Band Full-Duplex
ICI	Inter-Cell Interference
IDMA	Interleave Division Multiple Access
IoT	Internet Of Things
ITU-R	International Telecommunication Union
KKT	Karush-Kuhn-Tucker
LAP	Low Altitude Platform
LOS	Line-Of-Sight
MAB	Multi-Armed Bandits
MBS	Macro Base Station
mMTC	Massive Machine-Type Communications
mmWave	Millimeter Wave
MP-MAB	Multi-Player Multi-Armed Bandits
MTD	Machine-Type Device
NLOS	Non-Line-Of-Sight
NOMA	Non-Orthogonal Multiple Access
OFDM	Orthogonal Frequency Division Multiplexing

OMA	Orthogonal Multiple Access
PD-NOMA	Power-Domain NOMA
PDMA	Pattern Division Multiple Access
PER	Packet Error Rate
PF	Proportional Fairness
PSO	Particle Swarm Optimization
QoS	Quality Of Service
RACH	Random Access Channels
RL	Reinforcement Learning
RMS	Root-Mean-Square
RRH	Remote Radio Head
RT	Real-Time
SBS	Small Base Station
SC	Superposition Coding
SCMA	Sparse Code Multiple Access
SI	Self-Interference
SIC	Successive Interference Cancellation
SINR	Signal-To-Interference-Plus-Noise-Ratio
SON	Self-Organizing Network
TTI	Transmission Time Interval
UAV	Unmanned Aerial Vehicle
UCB	Upper-Confidence-Bound
URLLC	Ultra-Reliable Low-Latency Communications

Résumé étendu de la thèse en français

Chapitre 1: Introduction Générale

La demande d'accès sans fil omniprésent a fortement augmenté ces dernières années, en raison de la croissance exponentielle du nombre de dispositifs connectés intelligents tels que les téléphones mobiles, les voitures autonomes et les capteurs. Les nouveaux scénarios d'utilisation conçus pour une grande partie de ces nouveaux appareils diffèrent considérablement des services vocaux ou multimédias traditionnels. Les applications de réalité virtuelle et augmentée, les véhicules autonomes connectés, les robots télécommandés, les capteurs de surveillance et de contrôle et les drones sont quelques exemples de ces nouveaux usages. Comme certains de ces usages nécessitent un débit de données important, leur introduction dans les réseaux de communication sans fil mettra à rude épreuve la capacité des systèmes cellulaires sans fil existants. En outre, les futurs réseaux de communication doivent être capables de fournir des communications ultra fiables à faible latence (URLLC) aux services ayant de telles exigences [1, 2]. Les véhicules autonomes connectés, la chirurgie à distance et l'automatisation des usines sont des exemples de services nécessitant des URLLCs. Hormis l'URLLC, il est nécessaire de prendre en charge un grand nombre de dispositifs et de capteurs connectés, constituant l'environnement de l'Internet des objets (IoT) [3, 4]. En résumé, les futurs réseaux de communication doivent être correctement modélisés, conçus et optimisés pour pouvoir répondre aux exigences de qualité de service (QoS) des services mobiles à large bande (eMBB), de l'URLLC et des communications massives de type machine (mMTC) [5, 6]. En étendant les capacités des réseaux cellulaires et en satisfaisant les exigences de QoS requises, il est possible de libérer le potentiel de nouveaux cas d'usage capables d'améliorer considérablement notre qualité de vie.

Pour pouvoir répondre aux exigences des applications émergentes, les réseaux sans fil actuels doivent évoluer et adopter de nouvelles technologies. En particulier, les futurs réseaux de communication feront appel à de nouvelles techniques d'accès au spectre basées sur l'accès multiple, au déploiement de cellules denses, à des capacités d'auto-organisation des cellules déployées, aux communications par ondes millimétriques (mmWave), et à des communications assistées par des véhicules aériens sans pilote (UAV). Comment concevoir un système sans fil harmonieux en tirant parti de ces nouvelles techniques pour optimiser de manière adaptative l'utilisation des ressources sans fil est une question importante qui a fait l'objet d'une grande attention dans la littérature récente.

Le but principal de cette thèse est d'exploiter des techniques d'optimisation et d'apprentissage par renforcement pour optimiser l'allocation des ressources dans les

réseaux cellulaires sans fil de prochaine génération. Dans un premier temps, nous nous concentrons sur la gestion du spectre pour les futurs réseaux de communication, où des utilisateurs ayant des exigences de QoS différentes doivent être servis. Ensuite, nous abordons certains des défis qui se posent dans le contexte de l'accès multiple libre ou non coordonné et des réseaux auto-organisés (SON). Enfin, nous étudions l'utilisation de drones comme stations de base aériennes dans les systèmes de communication sans fil. Plus précisément, nous abordons le problème du positionnement du drone, de l'allocation des bandes fréquencielles et de puissance, en tenant compte de la liaison backhaul entre le drone et le réseau central.

Chapitre 2: Contexte et Outils

Ce chapitre couvre différents sujets en lien avec le travail présenté dans cette thèse. Un contexte général sur les réseaux cellulaires de prochaine génération est d'abord présenté. En particulier, un aperçu du principe de l'accès multiple non orthogonal (NOMA), des applications et des exigences de l'IoT et des réseaux de communication assistés par des UAVs est donné. La deuxième partie de ce chapitre apporte ensuite des informations générales sur les techniques d'analyse auxquelles les solutions proposées font appel. Plus précisément, des éléments sur les bases de la théorie des jeux, des techniques de bandits à plusieurs bras et de la théorie de l'appariement sont fournis.

Les Futurs Réseaux de Communication

Le moteur principal de l'évolution des systèmes sans fil dans le passé était le besoin de débits de données plus élevés. Cependant, avec la prolifération de nouvelles applications IoT, les exigences clés que doivent garantir les réseaux cellulaires de la cinquième génération (5G) et au-delà sont de plus en plus diverses. En particulier, les principaux éléments clés de la 5G, illustrés dans la Fig. 1, comprennent : des débits de données plus élevés, une meilleure couverture, une grande fiabilité, une faible latence, une faible consommation d'énergie et la prise en charge des objets connectés.

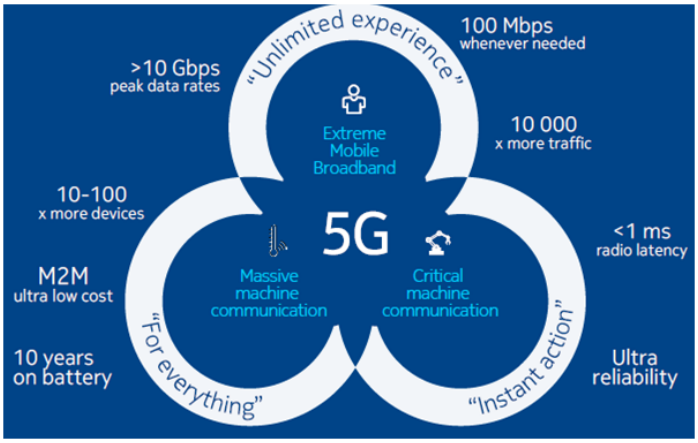


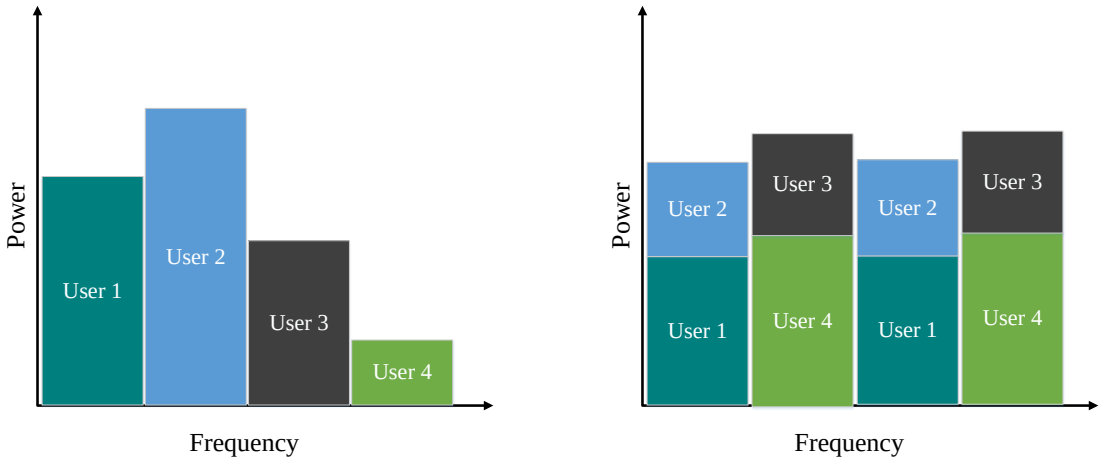
Figure 1 – Les cas d'usage et exigences de la 5G

Pour répondre à ces exigences, d'importants efforts de recherche ont été consacrés à

l'étude de nouvelles techniques et architectures de réseaux cellulaires, capables de faire face à la rapide évolution des systèmes sans fil. Ces techniques comprennent le déploiement dense de systèmes d'antennes distribuées (DAS) (par opposition au traditionnel système centralisé (CAS)), et de petites stations de base (BS), l'utilisation de drones comme BS aériennes et une utilisation plus efficace du spectre existant.

Accès Multiple Non Orthogonal (NOMA)

Le concept de base du NOMA repose sur l'exploitation du domaine de la puissance pour servir plusieurs utilisateurs de façon non orthogonale sur le même bloc de ressources OFDM. Le NOMA repose sur le codage par superposition (SC) [7] du côté émetteur et sur l'annulation successive des interférences (SIC) [8] du côté récepteur. Une illustration comparant les techniques NOMA et celles de l'accès multiple orthogonal (OMA) est présentée dans la Fig. 2.



(a) La technique OMA (b) La technique NOMA

Figure 2 – Comparaison entre les techniques OMA et NOMA

NOMA en Voie Descendante

En voie descendante, la BS utilise le SC pour transmettre une combinaison des messages superposés des utilisateurs. Du côté du récepteur, les utilisateurs exécutent le SIC, selon l'ordre croissant de leurs gains de canal, pour récupérer leur message. Dans cette thèse, nous limitons à deux le nombre d'utilisateurs multiplexés de façon non orthogonale sur chaque sous-bande. Soit $\mathcal{K}_s = \{k_s(1), k_s(2)\}$ l'ensemble des utilisateurs programmés sur la sous-bande s triés par ordre décroissant de gain de canal. Les débits atteints par $k_s(1)$ et $k_s(2)$ sur la sous-bande s sont respectivement donnés par :

$$R_{k_s(1),s} = B_c \log_2 \left(1 + \frac{P_{k_s(1),s} h^2_{k_s(1),s}}{N_0 B_c}\right), \tag{1}$$

$$R_{k_s(2),s} = B_c \log_2 \left(1 + \frac{P_{k_s(2),s} h^2_{k_s(2),s}}{P_{k_s(1),s} h^2_{k_s(2),s} + N_0 B_c}\right). \tag{2}$$

Comme on peut le voir dans les Eq. (1) et (2), seul l'utilisateur ayant le gain de canal le plus faible souffre d'interférences.

NOMA en Voie Montante

En liaison montante, la BS envoie des signaux de contrôle aux utilisateurs émetteurs pour déterminer la répartition de la puissance. Une fois la puissance d'émission déterminée, chaque utilisateur envoie son message. La BS reçoit dans ce cas un signal constitué par la superposition des signaux des utilisateurs émetteurs. En utilisant le SIC, la BS décode les messages des utilisateurs selon l'ordre décroissant des gains de canal. Dans le cas de deux utilisateurs programmés sur la sous-bande s, $k_s(1)$ et $k_s(2)$, $k_s(1)$ étant l'utilisateur ayant le gain de canal le plus élevé, les débits obtenus en liaison montante NOMA sont donnés par :

$$R_{k_s(1),s} = B_c \log_2 \left(1 + \frac{P_{k_s(1),s} h^2_{k_s(1),s}}{P_{k_s(2),s} h^2_{k_s(2),s} + N_0 B_c}\right), \tag{3}$$

$$R_{k_s(2),s} = B_c \log_2 \left(1 + \frac{P_{k_s(2),s} h^2_{k_s(2),s}}{N_0 B_c}\right), \tag{4}$$

où $P_{k_s(.),s}$ est la puissance de transmission de l'utilisateur $k_s(.)$ sur la sous-bande s.

L'internet des Objets et les Nouvelles Caractéristiques du Trafic Mobile

Les futurs réseaux de communication sans fil devront prendre en charge des milliards de dispositifs de type machine connectés (MTDs) [9], donnant naissance à l'IoT. Pour un déploiement efficace des services IoT, il faut résoudre de nombreux problèmes tels que l'analyse des données, les capacités de transmission, l'allocation des ressources, la sécurité et la confidentialité. En particulier, les exigences des dispositifs sont très différentes les unes des autres, et encore plus des exigences des utilisateurs humains. À cet effet, il est de la plus haute importance de trouver des techniques d'allocation des ressources pour les dispositifs IoT qui puissent répondre à leurs divers besoins de trafic en termes de débit, de latence et de fiabilité.

Réseaux de Communication Assistés par des Drones

Dans les systèmes de communication sans fil, des drones ou UAVs correctement déployés et exploités peuvent apporter des solutions à de nombreux problèmes. En effet, les drones peuvent être déployés comme des BS aériennes pour fournir des communications à la demande dans les zones qui en ont besoin. Grâce à leur altitude de vol ajustable, la probabilité d'établir des liaisons en visibilité directe ou "line-of-sight (LOS)" entre l'UAV et les utilisateurs qui lui sont assignés est accrue, ce qui se traduit par des communications plus efficaces par rapport aux BS terrestres. En outre, grâce à la souplesse de déploiement des UAVs, la capacité et la couverture des réseaux de communication équipés de drones peuvent être accrues. En raison de ces nombreux avantages, l'utilisation des UAVs en tant que BSs aériennes dans les réseaux de communication sans fil peut satisfaire de nombreuses applications. Toutefois, pour tirer parti des avantages des UAVs dans les systèmes de communication, plusieurs défis doivent être relevés. Parmi ces défis, nous pouvons citer : la modélisation du canal UAV vers les utilisateurs terrestres, le positionnement de l'UAV, l'optimisation de sa trajectoire et sa connectivité vers le réseau central.

Bases de la Théorie des Jeux

La théorie des jeux est un cadre mathématique qui étudie les interactions stratégiques entre des joueurs indépendants, égoïstes et rationnels. La théorie des jeux est utilisée dans différentes disciplines telles que l'économie, la politique, la biologie et la tarification des produits. Récemment, la théorie des jeux a reçu une attention significative dans la communauté des communications sans fil, en raison de l'émergence de réseaux sans fil distribués à grande échelle nécessitant des capacités d'auto-organisation [10]. Pour résoudre les problèmes d'accès au spectre et de contrôle de puissance non coordonnés de manière distribuée dans le chapitre 4 de cette thèse, nous proposons une solution basée sur les chaînes de Markov et la théorie des jeux, puisque cette dernière peut correctement modéliser les interactions entre différents agents indépendants.

Bandits à Plusieurs Bras

Le cadre "multi-armed bandits (MAB)" [11] est un cas particulier du cadre de l'apprentissage par renforcement (RL) . Dans le cadre MAB, l'agent vise à trouver l'action la plus rémunératrice dans un environnement à état unique. Le problème MAB a été largement étudié dans la littérature scientifique. Récemment, le cadre "multi-player multi-armed bandits (MP-MAB)", qui implique de nombreux agents essayant de trouver les actions les plus gratifiantes, a suscité un intérêt considérable dans la communauté de recherche sur les communications sans fil. Il a été appliqué pour résoudre de nombreux problèmes, parmi lesquels l'accès opportuniste au spectre dans les réseaux de radio cognitive [12, 13, 14], l'accès non coordonné au spectre [15], et l'optimisation des trajectoires dans les réseaux cellulaires assistés par des UAVs [16].

La Théorie de l'Appariement

La théorie de l'appariement ou "matching theory" [17] est un outil mathématique puissant qui s'est avéré utile pour résoudre le problème d'allocation de ressources décrit dans le chapitre 3 de ce manuscrit. Il s'agit d'un cadre mathématique qui décrit en économie la formation de relations mutuellement bénéfiques. En particulier, la théorie de l'appariement est utilisée pour résoudre les problèmes d'affectation. Dans la littérature sur les communications sans fil, cet outil a récemment attiré l'attention en raison des diverses propriétés utiles qu'il présente. Dans ces réseaux, les problèmes d'affectation des ressources consistent à attribuer les ressources du réseau (par exemple, les intervalles de temps, les canaux de fréquence, la puissance, les antennes de service) à des entités exigeantes (par exemple, des appareils ou des utilisateurs). L'objectif du problème d'allocation des ressources est d'allouer de manière optimale les ressources aux utilisateurs, compte tenu d'un ensemble de contraintes du réseau. Pour résoudre ce problème, plusieurs travaux récents ont proposé des algorithmes basés sur le cadre de la théorie de l'appariement [18, 19, 20, 21, 22].

Chapitre 3: Allocation de Ressources dans les Systèmes à Trafic Mobile Mixte

Dans ce chapitre, nous étudions un système de trafic mixte composé d'utilisateurs "real-time (RT)" et "best-effort (BE)". Alors que les utilisateurs RT se caractérisent par des

applications à latence limitée, les utilisateurs BE visent à maximiser les débits obtenus tout en optimisant l'équité du système. Nous proposons l'utilisation de la technique NOMA dans un système d'antennes distribuées pour servir les utilisateurs. En particulier, nous formulons un problème d'allocation de sous-bandes et d'antennes pour la coexistence de plusieurs types de trafic. Le problème d'affectation des ressources proposé vise à maximiser à la fois le niveau de satisfaction des utilisateurs de type RT et la performance des utilisateurs de type BE.

Algorithme Glouton de Faible Complexité

Comme le système est composé d'utilisateurs au trafic hétérogène, la technique d'allocation des ressources doit tenir compte de la différence de priorité entre les utilisateurs. C'est pourquoi nous avons conçu un algorithme glouton de faible complexité qui planifie d'abord les utilisateurs RT sur les sous-bandes et les antennes de manière OMA. Une fois la phase OMA terminée, nous passons à la phase de couplage par NOMA des utilisateurs. Lorsque plus aucun des utilisateurs RT n'a besoin d'être programmé, les utilisateurs BE sont programmés selon le principe d'équité proportionnelle ou "proportional fairness (PF)". Les étapes de l'algorithme peuvent être résumées comme suit :

1. Les utilisateurs et les sous-bandes sont affectés aux antennes distribuées. Avant tout, cette distribution est effectuée de manière à garantir les besoins des utilisateurs RT. Elle doit également maximiser l'utilité des utilisateurs BE, dans la mesure du possible.

2. Une phase de couplage NOMA suit, au cours de laquelle nous assignons les seconds utilisateurs aux sous-bandes attribuées aux utilisateurs RT. Durant ce couplage, l'assignation sous-bande-antenne est maintenue, et les débits requis par les utilisateurs RT déjà programmés sont garantis. L'objectif de cette étape est de satisfaire en priorité les utilisateurs RT non satisfaits, puis d'améliorer la performance des utilisateurs BE lorsque cela est possible.

Algorithme Basé sur la Théorie de l'Appariement

Pour améliorer la performance du système, nous faisons appel au cadre de la théorie de l'appariement bidimensionnel afin d'obtenir une solution efficace pour le problème considéré. La solution proposée fait appel à l'algorithme d'acceptation différée ou "deferred acceptance (DA)" comme suit :

1. Programmation des utilisateurs OMA RT

    - Les utilisateurs RT qui n'ont pas reçu les débits de données requis postulent pour leurs sous-bandes préférées.
    - Les sous-bandes prennent leurs décisions concernant l'affectation des utilisateurs RT en utilisant OMA et retiennent les utilisateurs qui maximisent leur utilité.
    - Les utilisateurs RT suppriment de leurs listes de préférences les sous-bandes auxquelles ils ont postulé.

2. Programmation des utilisateurs BE en utilisant un système NOMA hybride sur les sous-bandes restantes
    - Les utilisateurs BE et les paires des utilisateurs BE postulent pour leurs sous-bandes préférées.
    - Les sous-bandes prennent leurs décisions concernant l'affectation des utilisateurs BE et retiennent les utilisateurs qui maximisent leur utilité.
    - Les utilisateurs BE suppriment de leurs listes de préférences les sous-bandes auxquelles ils ont postulé.
3. Couplage NOMA sur les sous-bandes assignées aux utilisateurs RT
    - Les utilisateurs RT non satisfaits et les utilisateurs BE postulent à leurs sous-bandes préférées.
    - Les sous-bandes prennent leurs décisions concernant l'affectation des utilisateurs RT et BE et retiennent les utilisateurs qui maximisent leur utilité.
    - Les utilisateurs suppriment de leurs listes de préférences les sous-bandes auxquelles ils ont postulé.

Exemples de Résultats

La performance des techniques proposées est évaluée par le biais de simulations numériques. La technique basée sur la théorie de l'appariement, désignée par "MM", est testée dans les configurations OMA-CAS, OMA-DAS, NOMA-CAS et NOMA-DAS. Une variante de la méthode MM, désignée par "MM-FA", dans les configurations DAS est également testée. MM-FA adopte l'approche de [23] et n'optimise pas le nombre de sous-bandes par antenne. La performance obtenue par la méthode gloutonne de faible complexité est désignée par "GM".

Dans la Fig. 3, la performance des techniques proposées en termes de satisfaction des utilisateurs RT est évaluée. On peut noter que, jusqu'à ce que $K_{RT} = 15$, MM et GM ont des performances similaires quel que soit le scénario envisagé. Cependant, à mesure que la cellule devient plus congestionnée avec un plus grand nombre d'utilisateurs RT, MM surpasse GM dans toutes ses variations. Plus concrètement, lorsque $K_{RT} = 30$, GM n'obtient pratiquement aucune satisfaction pour les utilisateurs RT. Cependant, MM-OMA-CAS (resp. MM-OMA-DAS) surpasse son équivalent GM de près de 28% (resp. 62 %). En outre, dans le cas du NOMA, MM-NOMA-CAS (resp. MM-NOMA-DAS) surpasse son équivalent GM de près de 30 % (resp. 63 %). La Fig. 3 montre également les gains obtenus en optimisant le nombre de sous-bandes par antenne. Par exemple, lorsque $K_{RT} = 30$ utilisateurs, MM-OMA-DAS (resp. MM-NOMA-DAS) surpasse MM-FA-OMA-DAS (resp. MM-FA-NOMA-DAS) de près de 30 % (resp. 26 %). D'autre part, les résultats montrent le gain obtenu en utilisant un système DAS, par rapport à un système CAS, puisqu'il peut augmenter la performance de plus de 30%.

La Fig. 4a montre le débit total atteint par les utilisateurs BE au fur et à mesure de l'augmentation du nombre d'utilisateurs RT. Comme prévu, la somme des débits de toutes les méthodes diminue à mesure que K_{RT} augmente, car les ressources disponibles pour les utilisateurs BE sont moins nombreuses. Les deux méthodes (GM et MM) ont des performances similaires en ce qui concerne le débit total atteint par les utilisateurs

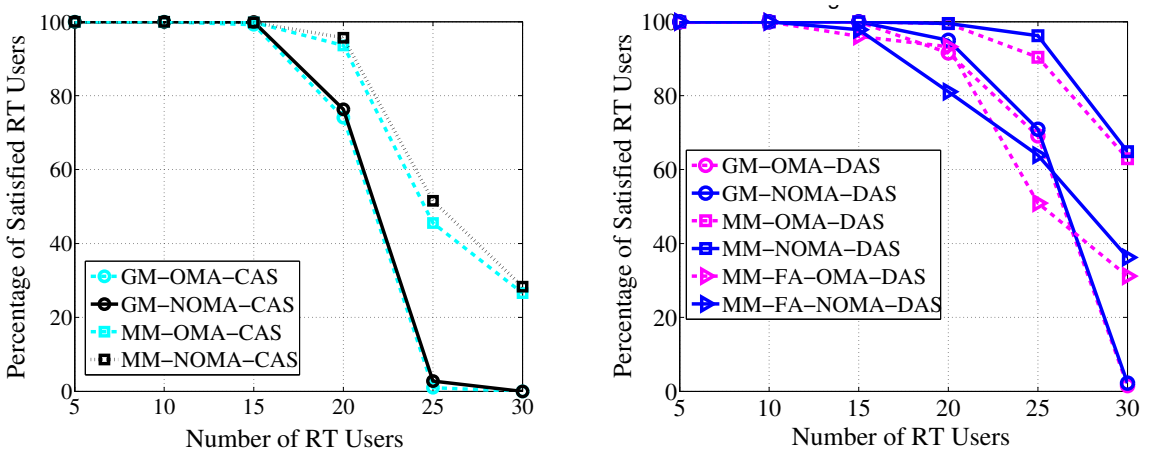
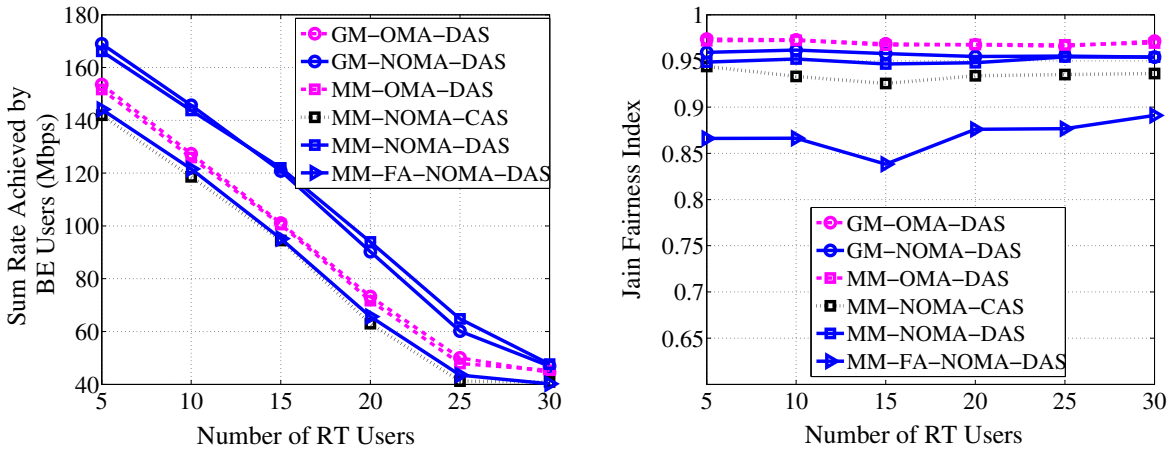


Figure 3 – Le pourcentage de satisfaction des utilisateurs RT pour les systèmes: (a) centralisés ou CAS , (b) et distribués ou DAS

BE. Par exemple, pour le cas NOMA-DAS, GM-NOMA-DAS réalise un gain de près de 1 Mbps par rapport à MM-NOMA-DAS lorsque $K_{RT} = 5$ ou 10 utilisateurs. Cependant, pour $K_{RT} = 20$ ou 25 utilisateurs, MM-NOMA-DAS réalise un gain de près de 3 Mbps par rapport à GM-NOMA-DAS. En outre, le MM-NOMA-DAS surpasse largement le MM-FA-NOMA-DAS. Dans la Fig. 4b, nous montrons l'équité obtenue par les différentes méthodes en fonction de K_{RT}. L'équité du système est évaluée par l'indice d'équité de Jain [24] qui varie entre 0 et 1 en prenant la valeur 1 dans le cas d'une équité optimale. On peut voir que MM-NOMA-DAS surpasse son homologue FA. En mettant MM-FA à part, la Fig. 4b montre que toutes les méthodes considérées ont de bonnes performances en termes d'équité avec un indice Jain supérieur à 0,9, avec un avantage pour les configurations DAS.

Chapitre 4: Accès Non Coordonné au Spectre Utilisant l'Outil Bandit à Plusieurs Bras

Dans ce chapitre, nous considérons la partie "fronthaul" d'un réseau sans fil auto-organisé où plusieurs points d'accès (AP) visent à organiser leurs transmissions en liaison ascendante avec une unité centrale de manière distribuée. Nous étudions à la fois les problèmes d'accès non coordonné au spectre et de contrôle distribué de la puissance. Une solution basée sur le cadre des bandits à plusieurs bras MAB, qui ne nécessite aucune coordination ou communication entre les APs, est proposée. Pour la première phase, soit celle de l'accès non coordonné au spectre, en plus de considérer les récompenses variables des canaux entre les APs, chaque AP est autorisé à accéder simultanément à plusieurs canaux. En outre, chaque canal peut servir plusieurs APs en même temps en utilisant NOMA, ce qui entraîne un problème de type MP-MAB avec des récompenses variables selon les joueurs, des jeux multiples et une récompense non nulle en cas de collision. Pour la phase de contrôle de la puissance, des récompenses variables selon les utilisateurs pour les différents niveaux de puissance sont envisagées et un algorithme pour résoudre le problème



(a) (b)

Figure 4 – (a) Débit total atteint par les utilisateurs BE en fonction de K_{RT}, (b) equité du système en fonction de K_{RT}

de contrôle de la puissance sur chaque canal est proposé.

Algorithme Proposé

Dans ce travail, nous partons du principe que l'horizon temporel de l'allocation n'est pas nécessairement connu à l'avance. Par conséquent, la solution proposée procède par époques, chaque époque comprenant trois phases, à savoir, *exploration, appariement et exploitation* :

1. **Phase d'exploration:** Pour l'accès non coordonné au spectre, cette phase vise à estimer les récompenses moyennes précédemment inconnues de chaque canal, ainsi que le nombre total d'APs en concurrence pour les ressources du système. Pour le contrôle de la puissance distribuée, elle vise à estimer les récompenses moyennes précédemment inconnues de chaque niveau de puissance. Pendant cette phase, chaque AP accède uniformément à un bras à la fois pour estimer sa récompense moyenne. L'AP k accédant au canal m reçoit en retour la récompense obtenue sur m ainsi que le nombre total des APs accédant simultanément au canal m. Cette phase se déroule sur un nombre constant de créneaux temporels. À la fin de la phase, pour l'accès non coordonné au spectre, tous les APs ont une estimation $\hat{\boldsymbol{\mu}}_M$ des récompenses moyennes des canaux et du gain de canal expérimenté. Chaque AP calcule également une estimation du nombre total d'APs, $\hat{K}$. À la fin de la partie d'exploration de la phase de contrôle distribué de la puissance, chaque AP a une estimation des valeurs moyennes de récompense des niveaux de puissance disponibles.

2. **Phase d'appariement:** Dans cette phase, une solution basée sur la théorie des jeux est proposée. Chaque AP k est associé à un état $[\bar{\boldsymbol{a}}_k, \bar{\boldsymbol{u}}_k, S]$, où $\bar{\boldsymbol{a}}_k$ et $\bar{\boldsymbol{u}}_k$ sont respectivement l'action de base et l'utilité de base de l'AP k. La variable $S \in \{C, D\}$ est l'humeur de l'AP k et reflète si k est satisfait ou non de l'action et de l'utilité actuelles. Un AP satisfait joue son action de base avec une grande

probabilité tandis qu'un AP non satisfait choisit son action uniformément au hasard. Ensuite, en fonction de l'action choisie et de l'utilité obtenue, un AP devient satisfait selon une probabilité qui augmente lorsque l'utilité obtenue s'approche de la valeur maximale réalisable.

3. **Phase d'exploitation:** Au cours de cette phase, les APs adoptent l'allocation ayant donné la meilleure performance lors de la phase d'appariement précédente.

Analyse du Regret

Dans le cadre du MP-MAB, le regret est défini comme la différence entre l'utilité obtenue en jouant l'allocation optimale pour toutes les plages temporelles et l'utilité réelle obtenue. Le regret encouru dans les trois phases de l'algorithme est alors évalué :

1. **Phase d'exploration:** Lors de l'accès non coordonné au spectre, le regret encouru est plafonné par :
$$R_C^1 \leq KNT_C^0 \log\left(T_C/c_2 + 2\right), \tag{5}$$
où K est le nombre d'AP, N est le nombre de canaux choisis pour chaque plage temporelle, T_C^0 est la longueur de la phase d'exploration, T_C est la longueur de la partie d'accès non-coordonné au spectre et c_2 est une constante.

   De même, le regret prévu pour tous les APs dans la phase d'exploration de l'allocation de puissance, R_P^1, est limité par la borne supérieure :
$$R_P^1 \leq KT_P^0 \log(T_P/c_2 + 2), \tag{6}$$
où T_P est la longueur de la partie de l'allocation distribuée de puissance.

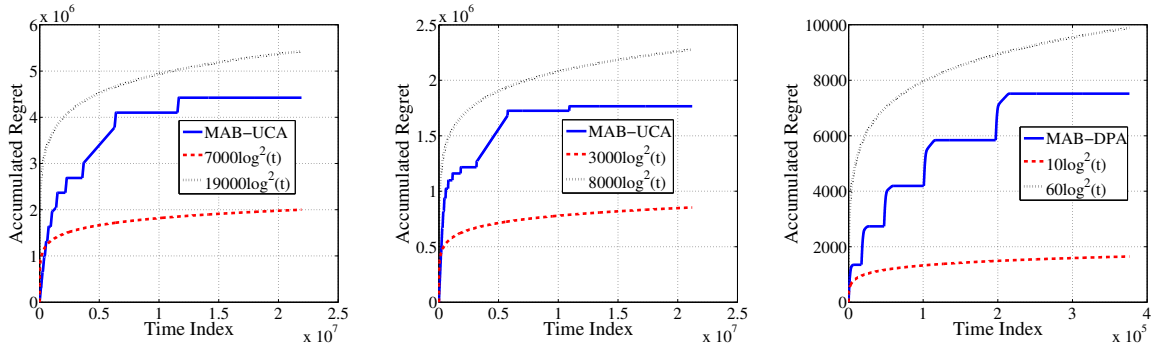
2. **Phase d'appariement:** Lors de l'accès au canal non coordonné, le regret encouru est limité par la borne supérieure :
$$R_C^2 \leq KNc_1 \log^{2+\delta}\left(T_C/c_2 + 2\right), \tag{7}$$
où c_1 et δ sont des constantes.

   De même, le regret prévu pour tous les APs dans la phase d'appariement de l'allocation de puissance, R_P^2, est limité par la borne supérieure :
$$R_P^2 \leq Kc_1 \log^{2+\delta}\left(T_P/c_2 + 2\right). \tag{8}$$

3. **Phase d'exploitation:** Le regret qu'éprouvent les APs dans la phase d'exploitation de l'accès non coordonné au spectre est $R_C^3 \leq A_3$. De même, le regret encouru par les APs dans la phase d'exploitation de l'allocation d'énergie est $R_P^3 \leq A_3$.

Regret de la Technique Proposée

Theorem 1. *Le regret prévu de la solution d'allocation proposée est sous-linéaire. Il peut être limité à la borne supérieure comme :*
$$R \leq R_C^1 + R_C^2 + R_C^3 + R_P^1 + R_P^2 + R_P^3 = \mathcal{O}\left(\log^{2+\delta}(T)\right). \tag{9}$$



(a) (b) (c)

Figure 5 – Regret accumulé en fonction du temps (a) pour la phase d'allocation des canaux avec une longueur de phase d'exploration constante, (b) pour la phase d'allocation des canaux avec une longueur de phase d'exploration dégressive, (c) pour la phase d'allocation de la puissance.

Exemples de Résultats

La Fig. 5 montre le regret moyen accumulé en fonction du temps dans la phase d'allocation des canaux pour les versions de la phase d'exploration à longueur constante et à longueur dégressive. Les résultats montrent que la moyenne des regrets accumulés pour les deux versions augmente avec le temps comme $\mathcal{O}(\log(t)^2)$. Plus précisément, le regret encouru dans le cas de la phase d'exploration à longueur constante est limité entre 7000 $\log(t)^2$ et 19000 $\log(t)^2$, comme le montre la Fig. 5a. Le regret encouru dans le cas de la phase d'exploration de longueur dégressive est limité entre 3000 $\log(t)^2$ et 8000 $\log(t)^2$. En fait, la plupart du regret est accumulé pendant la phase d'exploration où les APs choisissent un canal uniformément au hasard. Par conséquent, la réduction de la durée de la phase d'exploration diminue la valeur du regret accumulé, comme le montre la Fig. 5b, sans compromettre la précision de l'estimation.

Le regret encouru lors de la phase d'allocation de la puissance est limité entre 10 $\log(t)^2$ et 60 $\log(t)^2$, comme le montre la Fig. 5c. Le regret moins important observé lors de la phase de répartition de la puissance, par rapport à la phase de répartition des canaux, s'explique par le nombre moins élevé d'APs en concurrence pour un nombre de bras plus faible.

Chapitre 5: Systèmes de Communication Equipés de Drones

Dans ce chapitre, nous envisageons un scénario dans lequel l'infrastructure sans fil traditionnelle est absente, par exemple dans des régions éloignées ou à la suite d'une catastrophe ou d'une défaillance de la station de base. Un drone doté de capacités "full-duplex (FD)" est envoyé pour servir les utilisateurs ayant des besoins QoS dans cette zone, et un lien de liaison sans fil de type "in-band" est établi entre l'UAV et une macro station de base (MBS) pour fournir la capacité de liaison "backhaul" nécessaire. Afin de minimiser la puissance d'émission de l'UAV, un problème d'optimisation est résolu pour trouver l'affectation des sous-bandes dans les liaisons d'accès et de retour, la position

3D du UAV ainsi que les niveaux de puissance dans les liaisons d'accès et de retour. En outre, lorsque le budget de puissance de l'UAV n'est pas suffisant pour garantir les besoins de l'utilisateur en termes de débit, une étape d'appariement NOMA est effectuée pour maximiser les débits réalisables.

Algorithme Proposé

Le problème d'optimisation considéré se compose de trois sous-problèmes : le problème de placement de l'UAV, ainsi que les problèmes d'allocation des sous-bandes et de la puissance dans les liaisons d'accès et de backhaul. Toutefois, ce problème d'optimisation est mixte et multivarié. Pour le résoudre, nous proposons un algorithme multi-étapes qui cible les trois sous-problèmes et qui procède comme suit :

1. Effectuer l'assignation des sous-bandes dans le lien d'accès en utilisant l'algorithme hongrois. Cette assignation vise à minimiser l'interférence moyenne de la liaison de retour subie par les utilisateurs.

2. Résoudre un problème d'optimisation convexe pour vérifier si le budget de puissance disponible au niveau de l'UAV peut répondre aux besoins des utilisateurs sans tenir compte de l'interférence du backhaul.

3. Décider de l'attribution des sous-bandes de la liaison de retour. Pour ce faire :

   (a) En utilisant la méthode de recherche par bissection, déterminer le nombre minimum de sous-bandes requis pour la liaison de retour. Cette valeur initiale ne tient pas compte de l'impact de l'interférence du backhaul expérimenté du côté de l'utilisateur.

   (b) Pour toutes les affectations potentielles de sous-bandes dans la liaison de retour, trouver la valeur de puissance nécessaire pour l'UAV afin de satisfaire les exigences de débits des utilisateurs. Ensuite, conserver l'assignation des sous-bandes de la liaison de retour qui minimise la puissance nécessaire de l'UAV en tenant compte de l'interférence de la liaison de retour expérimentée du côté utilisateur.

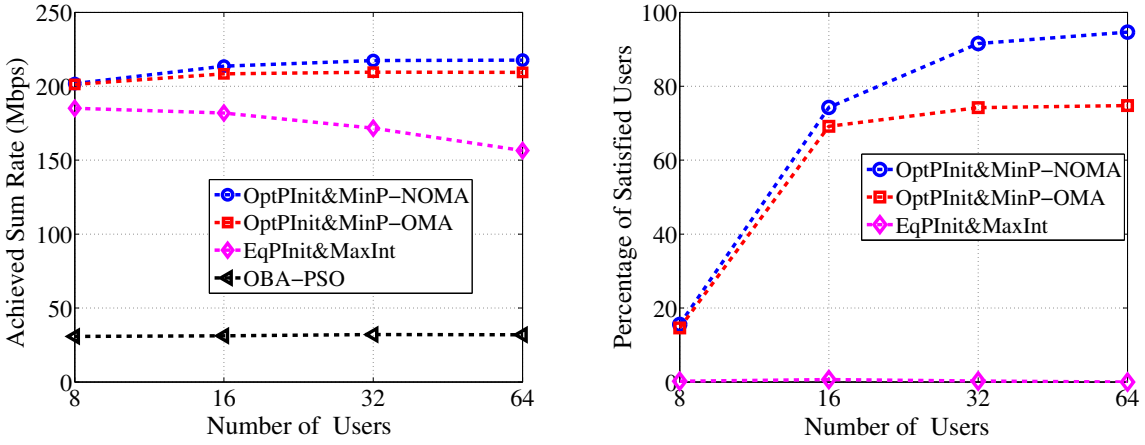
4. Trouver la position 3D de l'UAV et les valeurs de puissance dans les liens d'accès et de retour en résolvant numériquement un problème d'optimisation.

5. Si les besoins des utilisateurs en matière de débit ne peuvent être satisfaits avec le budget de puissance du drone, effectuer une étape d'appariement NOMA pour améliorer encore les débits de données obtenus par les utilisateurs.

Exemples de Résultats

La méthode proposée dans ce chapitre est dénommée OptPInit&MinP-NOMA. La performance d'une variante de cette méthode qui n'effectue pas l'étape d'appariement NOMA est également présentée, et est désignée par OptPInit&MinP-OMA. À titre de comparaison, deux méthodes différentes sont également simulées :

- OBA-PSO: cette solution est basée sur l'étude de [25]. Dans l'OBA-PSO, la bande de fréquence disponible est divisée orthogonalement entre les liaisons d'accès et de retour pour éviter les interférences entre les liaisons. Le placement 3D de l'UAV est ensuite effectué à l'aide de l'algorithme d'optimisation par essaim particulaire (PSO).
- EqPInit&MaxInt: il s'agit d'une solution simplifiée de la méthode proposée dans ce chapitre. EqPInit&MaxInt n'optimise pas l'affectation des sous-bandes dans la liaison d'accès, suppose que le budget de puissance du drone est suffisant pour garantir les exigences des débits des utilisateurs et n'effectue pas d'appariement NOMA.



(a) (b)

Figure 6 – (a) Somme des débits atteints, et (b) pourcentage d'utilisateurs satisfaits, en fonction de K pour $R^{\text{req}} = 220$ Mbps, $P_{UAV}^{\max} = 1W$ et $P_{MBS}^{\max} = 4W$

La Fig. 6 compare les performances des différentes méthodes pour un nombre d'utilisateurs allant de 8 à 64. Il a été supposé que 75% des utilisateurs avaient un besoin de 132 Mbps alors que les 25 % restants ont un besoin de 88 Mbps, ce qui donne $R^{\text{req}} = 220$ Mbps, pour toutes les valeurs de K. Fig. 6a montre que OBA-PSO donne le débit total le plus bas pour toutes les valeurs considérées de K. En fait, en raison de la division orthogonale du spectre entre les liaisons d'accès et de retour, la quantité de bande passante attribuée à chaque utilisateur est intrinsèquement plus petite que celle de la méthode que nous proposons. Contraints par le budget de puissance UAV, les utilisateurs ne peuvent pas être servis avec un débit de données suffisant. En revanche, les méthodes que nous proposons, basées sur le backhauling sans fil IBFD, permettent d'obtenir des débits de données beaucoup plus élevés. Cependant, EqPInit&MaxInt obtient le débit de données le plus faible parmi ces méthodes. Au fur et à mesure que K augmente, le débit obtenu avec EqPInit&MaxInt se détériore. D'autre part, le débit atteint par OptPInit&MinP-OMA et OptPInit&MinP-NOMA augmente en fonction de K parce que ces méthodes peuvent mieux exploiter la diversité multi-utilisateurs, OptPInit&MinP-NOMA étant plus performant que son homologue OMA. Pour $K = 64$, OptPInit&MinP-NOMA surpasse OptPInit&MinP-OMA, EqPInit&MaxInt et OBA-PSO de 8, 61 et 186 Mbps, respectivement. En ce qui concerne le pourcentage d'utilisateurs ayant reçu leur

débit de données, c'est-à-dire le pourcentage d'utilisateurs satisfaits, la Fig. 6b montre que EqPInit&MaxInt ne peut satisfaire aucun utilisateur. La performance de OBA-PSO n'est pas indiquée puisque le pourcentage de satisfaction est égal à zéro pour toutes les valeurs de K. OptPInit&MinP-NOMA atteint un pourcentage de satisfaction moyen de 94 % pour $K = 64$, surpassant OptPInit&MinP-OMA de près de 20%, ce qui montre l'avantage du couplage NOMA proposé.

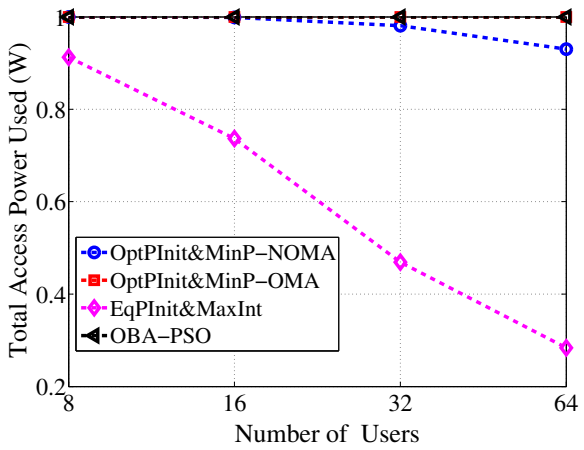


Figure 7 – Puissance utilisée par l'UAV en fonction de K pour $R^{\text{req}} = 220$ Mbps, $P_{UAV}^{\max} = 1W$ et $P_{MBS}^{\max} = 4W$

Dans la Fig. 7, la puissance d'émission nécessaire de l'UAV est présentée en fonction du nombre d'utilisateurs K. La puissance pour EqPInit&MaxInt est une fonction décroissante de K car le débit obtenu diminue également avec K pour cette méthode. OptPInit&MinP-OMA et OBA-PSO consomment tous deux le budget total pour maximiser le taux atteint. D'autre part, la puissance nécessaire pour OptPInit&MinP-NOMA est inférieure à celle de son homologue OMA lorsque $K = 32$ ou 64. Par conséquent, l'étape d'appariement NOMA permet non seulement d'augmenter le débit de données obtenu, mais peut le faire tout en consommant moins de puissance d'émission.

Conclusions et Perspectives

Dans cette thèse, nous avons abordé plusieurs problèmes complexes de gestion des ressources pour les futurs réseaux de communication sans fil.

Nous avons d'abord étudié un système de communication sans fil composé d'utilisateurs ayant des caractéristiques de trafic hétérogènes. Le premier groupe est constitué d'utilisateurs BE cherchant à maximiser les débits obtenus tout en préservant l'équité du système. Le second groupe est constitué d'utilisateurs RT qui visent à recevoir une certaine quantité de bits de données avant l'expiration de leur limite de latence. Pour améliorer la performance du système, un système d'antennes distribuées bénéficiant de NOMA a été envisagé. Pour résoudre le problème de l'allocation des ressources, nous avons d'abord proposé un algorithme glouton de faible complexité qui vise à satisfaire d'abord les utilisateurs RT, et si possible, à servir les utilisateurs BE. Nous nous sommes ensuite concentrés sur le problème de l'attribution des sous-bandes et des antennes et

avons proposé une solution basée sur la théorie de l'appariement pour résoudre ce problème. À l'aide des résultats de simulation, nous avons montré que les deux solutions proposées améliorent considérablement les performances du système, en particulier lorsqu'il s'agit d'accroître le niveau de satisfaction des utilisateurs RT, en surpassant les méthodes conventionnelles.

Ensuite, nous avons étudié l'accès non coordonné au spectre dans les réseaux sans fil. Nous avons considéré un réseau auto-organisé et proposé une nouvelle solution pour la répartition non coordonnée des canaux et de la puissance entre les APs déployés. Nous avons envisagé un réglage NOMA pour l'accès au canal, ce qui permet à plusieurs APs d'accéder au même canal et de recevoir une récompense non nulle. Nous avons modélisé et résolu le problème de l'accès non coordonné aux canaux en utilisant le cadre MP-MAB avec des récompenses variables sur les canaux à travers les APs, des jeux multiples et une récompense non nulle en cas de collision. Le problème de contrôle de la puissance a été modélisé et résolu en utilisant le cadre MP-MAB avec des récompenses variables sur les différents niveaux de puissance des APs. L'objectif de l'algorithme proposé est d'optimiser les performances du système tout en réduisant les collisions de transmission. En utilisant des dérivations théoriques, nous avons prouvé que la méthode proposée entraîne des regrets sous-linéaires. Nous avons ensuite validé les performances de l'algorithme proposé et les résultats théoriques à l'aide de simulations numériques.

Enfin, nous avons étudié un réseau de communication assisté par un drone. Un droneest déployé pour servir les utilisateurs qui ne peuvent pas être couverts par une station de base terrestre. Comme un drone n'a pas de liaison de retour câblée vers le réseau central, il doit s'appuyer sur une liaison de retour sans fil. C'est pourquoi nous avons accordé une attention particulière à cette liaison de retour sans fil entre le drone et une station de base servant de passerelle vers le réseau central. Pour augmenter les performances du système et l'efficacité spectrale, nous avons considéré un backhaul intra-bande, des capacités full-duplex du drone et la technique NOMA pour un accès multiple. Pour optimiser la position du drone et résoudre les problèmes d'allocation de ressources et de puissance dans les liaisons de backhaul et d'accès, nous avons formulé et résolu un problème d'optimisation qui minimise la puissance d'émission du drone. Le problème d'optimisation résolu prend en compte le besoin de l'utilisateur en termes de débit, la contrainte de capacité de backhaul, et les budgets de puissance d'émission pour le drone et la station de base. En utilisant des résultats de simulation, nous avons validé la supériorité des solutions élaborées par rapport aux techniques précédemment proposées dans la littérature.

De nombreuses perspectives peuvent être envisagées pour compléter ce travail. L'extension de l'étude sur le trafic mixte aux réseaux à grande échelle est très prometteuse. En effet, les futurs réseaux de communication devroent fournir des services à des milliards de dispositifs de type machine connectés [9], en plus des utilisateurs humains. Dans les réseaux à grande échelle, l'introduction d'entités intelligentes dans le réseau améliorerait considérablement les performances du système. L'étude de l'accès non coordonné au spectre peut également être étendue pour être applicable dans les réseaux à grande échelle. De plus, les travaux sur les réseaux assistés par des drones peuvent être poursuivis en optimisant la trajectoire du drone pour améliorer davantage les performances du système. Ces travaux peuvent également être élargis en envisageant des réseaux assistés par des drones multiples.

Chapter 1

Introduction

1.1 Motivations

The demand for ubiquitous wireless access has been greatly increasing in recent years due to the exponential growth in the number of intelligent connected devices such as mobile phones, autonomous cars and sensors. In addition to the proliferation of wireless devices, the emerging wireless use cases conceived for a large proportion of these new devices significantly differ from traditional voice or multimedia services. Examples of the emerging use cases include virtual and augmented reality applications, connected autonomous vehicles, remote controlled robots, surveillance and monitoring sensors and drones. As some of these use cases require a large data rate, their introduction in wireless communication networks will strain the capacity of existing wireless cellular systems. Moreover, in addition to applications requiring high data rates, future communication networks must be able to provide ultra-reliable low-latency communications (URLLC) to services having such requirements [1, 2]. Connected autonomous vehicles, remote surgeries and factory automation are illustrative examples of services requiring URLLC. In addition to URLLC, there is a need to support a massive number of connected devices and sensors of which consists the Internet of Things (IoT) environment [3, 4]. In a nutshell, future communications networks must be properly modeled, designed, and optimized to be able to support the quality of service (QoS) requirements of enhanced mobile broadband (eMBB) services, URLLC and massive machine-type communications (mMTC) [5, 6]. By expanding the capacities of cellular networks and providing the needed QoS requirements to the various services, the potential of the emerging use cases, capable of greatly improving the quality of life, can be unleashed.

To be able to fulfill the requirements of emerging applications, current wireless networks must evolve and adopt new technologies. In particular, future communication networks will be based on novel spectrum access and multiple access techniques, dense cell deployment, self-organizing capabilities for the deployed cells, millimeter wave (mmWave) communications, unmanned aerial vehicle (UAV) aided communications, and device-to-device (D2D) communications. How to design a harmonious wireless system leveraging these new techniques to adaptively optimize the use of wireless resources and provide the needed QoS requirements is an important question that has been given a lot of attention in the recent literature.

1.1.1 Radio Access Techniques

Radio access can be partitioned into two main categories: grant-based and grant-free access [26]. In grant-based radio access, devices contend over random access channels (RACH) to reserve radio resources before sending data over the granted resources. Since most IoT applications require short-packet transmissions, a grant-based connectivity for these applications results in a significant signaling overhead. Moreover, with the potential large number of active IoT devices, radio access congestion increases, resulting in communication delays and unnecessary energy consumption. Recently, grant-free communications for short-packet transmissions has been proposed and studied [26]. Grant-free radio access allows each device to transmit whenever it has new data without first reserving a radio resource. However, the problem of transmission collisions needs to be addressed. If properly designed, grant-free access techniques have the potential to enhance system performance by decreasing communication delays and signaling overhead.

1.1.2 Multiple Access Techniques

Future communication networks systems will exploit new multiple access schemes that can enhance performance, increase spectral efficiency, network capacity and the number of served users. These novel multiple access techniques fall under the umbrella of non-orthogonal multiple access (NOMA) [27]. As indicated by its name, NOMA aims at simultaneously serving multiple users on the same resource block. Depending on the considered NOMA category, the non-orthogonal multiplexing can be performed by exploiting the power domain or the code domain [28, 29, 30, 31, 32, 33]. By allowing for non-orthogonal user scheduling, NOMA allows for a higher number of served users while using the same amount of bandwidth. Moreover, NOMA increases system capacity, fairness and spectral efficiency.

1.1.3 Distributed Antennas and Small Cell Deployment with Self-Organizing Capabilities

In the third generation (3G) and the fourth generation (4G) of wireless communication systems, cellular networks relied on macro base stations (MBSs) to serve large geographical areas (few squared kilometers). However, such a deployment of base stations (BSs) may not achieve the optimal capacity and coverage [34]. Moroever, to serve cell-edge users, MBSs would require the use of high amounts of transmit power, thus increasing inter-cell interference. To overcome these challenges, a promising concept is to reduce the cell size by deploying low-power distributed antennas or small base stations (SBSs) [35, 36, 23, 37]. Such distributed antenna systems (DASs) decrease the average distance between the users and the serving antennas, allowing for high coverage, capacity and lower consumed power.

To partition resources among distributed antennas or SBSs, two main approaches exist. The first one relates to a central controller or a cloud that optimizes network resource partitioning among the distributed antennas [23, 38, 39]. The second approach consists of equipping the distributed antennas or the SBSs with self-organization capabilities [40, 41], allowing them to optimize their resource use in a distributed manner. This self-organizing approach limits human intervention and reduces planning and maintenance costs, which is of utmost importance in networks with dense SBSs deployment.

1.1.4 UAV-Aided Communications

The use of UAVs as flying BSs has emerged as a promising solution to provide reliable and cost-effective communications [42]. UAVs can be quickly and efficiently deployed to support cellular networks, making them ideal candidates to support public safety networks and disaster relief. Due to their high flying altitudes, UAVs can establish line-of-sight (LOS) communication links with ground users, thus enhancing the users QoS. Moreover, thanks to their mobility capabilities, UAVs can enhance network coverage and the number of served users by approaching users when needed.

1.2 Thesis Contributions

The main goal of this thesis is to leverage optimization and reinforcement learning techniques to optimize the resource allocation in next-generation wireless cellular networks. We focus on spectrum management for future communication networks, where users having different QoS requirements must be served. We further address some of the challenges that arise in the context of grant-free or uncoordinated multiple access, in addition to self-organizing networks (SONs). Furthermore, we study the use of UAVs as aerial base stations in wireless communication systems. Specifically, we tackle the problem of UAV placement, frequency and power allocation with backhaul consideration for UAV-aided communication networks.

The main contributions of this thesis can be summarized as follows:

- First, we study a wireless system consisting of users having different characteristics. According to their requirements, we distinguish between two different user types. The first type of users, called best-effort (BE) users, aim at maximizing their achieved rates. For these users, the system must ensure a certain level of fairness between them. The second type of users, called real-time (RT) users, seek receiving a certain quantity of data before the expiration of their latency limits. RT users applications are more urgent than BE ones, hence RT users are given a higher priority in the solutions. To serve the two categories of users, a DAS is considered so as to enhance system performance. Moreover, to increase performance and spectral efficiency, NOMA scheduling is performed. In fact, some RT applications have very strict latency constraints but require low data rates. Thus, these RT users will not make full advantage of a whole frequency band. By considering NOMA, multiple users can be scheduled on the same band, not jeopardizing spectral efficiency. In this first problem, our contributions can be summarized as follows:
  - We first propose a low-complexity greedy algorithm to perform resource allocation for the mixed traffic system.
  - We then focus on subband and antenna assignment and propose an efficient solution for this problem using matching theory. To the best of our knowledge, no previous study has considered the use of matching theory to resolve the mixed traffic resource allocation problem, combining DAS and NOMA.
  - We show, using simulation results, that both proposed solutions, i.e., the greedy algorithm and the matching-based solution, greatly outperform conventional methods, especially in terms of maximizing the satisfaction level of RT users.

- Second, we study uncoordinated spectrum access in wireless communication networks. We first focus on a SON where distributed SBSs partition the existing spectrum between them to increase the achieved system capacity while limiting inter-cell interference (ICI). Then, we study grant-free communications where users organize their transmissions, without any communication or coordination between them, so as to optimize their performance and reduce transmission collisions. For both settings, we propose a solution based on the multi-armed bandits (MAB) framework [11], which is closely related to reinforcement learning [43]. In this second problem, our contributions can be summarized as follows:
  - For the first setting, a novel solution for the uncoordinated channel and power allocation problems in a SON is proposed. The proposed technique is based on the multi-player multi-armed bandits (MP-MAB) framework and does not require any communication or coordination between the access points (APs). The case of varying channel rewards across APs is considered. In contrast to previous work on channel allocation using MAB, APs are permitted to choose multiple channels for transmission. Moreover, NOMA is used to allow multiple APs to access each channel simultaneously. This results in a MP-MAB model with varying channel rewards, multiple plays and non-zero reward on collision.
  - For the second setting, an algorithm based on the MP-MAB framework is also proposed to solve the uncoordinated spectrum access problem. The case of varying channel rewards across users is considered. In contrast to previous work, we allow users to choose multiple channels for transmission, resulting in a MAB model with multiple plays.
  - The proposed solutions in both settings have a sub-linear regret, validated by simulation results.
- Third, we focus on UAV-aided communications. Specifically, we study a system where a UAV is deployed to serve users in an area not covered by terrestrial BSs. We give particular attention to the wireless backhaul link between the UAV and an MBS, serving as gateway to the core network. In fact, this backhaul link is vital to receive the needed data to serve users. To increase spectral efficiency, we consider an in-band backhaul, where the same frequency band is used in both the access and the backhaul links. Furthermore, we assume that the UAV is equipped with full-duplex (FD) capabilities, allowing it to transmit and receive data at the same time. The purpose of this work is to find the optimal UAV placement as well as the best subband and power allocation to minimize the needed transmit power of the UAV while meeting user rate requirements. Here also, NOMA is used to enhance system performance. Our contributions in this third problem can be summarized as follows:
  - We formulate and solve an optimization problem that minimizes the UAV transmit power. This problem takes into account the rate requirement per user, the backhaul constraint, and the transmit power budget constraints for the UAV and the MBS.
  - A novel framework is introduced to find the optimal UAV placement, the best bandwidth assignment in both the access and the backhaul links, and the optimal power values in both links.

– When the UAV power budget cannot satisfy all users simultaneously, a NOMA pairing algorithm is proposed in order to maximize the achieved sum rate and the number of satisfied users.

– We validate, using simulation results, the superior performance of the proposed techniques, when compared to previously proposed algorithms.

1.3 List of Publications

As a byproduct of the above contributions, this dissertation has led to the following publications:

1.3.1 Journal Publications

- **M. J. Youssef**, V. V. Veeravalli, J. Farah, C. Abdel Nour, C. Douillard, "Resource Allocation in NOMA-based Self-Organizing Networks using Stochastic Multi-Armed Bandits", under review for publication in *IEEE Transactions on Communications*.

- **M. J. Youssef**, J. Farah, C. Abdel Nour, C. Douillard, "Full-Duplex and Backhaul-Constrained UAV-Enabled Networks using NOMA", *IEEE Transactions on Vehicular Technology*, vol. 69, no. 9, pp. 9667-9681, Sept. 2020.

- **M. J. Youssef**, J. Farah, C. Abdel Nour, C. Douillard, "Resource Allocation in NOMA Systems for Centralized and Distributed Antennas with Mixed Traffic using Matching Theory", *IEEE Transactions on Communications*, vol. 68, no. 1, pp 414-428, Jan., 2020.

1.3.2 Conference Publications

- **M. J. Youssef**, V. V. Veeravalli, J. Farah, and C. Abdel Nour, "Stochastic multi-player multi-armed bandits with multiple plays for uncoordinated spectrum acces", *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, London (Virtual)*, United Kingdom, Aug. 2020.

- **M. J. Youssef**, C. Abdel Nour, J. Farah, and C. Douillard, "Backhaul-constrained resource allocation and 3D placement for UAV-enabled networks", *IEEE 90th Vehicular Technology Conference*, Honolulu, Hawaii, United States, Sept. 2019.

- **M. J. Youssef**, J. Farah, C. Abdel Nour, and C. Douillard, "Resource allocation for mixed traffic types in distributed antenna systems using NOMA", *IEEE 88th Vehicular Technology Conference*, Chicago, United States, Aug. 2018.

Next, we give an overview on the structure of this thesis.

1.4 Thesis Structure

The rest of this manuscript is organized as follows:

- **Chapter 2** gives a general overview of the relevant topics to this thesis. In the first part of this chapter, a general overview on future communication networks is given. The second part of the chapter is dedicated to presenting the analytical frameworks some of the solutions of this thesis are based on, namely game theory, Markov chains, matching theory and multi-armed bandits.

- **Chapter 3** presents different resource allocation solutions for a distributed antenna wireless system with mixed traffic benefiting from NOMA scheduling. The conceived solutions aim at maximizing the number of RT users having received their requirements before the expiration of their latency limit, while enhancing both the data rates and fairness achieved by BE users.

- **Chapter 4** introduces an approach based on the MAB framework for uncoordinated spectrum access. In the first considered setting, the MAB framework is used to solve the uncoordinated spectrum access and power control in a SON benefiting from NOMA scheduling. The second considered setting consists of users trying to access multiple channels simultaneously without any communication or coordination. The goal in both settings is to optimize system performance and achieve a stable partitioning of resources between the competing entities, without any communication, and by avoiding collisions.

- **Chapter 5** studies a UAV-enabled communication network where the deployed UAV aims at serving users with their rate requirements while receiving the necessary data from the backhaul link. The proposed solutions benefit from NOMA, FD communications, and in-band wireless backhauling. The goal of these solutions is to find the optimal placement of the UAV and the optimal frequency and power allocation in the access and the backhaul links, so that users are served with their rate requirements while necessitating the least amount of UAV transmit power.

- **Chapter 6** summarizes the work presented in this thesis and provides some interesting directions for future research in the scope of this thesis.

Chapter 2

Background

In this chapter, we review the main topics that are relevant to the work presented in this thesis. The first part of this chapter presents a general background on next-generation cellular networks. In particular, an overview of novel multiple access techniques, namely non-orthogonal multiple access (NOMA), is given. We also present a general background on Internet of Things (IoT) applications and requirements. Finally, we focus on unmanned aerial vehicle (UAV)-aided communication networks and discuss the benefits and challenges of the introduction of UAVs in wireless systems. Background information on the underlying analytical techniques of the proposed solutions is presented in the second part of this chapter. Specifically, an overview of game theory basics, Markov chains, multi-armed bandits and matching theory is given.

2.1 Next-Generation Wireless Networks

The main driver of the evolution of wireless systems in the past was the need for higher data rates. However, with the proliferation of novel IoT applications, the key requirements that fifth generation (5G) and beyond cellular networks must guarantee are becoming more and more diverse. In fact, future communication networks have to support enhanced mobile broadband (eMBB) services, massive machine-type communications (mMTC) [44], and ultra-reliable low-latency communications (URLLC) [5, 6]. In particular, the main 5G key elements, illustrated in Fig. 2.1, can be summarized as:

- **Higher data rates and better coverage:** with an area capacity 1000 times higher than that provided by fourth generation (4G) mobile networks, and edge capacity exceeding that provided by 4G by 100 times.
- **High reliability:** with a packet error rate (PER) typically lower than 10^{-5}.
- **Low latency:** with a round-trip air latency of 1 ms for some latency-constrained applications.
- **Support for machine type devices:** for example autonomous vehicles, UAVs, virtual reality headsets, connected sensors.
- **Low energy consumption:** communications for energy-constrained IoT devices (e.g., sensors) must be designed in a way to extend their lifetime. Moreover, for green

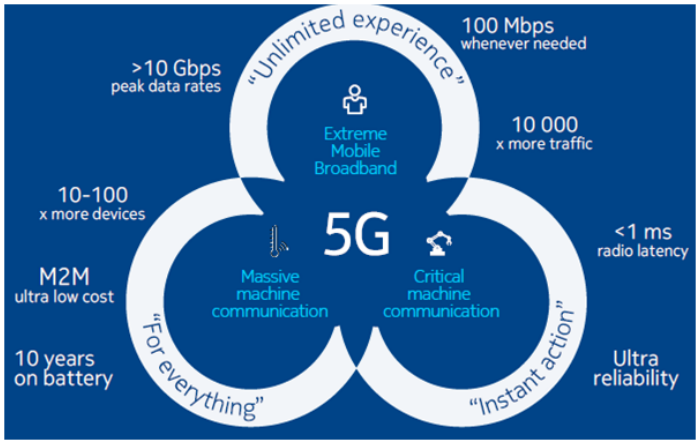


Figure 2.1 – 5G use cases and requirements

communications, future communication systems must be able to increase system capacity while maintaining a low energy consumption at the base stations (BSs) level.

To achieve these requirements, significant research efforts have been dedicated to the study of novel techniques and cellular network architectures that can cope with the rapid evolution of wireless systems. Among these, we cite:

- **Dense deployment of distributed antenna systems (DASs) and small base stations (SBSs):** Future communication networks will be characterized by a dense deployment of DASs and SBSs which will improve the area capacity and coverage, and decrease the system power consumption.
- **Use of UAVs as flying BSs:** The use of UAVs in cellular networks has been recently given a lot of attention. The applications of UAVs as flying BSs in communication networks are diverse and include among others, increasing the capacity and coverage in temporary hotspot areas, restoring communications in public safety scenarios, providing coverage to remote areas and gathering data from IoT devices.
- **More efficient use of the existing spectrum:** Novel multiple access techniques such as NOMA can be utilized to increase spectrum efficiency, the achieved rates and the number of served users.

Next, we give a general overview of NOMA, a novel multiple access technique to be used in future communication networks. This overview is followed by a discussion of IoT applications in future wireless communication networks. Finally, an overview of the use of UAVs in cellular networks is presented.

2.1.1 Non-Orthogonal Multiple Access

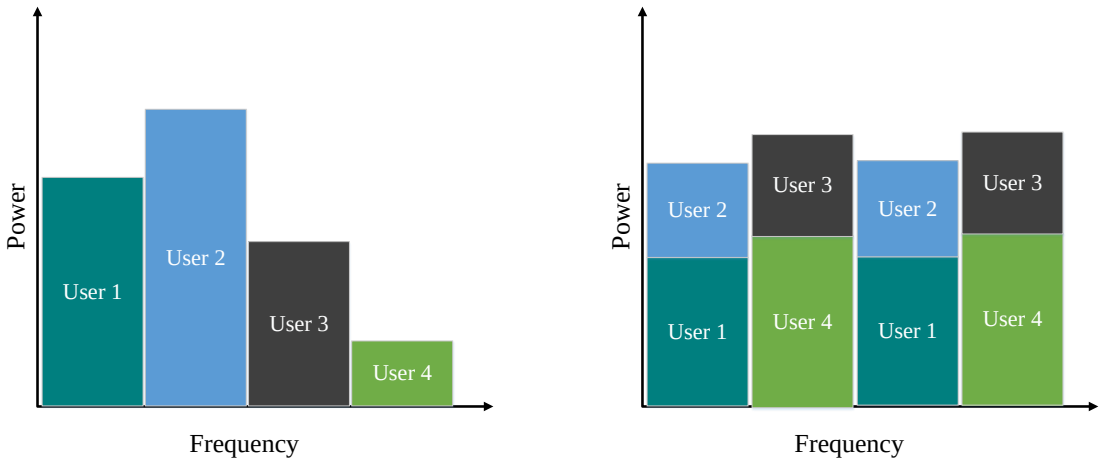
In this section, the basic principles of NOMA are introduced.

NOMA is based on the concept that multiple users can be simultaneously served on the same orthogonal resource block, whether the same resource block is a frequency band, a timeslot or an orthogonal spatial degree of freedom [45]. Multiple categories of NOMA have been considered in the literature, among which we cite:

- power-domain NOMA (PD-NOMA) [28, 29, 30],
- sparse code multiple access (SCMA) [31, 46, 47],
- pattern division multiple access (PDMA) [32],
- interleave division multiple access (IDMA) [33].

This thesis focuses on PD-NOMA, the concept of which will be discussed next. Note that for simplicity, PD-NOMA will be denoted by NOMA in this thesis.

The basic concept of NOMA relies on exploiting the power domain to serve multiple users non-orthogonally on the same orthogonal frequency division multiplexing (OFDM) resource block. NOMA relies on superposition coding (SC) [7] at the transmitter side and on successive interference cancellation (SIC) [8] at the receiver side. Although the key components of NOMA, i.e., SC and SIC, have been invented more than two decades ago, the application of NOMA scheduling in wireless communication systems, i.e., removing the orthogonality between users, is relatively new. An illustration comparing NOMA and orthogonal multiple access (OMA) scheduling is shown in Fig. (2.2).



(a) OMA scheduling (b) NOMA scheduling

Figure 2.2 – Comparison between OMA and NOMA scheduling

2.1.1.1 Basics of Downlink NOMA

In the downlink, the BS uses SC to transmit a combination of the superposed messages of scheduled users. At the receiver side, users perform SIC to retrieve their message.

Let $\mathcal{K}$ and $\mathcal{S}$ be the set of users and the set of subbands in the system, respectively. Moreover, let B_c be the bandwidth of each subband $s \in \mathcal{S}$. Also, let $\mathcal{K}_s = \{k_s(1), \ldots, k_s(n), \ldots, k_s(K_s)\}$ be the set of users scheduled over subband s, where $k_s(n)$ is the n^{th} user scheduled over subband s and $|\mathcal{K}_s| = K_s$. Denoting by $x_{k,s}$ the modulated signal relative to user k over subband s, the signal transmitted by the BS over subband s, x_s, is the superposition of the signals of all users in $\mathcal{K}_s$. The superposed signal can be expressed as:

$$x_s = \sum_{k \in \mathcal{K}_s} x_{k,s}, \tag{2.1}$$

with $P_{k,s} = \mathbb{E}\left[|x_{k,s}|^2\right]$ the power allocated to user k over subband s.

The signal received at the level of user $k \in \mathcal{K}_s$ is:

$$y_{k,s} = h_{k,s}x_s + w_{k,s}, \tag{2.2}$$

where $h_{k,s}$ is the channel gain between the BS and user k on subband s and $w_{k,s}$ is the noise experienced by user k on subband s.

Without loss of generality, let the users in $\mathcal{K}_s$ be sorted in the decreasing order of channel gain. Signal separation at the receiver side is done using SIC decoding in the increasing order of channel gains. In other words, user $k_s(K_s)$ having the lowest channel gain on s, does not perform SIC and decodes its message directly by treating the signals of all other users as noise. On the other hand, user $k_s(n),\ 1 \leq n < K_s$ decodes the signals of users coming before him in the decoding order, i.e., users with indices $n < i \leq K_s$ before decoding its own signal. Hence, user $k_s(1)$, having the highest channel gain among users in $\mathcal{K}_s$, can receive its signal without any interference. Assuming successful SIC decoding and no error propagation at the receiver side, the rate achieved by user $k_s(n)$ on subband s is given by:

$$R_{k_s(n),s} = B_c \log_2 \left(1 + \frac{P_{k_s(n),s} h^2_{k_s(n),s}}{\sum\limits_{j=1}^{n-1} P_{k_s(j)} h^2_{k_s(n),s} + N_0 B_c} \right), \tag{2.3}$$

where N_0 is the noise power spectral density and B_c is the subband bandwidth.

SIC increases significantly the complexity at the receiver side. Moreover, the achieved performance gain diminishes when the number of non-orthogonally multiplexed users increases beyond two, as was shown in [28, 48]. Therefore, as was done in most work in the literature, the number of non-orthogonally multiplexed users on each subband is limited to two in this thesis. Hence, $\mathcal{K}_s = \{k_s(1), k_s(2)\}$ and the rates achieved by $k_s(1)$ and $k_s(2)$ over subband s are respectively given by:

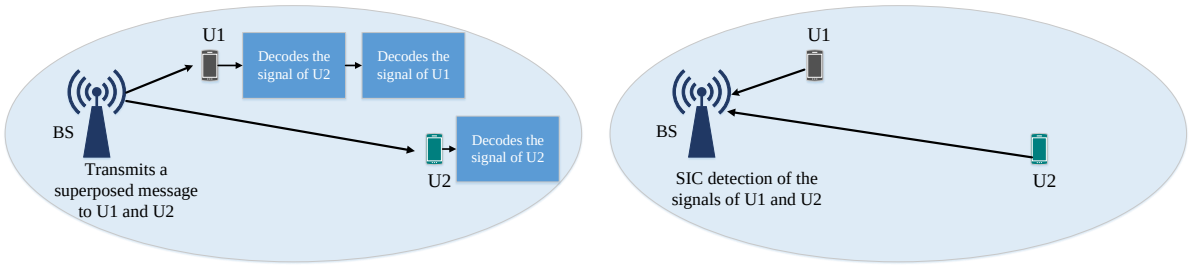
$$R_{k_s(1),s} = B_c \log_2 \left(1 + \frac{P_{k_s(1),s} h^2_{k_s(1),s}}{N_0 B_c} \right), \tag{2.4}$$

$$R_{k_s(2),s} = B_c \log_2 \left(1 + \frac{P_{k_s(2),s} h^2_{k_s(2),s}}{P_{k_s(1),s} h^2_{k_s(2),s} + N_0 B_c} \right). \tag{2.5}$$

As seen from Eq. (2.4) and Eq. (2.5), only the user with a lower channel gain suffers from co-channel interference. Fig. 2.3a shows a setting with two users scheduled using NOMA on the same subband, where only the user closer to the BS, hence having a better channel gain, denoted by strong user, performs SIC. The user with a lower channel gain, denoted by weak user, decodes its signal directly by treating the signal of the strong user as noise.

As seen from Eq. (2.4) and Eq. (2.5), both the choice of the multiplexed users over subband $s \in \mathcal{S}$ and the amount of power allocated to each user affect system performance. Hence, to reap the performance gain promised by NOMA systems, the frequency resource allocation, scheduling, user pairing and power allocation must be carefully designed. Note that it has been shown in [49] that, to guarantee successful decoding at the user side, also called SIC stability, the power allocated to the weak user must be higher than the power allocated to the strong user, i.e., the power allocation must respect the following constraint:

$$P_{k_s(2),s} > P_{k_s(1),s},\ \forall s \in \mathcal{S}. \tag{2.6}$$



(a) Downlink NOMA (b) Uplink NOMA

Figure 2.3 – SIC decoding

In the opposite case, as shown in [30], the outage probabilities of the users will be always one.

2.1.1.2 Basics of Uplink NOMA

In the uplink, the BS sends control signals to the transmitting users to determine power allocation. Once the transmit power is determined, each transmitting user sends its message. The BS in this case receives a signal constituted by the superposition of the signals of transmitting users. Using SIC, the BS decodes the messages of the users according to the decreasing order of channel gains. In other words, the signal of the user having the highest channel gain is first decoded, while considering the signals of all other users as interference. The signal of the user with the highest channel gain is then subtracted from the total received signal before decoding the signals of the remaining users. In uplink NOMA, the signal of the user with the lowest channel gain is decoded without interference. Considering the case of two scheduled users on subband s, $k_s(1)$ and $k_s(2)$ with $k_s(1)$ being the user with the higher channel gain, the achieved rates in uplink NOMA are given by:

$$R_{k_s(1),s} = B_c \log_2 \left(1 + \frac{P_{k_s(1),s} h^2_{k_s(1),s}}{P_{k_s(2),s} h^2_{k_s(2),s} + N_0 B_c} \right), \tag{2.7}$$

$$R_{k_s(2),s} = B_c \log_2 \left(1 + \frac{P_{k_s(2),s} h^2_{k_s(2),s}}{N_0 B_c} \right), \tag{2.8}$$

where $P_{k_s(.),s}$ is the transmit power of user $k_s(.)$ on subband s. The basic principle of NOMA in the uplink is illustrated in Fig. 2.3b.

To achieve SIC stability in uplink NOMA, i.e., successful decoding at the BS level, the following condition must be satisfied in the case of two scheduled users on subband s:

$$P_{k_s(1),s} h^2_{k_s(1),s} > P_{k_s(2),s} h^2_{k_s(2),s}. \tag{2.9}$$

Although most studies related to NOMA focus on the downlink setting, there has been some work that recently considered the application of NOMA in uplink scenarios. For

example, [50] derived the power allocation factors using the Karush-Kuhn-Tucker (KKT) conditions to maximize the sum throughput. In [51], the authors proposed an algorithm for subband and power allocation to maximize the system throughput. However, most studies on uplink NOMA assume that subband assignment and power allocation are performed in a centralized manner, with the BS carefully allocating power and users over multiple channels with full channel state information (CSI). Hence, these studies are not suitable for uncoordinated spectrum access using NOMA.

Recently, [52] proposed a power allocation algorithm for random access using NOMA where the BS does not coordinate uplink transmissions. Assuming that all users have a target signal-to-interference-plus-noise-ratio (SINR) requirement Γ, a set of predefined received power levels v_l, allowing up to L users to simultaneously access each subband are found according to:

$$v_l = \Gamma N_0 B_c (\Gamma + 1)^{L-l}. \tag{2.10}$$

User k choosing power level v_l on subband s transmits using the following transmit power value:

$$P_{k,s,l} = \frac{v_l}{h_{k,s}^2}. \tag{2.11}$$

The work in [52] shows that, if the L users coordinate their transmissions in such a way that each user chooses a unique received power level, all L users achieve the same SINR requirement Γ.

In Chapter 4, we extend the power allocation scheme proposed in [52] to the case of L distinct SINR requirements. In the context of a self-organizing network (SON), we then propose an algorithm that allows access points (APs) to coordinate their transmissions so that each AP transmits using a unique SINR requirement.

2.1.2 Internet of Things and New Mobile Traffic Characteristics

Future wireless communication networks are expected to support billions of connected machine-type devices (MTDs) [9], giving rise to the IoT. These MTDs will be able to connect with each other and with the BSs over wireless links, allowing them to collect and exchange information in real-time. The applications considered for IoT devices are countless: smart grids, smart cities, smart homes, autonomous cars, health sensors, etc. These novel applications will help significantly enhance the quality of people's lives. However, for efficient deployment of IoT services, many challenges such as data analytics, transmission capabilities, resource allocation, security and privacy, must be solved. In particular, the mobile traffic requirements of IoT devices are application-dependent. For example, the necessary traffic for efficient deployment of autonomous cars must benefit from a high reliability and a low latency, while normally requiring low data rates. On the other hand, applications like virtual or augmented reality require both high data rates and low latency. Hence, the quality of service (QoS) requirements of IoT devices are very different from one another, and even more so from the QoS requirements of human users. To this effect, coming up with resource allocation techniques for IoT devices that can meet their diverse traffic requirements, in terms of throughput, latency and reliability, is of utmost importance.

2.1.3 UAV-Aided Communication Networks

Recently, the use of UAVs has received significant attention in various application domains spanning the military, surveillance and monitoring, delivery, and telecommunication domains, among others [42, 53, 54, 55]. Depending on the requirements of each application, different types of UAVs must be used. Based on their flying altitudes, UAVs can be partitioned into two categories: high altitude platforms (HAPs) and low altitude platforms (LAPs). HAPs are deployed on altitudes exceeding 17 km and are quasi-stationary. LAPs on their turn can occupy much lower altitudes, from 50 m and up to a few kilometers, have a flexible deployment and can move quickly. On the one hand, LAPs benefit from a rapid deployment, when compared to HAPs, making them suitable for time-sensitive applications. On the other hand, HAPs benefit from a longer battery life, making them suitable for long-term applications. UAVs can be also partitioned among rotary-wing and fixed-wing UAVs. Fixed-wing UAVs are heavier than rotary-wing UAVs, benefit from higher speeds, and must keep moving in order to stay aloft. On the contrary, rotary-wing UAVs can hover over a certain area if the application requires it.

In wireless communication systems, properly deployed and operated UAVs can provide solutions to many challenges. Indeed, UAVs can be deployed as aerial BSs to provide on-demand communications to areas in need. Thanks to their adjustable flying altitude, the probability of establishing line-of-sight (LOS) links between the UAV and its assigned users is increased, resulting in more efficient communications, when compared to ground BSs. Moreover, due to the flexible deployment of LAP UAVs, the capacity and coverage of UAV-enabled communication networks can be enhanced. Due to these numerous advantages, the use of UAVs as aerial BSs in wireless communication networks admits many potential applications. For example, UAVs can increase the capacity and coverage in temporary hotspot areas such as sport stadiums. UAVs can also restore communications in public safety scenarios to support search and rescue operations when terrestrial communication networks get damaged from natural disasters for example. UAVs can further aid in the IoT networks where devices normally have a low transmit power and are not able to communication over long ranges. In this case, UAVs can get close to IoT devices, allowing them to transmit their data efficiently.

2.1.3.1 Challenges of UAV Communications

To reap the benefits of UAVs in cellular communication systems, several challenges must be met. Among these challenges, we cite:

1. **UAV-to-ground channel modeling:** The first challenge relates to finding accurate models for the air-to-ground (A2G) channel between the UAV and the ground users. In fact, the A2G channel characteristics differ significantly from classical terrestrial channel models and any movement of the UAV can affect them. Furthermore, the A2G channel depends on the altitude of the UAV, its elevation angle and the propagation environment. Hence, finding generic and accurate channel models for UAV-to-ground communications is a challenging task.

2. **Placement of the UAV:** The optimal placement of the UAV in a wireless communication network depends on the objective of its deployment (e.g., coverage and capacity enhancement, power minimization of IoT communications). Naturally, the optimal placement of the UAV depends on the locations of the ground users and

their traffic characteristics (e.g., latency, reliability and data rate requirements). Moreover, the deployment of a UAV should take into account managing inter-cell interference from other potentially deployed UAVs as well as from terrestrial BSs. In addition, the placement of the UAV must be optimized in 3D space and must account for the UAV energy constraints. Hence, the placement problem of a UAV-BS is much more challenging than that of a terrestrial BS.

3. **Trajectory optimization:** Optimizing the flight trajectory of the UAV is a challenging task as it necessitates accounting for the channel variation, the energy and the flight constraints.

4. **Wireless backhaul connectivity:** Unlike most terrestrial communication networks, UAV-BSs must rely on a wireless backhaul connectivity to the core network. To effectively operate UAV-enabled cellular networks, this wireless backhaul link needs to be dynamically managed according to the traffic state of the network. In fact, if not configured properly, the backhaul link introduces additional interference and limits the performance of the UAV in the access link.

Next, we will discuss some recent work proposing channel models for UAV communication systems. We will also present in detail the A2G channel model that was most adopted in the literature.

2.1.3.2 Air-to-Ground Channel Modeling

The channel quality between the UAV and its users determines the performance of a UAV-enabled network in terms of capacity, coverage and power consumption. That said, the A2G channel model differs significantly from terrestrial channel models. Therefore, several recent work studied A2G channel modeling. In [56] and [57], the authors focused on path loss modeling for HAPs. The work in [56] showed that due to the high altitudes of the UAVs leading to the high elevation angles between the UAVs and the users, A2G links experience lower path loss and shadowing when compared with terrestrial communication links. In [58], the probability of having a LOS link in A2G communications was derived as a function of the elevation angle between the UAV and each user, and of the average buildings height. Some studies as in [59, 60, 61] performed measurement-based studies and identified characteristics of A2G channel links. However, the results of these studies are valid for specific system setups and do not provide a generic channel model that can be used to design and study UAV-aided communication networks.

The most widely adopted channel model for LAPs is proposed in [62]. The work in [62] showed that the path loss between the UAV and the ground user depends on:

1. the location of the UAV,

2. the location of the ground user,

3. the propagation environment (e.g., rural, suburban, urban or high-rise urban).

Depending on these factors, the work in [62] showed that the A2G communication link can be either a LOS link or a non-line-of-sight (NLOS) link. However, without additional information about the exact positions, heights and number of obstacles, classifying a link as LOS or NLOS is not straightforward. Therefore, one must consider the randomness

associated with the occurrence of LOS and NLOS links. Many works in the literature on UAV-enabled communications adopted the probabilistic path loss model given in [62, 63], where the LOS and NLOS links are considered separately according to their probabilities of occurrence. The latter is a function of the environment, the density and heights of buildings, and the elevation angle between the UAV and the ground user. The common probabilistic path loss model is based on the general geometrical statistics of different propagation environments provided by the Radio communication sector of the International Telecommunication Union (ITU-R) [64].

According to [62], the probability of having a LOS link between the UAV and ground user k is given by:

$$P_{LOS} = \frac{1}{1 + \alpha \exp\left(-\beta\left(\frac{180\theta_k}{\pi} - \alpha\right)\right)}, \tag{2.12}$$

where $\theta_k = \arctan\left(\frac{H}{r_k}\right)$ is the elevation angle, $r_k = \sqrt{(x_k - x_{UAV})^2 + (y_k - y_{UAV})^2}$ is the horizontal distance between user k and the UAV, and H is the UAV altitude. Variables α and β are constants determined by the environment (e.g., rural, urban, suburban). The NLOS probability is hence given by: $P_{NLOS} = 1 - P_{LOS}$.

Eq. (2.12) shows that the LOS probability increases with the elevation angle. Fig. 2.4 shows a snapshot of a UAV-enabled communication network with both LOS and NLOS communication links.

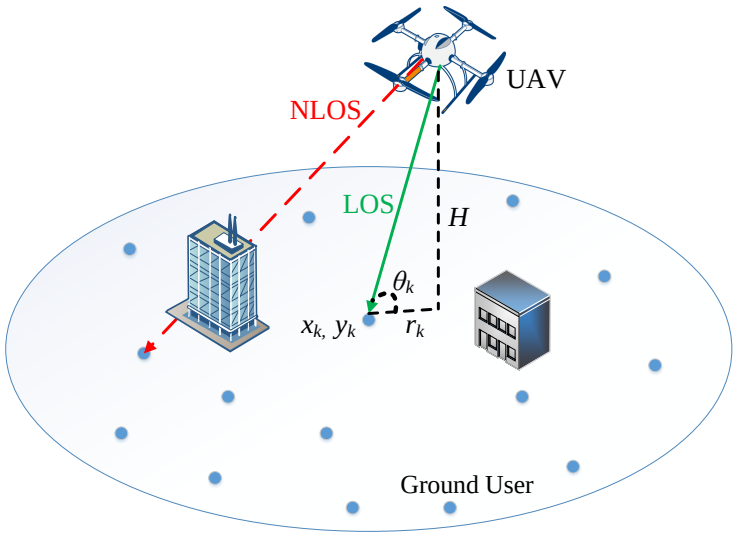


Figure 2.4 – UAV-aided communication system

Next, we discuss several concepts and techniques used to find solutions for the problems considered in this thesis. First, we briefly discuss the general concepts of game theory in Section 2.2. Then, we give an overview of Markov Chains and related concepts in Section 2.3 before presenting multi-armed bandits in Section 2.4 and matching theory in Section 2.5.

2.2 Basics of Game Theory

Game theory is a mathematical framework that studies the strategic interactions among independent selfish and rational players [65]. Game theory is used in different disciplines that include economics, politics, biology and product pricing. Recently, game theory received significant attention in the wireless communication community, due to the emergence of large-scale distributed wireless networks requiring self-organizing capabilities [10]. To solve the uncoordinated spectrum access and power control problems in a distributed manner in Chapter 4 of this thesis, we propose a solution based on game theory, since the latter can properly model interactions between different independent agents.

In its basic form, a game consists of:

- **Players:** with conflicting interests or mutual benefit, seeking to maximize their profits.
- **Actions or strategies:** set of actions available for each player.
- **Utility or payoff:** corresponds to the benefit or profit a player can obtain when playing a particular action.

Games can be broadly partitioned into two categories:

1. **Non-cooperative games:** where players have conflicting interests,
2. **Cooperative games:** where different players have mutual benefits and cooperate to enhance their performance.

Different solution concepts exist for different types of games. Next, we discuss the most commonly adopted solution concepts for non-cooperative games that have a particular importance in wireless communication networks. These concepts can be summarized by:

- **Nash equilibrium:** The Nash equilibrium [65] is a fundamental solution concept for strategic form games. It refers to the stable state of a system where multiple players interact, and where no player can profit from a unilateral deviation of its strategy, given that the strategies of other players remain unchanged. Hence, when reaching the Nash equilibrium, the action of each player is a *best response* to the actions taken by the remaining players. Therefore, the Nash equilibrium can be viewed as a stable state of the strategic interaction between the players. Different types of Nash equilibria exist, namely pure, mixed and behavioral Nash equilibria. A pure strategy Nash equilibrium involves each player playing one specific strategy unconditionally. On the other hand, a mixed strategy Nash equilibrium involves each player assigning a probability to each pure strategy. In other words, in a mixed strategy Nash equilibrium, at least one player randomizes over some or all of its pure strategies. Finally, a behavorial strategy Nash equilibrium exists in dynamic games. In such an equilibrium, each player assigns independent probabilities to its set of actions at each state of the game.
- **Pareto optimality:** The outcome of a game is said to be Pareto optimal [66] if no other outcomes provide every player with at least the same gain, while one player makes a strictly higher profit. Put differently, Pareto optimality implies a point where no player can achieve a higher gain without decreasing the gain of other players.

Next, we discuss concepts related to Markov chains and adaptive play.

2.3 Markov Chains and Related Concepts

2.3.1 Markov Chain: Definition

A Markov chain [67] is a random process that satisfies the Markov property which states that, conditioned on the present value of the random process, its future is independent from its past. Let $\{Z_0, Z_1, \ldots\} \in \mathcal{Z}$ be a sequence of discrete random variables taking values in a countable set $\mathcal{S}$, called the set space. Each random variable Z_n takes one of N possible values with $N = |\mathcal{S}|$. Definition 1 gives the formal description of a Markov chain:

Definition 1. *The process $\mathcal{Z}$ is a Markov chain if it satisfies the Markov property, i.e., if it satisfies the following condition:*

$$\mathbb{P}\left(Z_n = s \mid Z_0 = z_0, Z_1 = z_1, \ldots, Z_{n-1} = z_{n-1}\right) = \mathbb{P}\left(Z_n = s \mid Z_{n-1} = z_{n-1}\right), \quad \forall n \geq 1, \text{ and } \forall s, z_0, z_1, \ldots, z_{n-1} \in \mathcal{S}. \tag{2.13}$$

In Definition 1, $\mathbb{P}\left(Z_n = j \mid Z_{n-1} = i\right)$ is the *transition probability* of the Markov chain that describes its evolution from *state i* to *state j*. The transition matrix of the Markov chain $\mathcal{Z}$ can be defined as:

Definition 2. *The **transition matrix** $\boldsymbol{P} = \{P_{ij}\},\ \forall i, j \in \mathcal{S}$ is the $|\mathcal{S}| \times |\mathcal{S}|$ matrix of transition probabilities:*

$$P_{ij} = \mathbb{P}\left(Z_n = j \mid Z_{n-1} = i\right). \tag{2.14}$$

Moreover, the transition matrix $\boldsymbol{P} = \{P_{ij}\}$ is a stochastic matrix, i.e.:

- *$\boldsymbol{P}$ has non-negative entries, i.e., $\{P_{ij}\} \geq 0,\ \forall i, j \in \mathcal{S}$,*
- *the rows of $\boldsymbol{P}$ sum up to 1, i.e., $\sum_j P_{ij} = 1$.*

Definition 3. *The **n-step transition matrix** $\boldsymbol{P}(n)$ is the matrix of **n-step transition probabilities** $P_{ij}(n)$ defined as:*

$$P_{ij}(n) = \mathbb{P}\left(Z_n = j \mid Z_0 = i\right). \tag{2.15}$$

2.3.2 Classification of states

The development of a Markov chain is described by its transition between the different possible states. A state can first be classified as either *recurrent* or *transient*:

Definition 4. *A state i is called **recurrent** or **persistent** if:*

$$\mathbb{P}(Z_n = i \text{ for some } n \geq 1 \mid Z_0 = i) = 1, \tag{2.16}$$

*i.e., the probability of returning to state i, having started from this state, is 1. If this probability is strictly less than 1, the state i is called **transient**.*

Let $T_i = \min\{n \geq 1, \text{ such that } Z_n = i\}$ denote the time of the first visit to state i. Moreover, let $T_i = \infty$ if state i is never visited, which happens if and only if state i is transient.

Definition 5. *The **mean recurrence time** of a state i, $\omega_i = \mathbb{E}(T_i \mid Z_0 = i)$, is defined as:*

$$\omega_i = \begin{cases} \sum_n n\mathbb{P}(Z_1 \neq i, \ldots, Z_{n-1} \neq i, Z_n = i \mid Z_0 = i) & \text{if } i \text{ is persistent,} \\ \infty \quad \text{if } i \text{ is transient.} \end{cases} \tag{2.17}$$

Note that ω_i can be equal to ∞ even if state i is recurrent or persistent. A persistent state can be further classified based on the following definition.

Definition 6. *A persistent state is called **positive or non-null** if $\omega_i < \infty$ and **null** otherwise.*

Definition 7. *The **period** $d(i)$ of state i is defined as:*

$$d(i) = gcd\{n \text{ such that } P_{ii}(n) > 0\}, \tag{2.18}$$

*where gcd refers to the greatest common divisor. In other words, the period of a state i is the greatest common divisor of the time instants at which return is possible to state i. State i is said to be **periodic** if $d(i) > 1$ and **aperiodic** if $d(i) = 1$.*

Definition 8. *A state i is called **ergodic** if it is **persistent, positive or non-null, and aperiodic.***

Definition 9. *A state i is called **absorbing** if there are no outgoing transitions from the state. In other words, a state i is called **absorbing** if, once entered, it cannot be left.*

2.3.3 Classification of Chains

Next, we consider the ways that relate the states of a Markov chain with one another.

Definition 10. *State i is said to **communicate** with state j if the Markov chain may visit state j with positive probability, having started from state i. In other words, state i **communicates** with state j if, for some $m \geq 0$, we have $P_{ij}(m) > 0$. We refer to the **communication** of state i to state j by: $i \to j$. States i and j **intercommunicate** if $i \to j$ and $j \to i$. We refer to the **intercommunication** of states i and j by: $i \leftrightarrow j$.*

Definition 10 divides the states of a Markov chain into a set of *classes*. Within each *class*, all states intercommunicate, but no pair of states in different *classes* intercommunicates.

Proposition 1. *If state i is recurrent, and state i communicates with state j, i.e., $i \to j$, then state j is also recurrent. Therefore, in any class, either all states are recurrent or all states are transient. Hence, classes of a Markov chain can be classified as either recurrent or transient.*

Definition 11. *A Markov chain is **irreducible** if all its states **intercommunicate**, i.e., if it consists of one **class** only.*

Based on Proposition 1 and Definition 11, if the Markov chain is irreducible, then either all states are recurrent or all states are transient.

Definition 12. *A Markov chain is **ergodic** if it is **irreducible** and all its states are **ergodic**.*

2.3.4 Adaptive Play

Consider a game with K players that is played once at each period. Let $\mathcal{S}_k$ be the finite set of strategies or actions available for player $k \in \mathcal{K}$ and let $\boldsymbol{s}^t = (s_1^t, \ldots, s_K^t)$ be the strategy profile of all K players at time t. Based on the strategy profile $\boldsymbol{s}^t$, player k receives a utility $u_k(\boldsymbol{s}^t)$.

At each round of play, each player chooses an optimal strategy in an *adaptive* manner, based on its achieved performance in the previous round of the game. Hence, the action played in the previous round has a feedback effect on the action chosen by each player in the current round of the game. These dynamics constitute, for each player, a Markov chain whose states are the historical actions chosen in the previous periods of play.

For general K-player games, this kind of adaptive play does not necessarily converge to a Nash equilibrium. However, for certain classes of games, starting from any initial choice of strategies, a sequence of best replies leads to a pure Nash equilibrium with probability one. Once the equilibrium is reached, it becomes known as the *conventional* way of game play. This equilibrium hence constitutes an absorbing state of the Markov process.

The above analysis involves players always responding with the optimal strategy, given a certain history of actions. However, it may be the case that some or all players occasionally experiment or make mistakes. In this case, the game has no absorbing states, but has a stationary distribution π that describes the frequency with which different states are observed. As shown in [68], if the probability of experimentation or exploration or mistakes is small, this stationary distribution π is typically concentrated around one pure strategy Nash equilibrium, deemed the *stochastically stable equilibrium*. A state Z is *stochastically stable* if and only if the stationary distribution of the Markov process concentrated around Z remains positive when the probability of experimentation tends towards 0. If the probability of experimentation or exploration or mistakes is small, [68] showed that the *stochastically stable equilibrium* is observed with probability close to one.

2.3.4.1 Adaptive Play When There Are No Mistakes

Let the state of the Markov process at time t be denoted by $Z_t = \boldsymbol{s}^t$. Given this state Z_t, at time $(t+1)$, the players decide on their chosen actions and the system transitions to state Z_{t+1}. Let $p_k(.)$ be a *best reply* distribution, meaning that $p_k(s_k \mid Z_t) > 0$ if and only if s_k is the best reply strategy of user k to state Z_t and that $p_k(s_k \mid Z_t)$ is independent of time t. Assuming that players do not make mistakes, the transition probability from state Z_t to state Z' is given by:

$$P^0_{Z_t Z'} = \prod_{k=1}^{K} p_k(s_k \mid Z_t). \tag{2.19}$$

Note that $P^0_{Z_t Z'} = 0$ if state Z' is not a successor of state Z_t, i.e., if there is no strategy profile for all users $\boldsymbol{s}$ that allows the process to reach state Z' from state Z_t.

Variable $\boldsymbol{P^0}$ is the transition matrix of the *unperturbed Markov process*, i.e., when no mistakes take place.

2.3.4.2 Adaptive Play With Mistakes

When players experiment with their chosen actions, mistakes happen. A mistake in the transition from state Z_t to state Z' occurs when one or more players do not respond with

the best reply action, or the optimal response, when faced with state Z_t. The committed mistakes *perturb* the process away from equilibrium. When the probability of mistakes is kept small, [68] showed that the process converges to one *stochastically stable pure Nash equilibrium*.

In the remainder of this section, we present some concepts relative to both adaptive play with mistakes, and Markov chains that will prove useful for the understanding of the proposed solution to the uncoordinated spectrum access problem in Chapter 4. We start by defining the *mixing time* of a Markov chain.

Definition 13. *The **mixing time** of a Markov chain $\mathcal{Z}$ refers to the time necessary for the Markov chain to reach its stationary distribution π with a certain accuracy δ. Let $T_m(\delta)$ be the mixing time of the Markov chain $\mathcal{Z}$ with an accuracy δ. The mixing time is defined as [69]:*

$$T_m(\delta) = \min\left\{t \text{ such that } ||\phi Z_t - \pi||_1 \leq \delta\right\}, \tag{2.20}$$

where ϕ is the initial distribution of the Markov chain.

The modeling of the adaptive play when miskates take place is given next.

Let $\epsilon\lambda_k$ be the probability with which player k experiments by choosing its played action randomly from $\mathcal{S}_k$. Variable ϵ determines the probability with which players generally experiment. For every player k, let $q_k(s \mid Z_t)$ be the conditional probability that player k chooses strategy $s \in \mathcal{S}_k$ given that player k experiments when the system is in state Z_t. For every player k, $q_k(s \mid Z_t)$ satisfies the following:

1. $\sum\limits_{s\in\mathcal{S}_k} q_k(s \mid Z_t) = 1$, for every state Z_t,
2. $q_k(s \mid Z_t)$ is independent of time t,
3. $q_k(s \mid Z_t) > 0$ for every strategy $s \in \mathcal{S}_k$.

Next, we describe the *perturbed* Markov process, i.e., the Markov process that involves players sometimes committing mistakes. The probability of having exactly J users, such that $1 \leq J \leq K$, experimenting is given by: $\epsilon^J \left(\prod\limits_{j\in\mathcal{J}} \lambda_j\right)\left(\prod\limits_{k\in\mathcal{K}\setminus\mathcal{J}} (1-\epsilon\lambda_k)\right)$. Assuming that J players experiment when the process is in state Z_t, the transition probability to state Z' is given by:

$$Q^J_{Z_tZ'} = \prod_{j\in\mathcal{J}} q_j(s_j \mid Z_t) \prod_{k\in\mathcal{K}\setminus\mathcal{J}} p_k(s_k \mid Z_t). \tag{2.21}$$

In the case where no strategy profile can ensure the transition form state Z_t to state Z', $Q^J_{Z_tZ'} = 0$.

If no player experiments, which happens with a probability of $\prod_{k\in\mathcal{K}}(1-\epsilon\lambda_k)$, the transition probability is given by $P^0_{Z_tZ'}$ from Eq. (2.19). Hence, the transition matrix of the *perturbed Markov process*, with a perturbation ϵ, can be given by:

$$P^\epsilon_{Z_tZ'} = \left(\prod_{k\in\mathcal{K}}(1-\epsilon\lambda_k)\right) P^0_{Z_tZ'} + \sum_{\mathcal{J}\in\mathcal{K},\ \mathcal{J}\neq\emptyset} \epsilon^{|\mathcal{J}|}\left(\prod_{j\in\mathcal{J}}\lambda_j\right)\left(\prod_{k\in\mathcal{K}\setminus\mathcal{J}}(1-\epsilon\lambda_k)\right) Q^{|\mathcal{J}|}_{Z_tZ'}. \tag{2.22}$$

Next, we define a regular perturbed Markov process.

Definition 14. *A perturbed Markov process with a transition matrix* $\boldsymbol{P}^{\epsilon}$ *is called a* ***regular perturbed Markov process*** *if it is ergodic for all sufficiently small perturbations* $\epsilon > 0$*, and if* $\boldsymbol{P}^{\epsilon}$ *approaches* $\boldsymbol{P}^{\mathbf{0}}$ *at an exponentially smooth rate, i.e., if:*

1. $\lim_{\epsilon \to 0^+} P^{\epsilon}_{ZZ'} = P^{0}_{ZZ'},\ \forall Z, Z' \in \mathcal{Z}$,

2. $P^{\epsilon}_{ZZ'} > 0$ *for some* $\epsilon > 0 \Rightarrow \lim_{\epsilon \to 0^+} \frac{P^{\epsilon}_{ZZ'}}{\epsilon^{r(Z \to Z')}} < \infty,\ \forall Z, Z' \in \mathcal{Z}$,

where $r(Z \to Z')$ *is the resistance of the transition* $Z \to Z'$*, i.e., from state* Z *to state* Z'*. The definition of the resistance is given next.*

Definition 15. *The* ***resistance*** *of the transition* $Z \to Z'$*,* $r(Z \to Z')$*, denotes the total number of mistakes involved in this transition if such a transition is possible. If the transition* $Z \to Z'$ *is not possible,* $r(Z \to Z') = \infty$*.*

Next, we define the stochastic potential of a recurrence class. Recall that a recurrence class is a set of recurrent states that intercommunicate with one another.

Let the recurrence classes of the unperturbed Markov process with a transition probability $\boldsymbol{P}^{\mathbf{0}}$ be denoted by $C_1, C_2, \ldots C_M$. For each pair of distinct recurrence classes C_i and C_j, $i \neq j$, define an *ij-path* as a sequence of states $\zeta = (Z_1 \to Z_2 \to \ldots \to Z_m)$, such that $Z_k \in \mathcal{Z},\ \forall k \in \{1, \ldots, m\}$, $Z_1 \in C_i$ and $Z_m \in C_j$. The resistance of this *ij-path* is the sum of the resistances of its edges, i.e.:

$$r(\zeta) = r(Z_1 \to Z_2) + r(Z_2 \to Z_3) + \ldots + r(Z_{m-1} \to Z_m). \tag{2.23}$$

Let $\rho_{ij} = \min r(\zeta)$ be the least resistance over all ij-paths ζ. Since transitioning from one recurrence class to another involves at least one player experimenting, ρ_{ij} is positive.

Let G be a weighted directed graph with M vertices, where each vertex corresponds to one recurrence class. Denote the vertex corresponding to class C_j by j and let the weight of the directed edge $i \to j$ be given by ρ_{ij}.

Definition 16. *A* ***j-tree*** T *is a set of* $M - 1$ *directed edges such that from every vertex in* G *other than vertex* j*, there is a unique directed path in the* ***j-tree*** *to vertex* j*.*

Note that the resistance of the *j-tree* is the sum of the resistances of the $M - 1$ edges composing it.

Definition 17. *The* ***stochastic potential****,* γ_j*, of the recurrence class* C_j *is the minimum resistance over all trees rooted at vertex* j*.*

The concepts presented in this section, along with the framework of multi-armed bandits, presented in the next section, will be useful for the understanding of the solution proposed for the uncoordinated spectrum access problem in Chapter 4 of this thesis.

2.4 Multi-Armed Bandits

Reinforcement learning (RL) is a framework where one or multiple learning agents interact with an environment in order to achieve a certain goal, normally represented by the maximization of a numerical reward. RL involves learning how to choose the best actions,

when interacting with an environment, in a way to achieve the goal of the learning agent [43]. The process of RL is illustrated in Fig. 2.5. After observing the state of the environment, the learning agent chooses an action based on its policy or strategy. Based on the chosen action, the environment reacts by changing its state according to a probability distribution that depends on the chosen action of the agent. Moreover, the environment generates a numerical reward that is fed back to the learning agent. Having observed the new state and the achieved reward, the agent chooses a new action. The agent does not know in advance which actions yield the highest reward. Hence, it must learn which actions are the best by trying them out. The aim of the agent in RL is to maximize the long-term perceived reward.

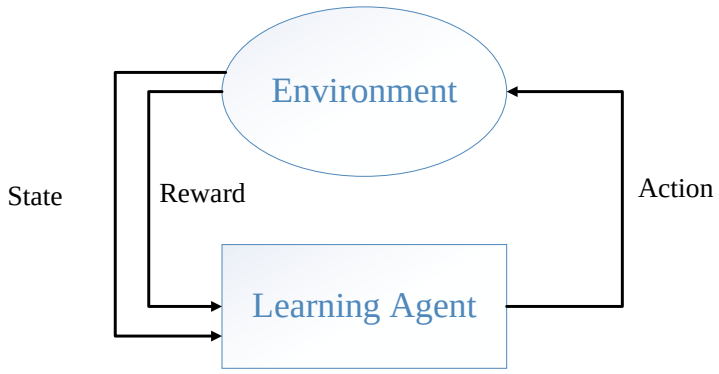


Figure 2.5 – Reinforcement Learning Process

Since the agent does not know *a priori* which actions yield the highest rewards, one of the challenges that arise in RL is the trade-off between exploration and exploitation. To increase the achieved reward, the learning agent must choose more frequently the actions that were found to be effective when tried in the past. However, to discover such actions, the agent must try actions that were not chosen in the past. In other words, the learning agent must exploit the actions that were already tried and that generated satisfactory rewards, and explore un-chosen or rarely chosen actions to make better selections in the future and increase its long-term reward. Hence, the policy followed by the learning agent for choosing actions must strike a trade-off between exploration and exploitation.

The multi-armed bandits (MAB) framework [11] is a particular case of the RL framework corresponding to the single state case. In this setting, the agent aims at finding the action yielding the highest reward in a single-state environment. The MAB problem was extensively studied in the literature. Recently, the multi-player multi-armed bandits (MP-MAB) framework, which involves many learning agents trying to find the actions yielding the highest reward, has gained significant attention in the wireless communication research community. It has been applied to solve many problems, among which we cite:

- opportunistic spectrum access in cognitive radio networks [12, 13, 14],
- uncoordinated spectrum access [15],

- trajectory optimization in UAV-enabled cellular networks [16],
- beam alignment in millimeter wave communications [70].

There exists two main frameworks for the MAB problem that depend on the reward feedback model: stochastic MAB when the reward of an arm is sampled from an unknown probability distribution, and adversarial MAB when the reward of the arm is chosen by an adversary and does not necessarily follow a probability distribution. Wireless communication networks are usually modeled as stochastic environments. Hence, in this thesis, we are interested in the stochastic MP-MAB problem. In this problem, the arm rewards are supposed to be independent and identically distributed (i.i.d.) variables following some unknown distribution, specific to each arm and each user. The stochastic MP-MAB framework consists of:

- a set of K players $\mathcal{K} = \{1, \ldots, K\}$,
- a set of M arms $\mathcal{M} = \{1, \ldots, M\}$,
- the mean reward achieved by player $k \in \mathcal{K}$ when pulling arm $m \in \mathcal{M}$, $\mu(k, m)$,
- the unknown probability distribution of arm $m \in \mathcal{M}$ relative to player k, $\nu(k, m)$,
- the action space of player $k \in \mathcal{K}$, $\mathcal{A}_k$,
- the action space of all players, $\mathcal{A} = \prod_{k \in \mathcal{K}} \mathcal{A}_k$,
- the action of player k at timeslot t, $\boldsymbol{a}_k^{(t)} \in \{0, 1\}^{1 \times M}$, such that $a_k^{(t)}(m) = 1$ if player k chooses arm m and 0 otherwise,
- the strategy profile of all players at timeslot t, $\boldsymbol{a}^{(t)} = \{\boldsymbol{a}_1^{(t)}, \ldots, \boldsymbol{a}_K^{(t)}\}$,
- the collision indicator on channel $m \in \mathcal{M}$ under $\boldsymbol{a}^{(t)}$, $\eta_m(\boldsymbol{a}^{(t)})$ which is equal to 0 if multiple players pull m simultaneously and 1 otherwise.

The players do not a priori know the mean reward associated with each arm. When player k pulls arm m, k receives a reward in return generated according to $\nu(k, m)$ with mean $\mu(k, m)$. In almost all previous work on stochastic MP-MAB in wireless systems, it is assumed that if two or more players pull the same channel, all of the players receive a zero reward. The problem in a MP-MAB setting is hence: how should the players choose arms in order to maximize the accumulated rewards? Here also, the learning agents are faced with the trade-off between exploitation and exploration. In fact, if the players tend to exploit arms that proved effective in the past, they will not gain information on other arms that might provide better rewards. Hence, the players might get stuck in a local optimum. On the other hand, if the players resort to exploring arms all the time, they will gain information about all arms. However, the players will not maximize their accumulated rewards since they will not be exploiting the arms providing the highest mean rewards. Hence, the strategies of the players must strike a trade-off between exploration and exploitation.

To evaluate the performance of the strategy followed by the learning agent, the notion of accumulated regret is introduced. Broadly speaking, the accumulated regret is quantified as the difference between the accumulated reward generated by playing the optimal

arm at each time instance, and the actual achieved reward of the player. The objective of the strategy of the learning player is to refine its strategy in a way to minimize its accumulated regret.

In the stochastic MP-MAB framework, a common assumption is to have the number of arms M be larger than or equal to the number of players K. Hence, in this case, the optimal partition of the arms among the users, without collision, is possible. Let $\boldsymbol{a}^*$ denote the optimal arm selection of all players:

$$\boldsymbol{a}^* = \underset{\boldsymbol{a}\in\mathcal{A}}{\operatorname{argmax}} \sum_{k\in\mathcal{K}} \sum_{m\in\mathcal{M}} a_k(m)\mu(k,m)\eta_m(\boldsymbol{a}). \tag{2.24}$$

In the stochastic MP-MAB framework, the accumulated regret during a time horizon T is given by [11]:

$$R = T \sum_{k\in\mathcal{K}} \sum_{m\in\mathcal{M}} a_k^*(m)\mu(k,m) - \mathbb{E}\left[\sum_{t=1}^{T} \sum_{k\in\mathcal{K}} \sum_{m\in\mathcal{M}} a_k^{(t)}(m)\mu(k,m)\eta_m(\boldsymbol{a}^{(t)})\right]. \tag{2.25}$$

The problem in a stochastic MP-MAB framework is to come up with a strategy for all players to minimize the regret in Eq. (2.25). There exists many strategies in the literature for the stochastic MAB problem, such as the Upper-Confidence-Bound (UCB) [71], Thompson Sampling [72], and ϵ-greedy strategies [73]. Next, we will briefly discuss the UCB algorithm as it is used for comparison in Chapter 4 of this thesis.

UCB Algorithm

The UCB algorithm is based on the principle of *optimism in the face of uncertainty* [11]. This principle uses the rewards observed so far to assign to each arm an upper confidence bound that constitutes, with high probability, an overestimate of the unknown arm mean reward. At time t, let $N_{k,m}^{(t-1)}$ be the number of observations learning agent k has of arm m. Moreover, let $\hat{\mu}(k,m)$ be the estimated mean of arm m by player k. The UCB index assigned by player k to arm m, $UCB^{(t)}(k,m)$, is given by:

$$UCB^{(t)}(k,m) = \begin{cases} \infty & \text{if } N_{k,m}^{(t-1)} = 0, \\ \hat{\mu}(k,m) + \sqrt{\frac{\phi \log(t)}{N_{k,m}^{(t-1)}}}, & \text{otherwise,} \end{cases} \tag{2.26}$$

where ϕ is a parameter that provides a trade-off between exploration and exploitation. Each learning agent adheres to Algorithm 1 to choose its action.

Algorithm 1 UCB Algorithm

Input: ϕ, M

1: **for** $t = 1, \ldots, T$ **do**
2: Choose the arm satisfying $m_k^* = \underset{m\in\mathcal{M}}{\operatorname{argmax}}\ UCB^{(t)}(k,m)$.
3: Observe the achieved reward and update the confidence bound of m_k^*.
4: **end for**

The index of each arm m for each player k, $UCB^{(t)}(k,m)$, is the sum of the empirical mean of rewards experienced so far on m and an exploration bonus, which is also known

as the confidence width. By following Eq. (2.26), a learning agent k explores arm m more often if m is promising because its estimated mean $\hat{\mu}(k,m)$ is large or because it has not been well explored, i.e., if $N_{k,m}^{(t-1)}$ is small. As time progresses, the learning agent tends to exploit the arms having high estimated means. Hence, the UCB algorithm succeeds at striking a trade-off between exploration and exploitation. In the case where the arm means of all players follow the same probability distribution, the UCB algorithm achieves sub-linear regret. However, this sub-linear regret does not scale to the situation of varying arm rewards between users, which is often the case in wireless communication networks. Indeed, for such a case, no theoretical guarantees for the regret exist for the UCB algorithm.

2.5 Matching Theory

In this section, we overview matching theory [17], a powerful mathematical tool that will prove useful in solving the resource allocation problem in Chapter 3 of this manuscript.

Matching theory is a mathematical framework in economics that describes the formation of mutually beneficial relationships. In particular, matching theory is used to solve assignment problems.

In wireless communications literature, matching theory has recently drawn significant attention due to the various useful properties it exhibits. In wireless networks, resource allocation problems consist of assigning network resources (e.g., timeslots, frequency channels, power, serving antennas) to demanding entities (e.g., devices or users). The goal of the resource allocation problem is to optimally allocate resources to users, given a set of network constraints. To solve this problem, several recent works proposed algorithms based on the matching theory framework [18, 19, 20, 21, 22].

As discussed in Section 2.1 of this chapter, future communication networks are expected to incorporate new technologies such as DASs and SBSs to improve their performance. To be able to support these new technologies, novel resource allocation techniques must be adopted. First, these novel resource allocation techniques must support a distributed implementation. In fact, existing cellular networks rely on macro base stations (MBSs) to exchange resource allocation information among one another and allocate network resources to their assigned users while limiting inter-cell interference. However, such a conventional centralized resource management will not effectively work in future wireless networks, due to their distributed nature, resulting from the dense deployment of DASs, SBSs and UAVs. In fact, achieving full coordination among these distributed entities requires a significant backhaul infrastructure, leading to substantial signaling costs. Second, with the densification of cellular networks, the cell association of users becomes more complex as the choices available for each user increase with the number of distributed antennas and SBSs, greatly increasing system complexity.

Therefore, novel resource allocation techniques should benefit from a distributed implementation, have a low-complexity and exhibit fast-convergence behavior. New optimization techniques and game-theoretic solutions were recently introduced in the literature to account for the new challenges of wireless communication networks. That said, resource allocation solutions based on mathematical optimization techniques are mostly suitable for centralized networks and require full information of the whole network. Applying them to distributed systems greatly increases complexity and signaling overhead.

Recently, matching theory, which benefits from a distributed, low-complexity implementation, emerged as a promising approach to solve the resource allocation problems in future communication networks. Next, the key components and terminologies of a matching problem in the context of wireless resource allocation are defined:

- **Two disjoint sets of players:** A matching problem Ψ can be visualized using a bipartite graph where the resources and the users, known as players in game theory terminology, form two disjoint sets. The matching problem Ψ consists of assigning resources from one set ($r \in \mathcal{R}$) to the players of the other set (users $k \in \mathcal{K}$). Let $\Psi(k) \in \mathcal{R}$ be the set of resources allocated to user k. Similarly, let $\Psi(r) \in \mathcal{K}$ be the set of users assigned to resource r. Examples of resources in wireless communication networks include timeslots, frequency channels and serving antennas.

- **Quota:** The quota of each player p determines the maximum number of players from the other set, with which it can be matched. Let q_k and q_r be the quota of user k and of resource r, respectively. A matching outcome Ψ must ensure: $|\Psi(k)| \leq q_k$ and $|\Psi(r)| \leq q_r$.

- **Preference relation and strategy:** A player uses a preference relation $\succ$ to rank players from the other set based on some metric that the player seeks to maximize. The *strategy* of the player consists of this ranking. Player p quantifies the preference relation by assigning a utility U to each player from the other set. Let $\succ_k$ and $U_K(k, r)$ denote the preference relation of user k and the utility assigned by user k to resource r respectively. User k prefers resource r_1 over resource r_2, i.e., $r_1 \succ_k r_2$, if $U_K(k, r_1) \geq U_K(k, r_2)$. Similarly, let $\succ_r$ and $U_R(r, k)$ denote the preference relation of resource r and the utility assigned by resource r to user k respectively. Resource r prefers user k_1 over user k_2, i.e., $k_1 \succ_r k_2$, if $U_R(r, k_1) \geq U_R(r, k_2)$.

- **Utility function:** Each player assigns a utility function to the players from the other set. The utility function is the objective function of user k. By maximizing its utility, each user maximizes its objective function.

- **Solution of the matching problem:** The solution of the matching problem is a function $\Psi : \mathcal{K} \to \mathcal{R}$ such that:

  1. $\Psi(k) \subseteq \mathcal{R}, \forall k \in \mathcal{K}$,
  2. $\Psi(r) \subseteq \mathcal{K}, \forall r \in \mathcal{R}$
  3. $|\Psi(k)| \leq q_k, \forall k \in \mathcal{K}$
  4. $|\Psi(r)| \leq q_r, \forall r \in \mathcal{R}$
  5. $r \in \Psi(k) \Leftrightarrow k \in \Psi(r)$

  If user k and resource r are matched together under Ψ, we say that $(k, r) \in \Psi$ forms a matching pair. If user k and resource r are not matched together under Ψ, then $(k, r) \notin \Psi$.

- **Two-sided stable matching:** Consider there exists a pair $(k, r) \notin \Psi$, such that $k \succ_r k'$, where $k' \in \Psi(r)$, and $r \succ_k r'$, where $r' \in \Psi(k)$. In this case, (k, r) forms a *blocking pair* in the sense that they can leave their current assigned players, and create a new matched pair $(k, r) \in \Psi$. A matching Ψ is two-sided *stable* if there does not exist any pair that can block the matching.

Classifying a matching problem is commonly based on the quota of the players. If $q_k = q_r = 1, \forall k \in \mathcal{K}, \forall r \in \mathcal{R}$, Ψ is referred to as a one-to-one matching. If $q_k = 1, \forall k \in \mathcal{K}$ and $q_r > 1,\ \forall r \in \mathcal{R}$ or $q_k > 1, \forall k \in \mathcal{K}$ and $q_r = 1,\ \forall r \in \mathcal{R}$, Ψ is referred to as a many-to-one matching. Moreover, if $q_k > 1$ and $q_r > 1, \forall k \in \mathcal{K}, \forall r \in \mathcal{R}$, Ψ is called to many-to-many matching. For one-to-one and many-to-one matching problems, the *deferred acceptance (DA) algorithm*, proposed in [17], is always guaranteed to converge to a two-sided stable matching.

A new wireless-oriented method for classifying problems is presented in [74]. This method partitions matching problems into three classes:

1. **Canonical matching:** This encompasses the basic form of matching problems where the preferences of the players do not change during the resource allocation problem. An example of such a matching problem is the allocation of orthogonal frequency bands.

2. **Matching with *externalities***: This category includes scenarios in which player preferences are interdependent, leading players to change their preferences and strategies within the timeframe of the resource allocation. Examples of matching problems with *externalities* include the allocation of non-orthogonal frequency bands, cell association and load balancing.

3. **Matching with dynamics**: This third class represents matching problems in which the strategy of players in the current resource allocation may depend on the strategies in the past resource allocations. Resource allocation with changing environmental conditions is an example of such a matching.

2.6 Summary

As discussed in Chapter 1 and earlier in this chapter, future wireless communication networks must evolve by adopting new technologies to fulfill the requirements of emerging applications. Among these new technologies, future communication networks must leverage new radio access techniques such as NOMA, distributed systems with self-organizing capabilities and UAV-aided communications. In this thesis, we study the integration of these technologies in future wireless communication networks. In particular,

- In Chapter 3, we adopt the matching theory framework to jointly solve the antenna association and the channel allocation problems in a DAS characterized by heterogeneous traffic requirements and employing NOMA signaling. We show that, because of the co-channel interference resulting from NOMA, this problem belongs to the class of matching problems with *externalities*. An algorithm that solves the antenna association and channel allocation problems, and accounts for the user preferences' *externalities*, is then proposed.

- In Chapter 4, we adopt the MP-MAB framework to solve the uncoordinated spectrum access and power control problems in wireless communication networks. A game-theoretic solution, based on Markov chains and achieving a sub-linear regret, is proposed.

- In Chapter 5, we study a UAV-aided communication system, where a UAV is deployed to serve users in an area not covered by terrestrial BSs. The backhaul link between the UAV and an MBS, serving as gateway to the core network, is given particular attention. A solution that minimizes the transmit power of the UAV while satisfying the QoS requirements of the deployed users as well as the backhaul capacity is proposed.

Chapter 3

Resource Allocation for Mixed Traffic

3.1 Introduction

As mentioned in Chapter 1, the explosive growth in connectivity and information sharing brought by the proliferation of Internet of Things (IoT) applications has been paving the way towards future and more evolved generations of cellular networks. In addition to increasing the achieved data rates, future communication systems are expected to accommodate a massive number of connected devices deployed to enable different applications. These span various sectors (e.g., autonomous vehicles, automated control, e-health, virtual reality) and should co-exist with traditional applications (e.g., file download, web browsing). However, the new envisioned applications have very different requirements, compared to traditional services, in terms of data rate, latency and reliability. As a result, mobile traffic is evolving into a more heterogeneous or mixed model requiring the adoption of new technologies by fifth generation (5G) and beyond communication systems to cope with these diverse requirements.

Non-orthogonal multiple access (NOMA) has recently emerged as a promising multiple access technology for future communication systems [28, 75, 76]. By allowing multiple users to access the same resource, NOMA enhances spectral efficiency and increases the number of admitted users which is necessary to achieve massive connectivity, rendering NOMA a promising solution to support mixed traffic systems. In fact, in a mixed traffic system where some users have rate requirements and others aim to maximize theirs, NOMA enables the sharing of one subband between two users of the two categories. That way, if the rate requirement of a user is low, the system spectral efficiency is not penalized as in an orthogonal multiple access (OMA) system, since another user can benefit from the same subband. Moreover, NOMA enables the spectrum to be overloaded which ensures the accommodation and the satisfaction of a higher number of users when compared to OMA scheduling.

In addition to NOMA, distributed antenna systems (DASs) [23, 37, 77, 78], and their evolution to cloud radio access networks (C-RAN) [38, 39, 79, 80] were recently introduced as promising network architectures. By using multiple remote radio heads (RRHs) coordinated by a central controller, DASs enable higher capacities and increased coverage.

In this chapter, we study a mixed traffic system consisting of real-time (RT) users running latency-constrained applications, in addition to best-effort (BE) users aiming to

maximize their achieved rates while optimizing system fairness. We propose the use of NOMA signaling in a distributed antenna system to serve users. In particular, we formulate a subband and antenna assignment problem for the coexistence of multiple traffic types. The proposed resource allocation problem aims at maximizing both the satisfaction level of RT users and the performance of BE users. The work presented in this chapter has appeared in the Proceedings of the IEEE Vehicular Technology Conference Fall 2018 [81] and an extended journal version has been published in the IEEE Transactions on Communications in January 2020 [82]. Next, we give an overview of the related literature before presenting our problem statement and contributions.

3.1.1 Related Work

Resource allocation for mixed traffic types was previously investigated in the literature, mostly in an OMA system. In [83], the authors adopted utility theory for a system consisting of RT and BE users, and proposed a heuristic algorithm to solve the resource allocation problem, based on Lagrange multipliers. In [84] and [85], after partitioning users among different classes based on their requirements, the priority of each user was calculated using fuzzy logic before scheduling the most urgent ones. In [86], the authors proposed a heuristic to perform quality of service (QoS)-based scheduling for small-cell users. They also developed an admission control algorithm to enhance the scheduling policy. Network coordination was employed to enhance the performance of RT users in [87] and minimize the amount of resources needed by RT users, thereby increasing their availability for BE users. This minimization was also the target of [88] where a scalable transmission time interval (TTI) was adapted to the data and latency requirements of the users.

Resource allocation for NOMA systems has been extensively studied with different performance measures. To name a few, the weighted sum rate of a NOMA system was maximized in [89]; however the proposed method has an exponential complexity with both the number of users and the number of available subbands. Maximizing system fairness was the target of [90], while minimizing the used power subject to rate requirements was the target of the works in [91] and [92]. Considering a millimeter wave system with mixed traffic, [93] proposed an algorithm for user grouping, then determined the optimal power allocation to maximize the spectral efficiency of BE users while serving RT users with their rate requirements. However, [93] restricted RT users to be scheduled as second users in NOMA, i.e., users not performing successive interference cancellation (SIC), which decreases the probability of satisfying their needs.

Regarding the combination of NOMA and DASs, in [39] and [94], the authors used NOMA in the transmission from the central controller to various RRHs. They proposed a power allocation scheme between the RRHs as well as an algorithm that finds the optimal number of base stations (BSs) in order to guarantee the cell-edge user requirement in [39] or maximize the energy efficiency (EE) in [94]. The work in [80] considered an uplink setting, where the RRHs cooperate to remove the interference brought by NOMA. In [38], the outage probability of a downlink two-user cloud radio access networks (C-RAN) system was derived using stochastic geometry. In [95], several joint subcarrier, RRH and power allocation techniques were proposed to reduce the total transmit power in each cell, using proper combinations of NOMA with DASs. However, none of these works investigated the use of NOMA and DASs to accommodate mixed traffic systems.

Matching theory-based algorithms for resource allocation have recently gained significant attention. In [18], the authors considered a hybrid C-RAN system with device-to-device (D2D) communications and adopted matching theory to perform the subband-RRH assignment. However, to simplify the problem, they supposed that the user-RRH association is done beforehand and restricted each user to be assigned to one antenna and access one subband only. Similarly, in the context of NOMA, [19] developed a user pairing technique based on matching theory. In [20] and [21], an algorithm based on matching theory was proposed to perform subband allocation for users and D2D pairs respectively. However, to the best of our knowledge, no previous work considered the application of matching theory to solve the subband assignment problem for mixed traffic in a NOMA-DAS. More specifically, none of the previously proposed matching theory-based algorithms for NOMA systems ensured that rate requirements are met. Furthermore, in the DAS settings, the restrictions made by previous algorithms on antenna and subband assignment are unrealistic in practice, and are introduced only to simplify the resource allocation problem. Indeed, as already mentioned, the work in [18] assumed that the assignment of users to the distributed antennas is done beforehand. In addition, almost all previous studies in this context restricted each user to be assigned to one antenna and to access one subband only, while others [37] considered that the spectrum consists of one subband to bypass the subband assignment step.

3.1.2 Problem Statement and Contributions

The main objective of this chapter is to propose an efficient and practical solution for the resource allocation problem. An efficient solution for the formulated problem aims at satisfying the maximum number of RT users, while still providing BE users with an acceptable service. A practical solution means that the proposed technique is readily applicable in wireless systems without incurring a prohibitive complexity. That said, to achieve the full-potential of NOMA in the DAS settings, RRH and sub-channel assignment, as well as power allocation must be optimized jointly. However, this results in a mixed-integer optimization problem which is NP-hard and for which the optimal solution is found by exhaustive search. The latter has a prohibitive complexity for practical systems; hence, suboptimal but more efficient resource allocation techniques are preferred in practice.

In this chapter, we first propose a low-complexity greedy algorithm to perform resource allocation for a mixed traffic system consisting of RT and BE users. We then study the resource allocation problem with a focus on antenna and subband assignment under different system configurations. In particular, different system settings combining DASs or centralized antenna systems (CASs) with different signaling technologies, namely OMA and NOMA, are studied to compare their performance in the mixed traffic context. To tackle the assignment problem in an efficient manner, we then cast the antenna and subband assignment problem as a matching game.

Contrary to previous works, the matching theory-based solution proposed in Section 3.4 of this chapter is the first to provide a joint solution for the assignment problem at hand while allowing each user to access any number of RRHs and subbands simultaneously. This solution takes into account the NOMA power multiplexing constraints for users scheduled on each subband, mostly neglected in previous studies. Moreover, for each system setting, i.e., for OMA-CAS, OMA-DAS, NOMA-CAS and NOMA-DAS, the resource allocation

problem is cast as a one-to-many matching game and a technique based on the deferred acceptance (DA) algorithm is introduced. To the best of our knowledge, no previous study has considered the use of matching theory to resolve the mixed traffic resource allocation problem, combining DAS and NOMA. It should also be noted that none of the previous works applying matching theory to solve the resource allocation problem for a NOMA system incorporated rate requirements into their analysis, while the current work does. For the NOMA-CAS and NOMA-DAS settings, a hybrid NOMA system is devised using matching theory, where subbands are either allocated to single users or user-pairs in such a way to optimize system performance. Moreover, an algorithm that overcomes the need for a swapping phase to deal with the interdependencies between users' preferences is introduced. For the DAS setting, an approach to optimize the number of subbands per antenna is also introduced. We prove that the proposed algorithm, for each system setting, converges to a stable solution, within a limited number of iterations. Moreover, simulation results show that both proposed solutions, i.e., the greedy algorithm and the matching-based solution, greatly outperform conventional methods, especially in terms of maximizing the satisfaction level of RT users.

The rest of this chapter is organized as follows. In Section 3.2, the system model is described. Section 3.3 presents the low-complexity greedy solution, while Section 3.4 is dedicated to introducing the matching-based solution. Conclusions are drawn in Section 3.5.

3.2 System Model and Problem Formulation

3.2.1 System Description

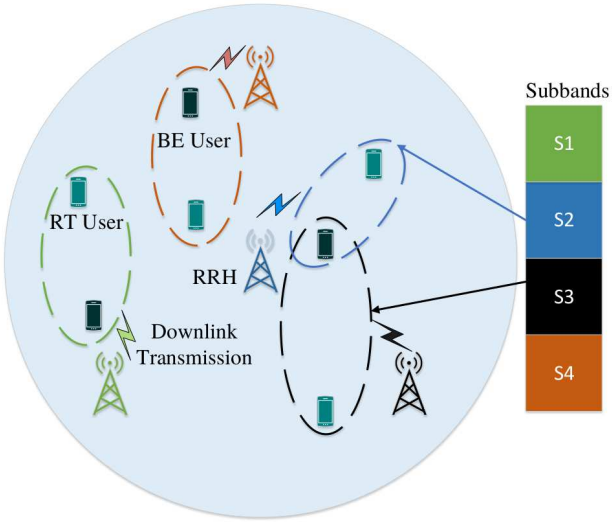


Figure 3.1 – System Model

Consider a downlink system as shown in Fig. 3.1 with K single-antenna users uniformly deployed over a cell. The total system bandwidth B is divided into S subbands, leading to a bandwidth of $B_c = B/S$ per subband. In this chapter, different system configurations are studied. More precisely, we consider both CASs and DASs settings: CASs

consists of one antenna located at the cell center, with a power budget P_{CAS}, whereas in DAS, A single-antenna RRHs are uniformly deployed over the cell. Each antenna has a total power budget of P_{DAS}. For multiple access, OMA as well as NOMA, which enables up to N_s users to be non-orthogonally multiplexed over a subband s, are considered. Hence, the different studied system configurations are OMA-CAS, OMA-DAS, NOMA-CAS and NOMA-DAS. The sets of users, subbands and RRHs will be respectively denoted by $\mathcal{K}, \mathcal{S}$ and $\mathcal{A}$.

When DAS is considered, a subband can only be assigned to one antenna during a scheduling slot to limit intra-cell interference. When NOMA is adopted for multiple access, the messages of up to N_s users are superposed and transmitted over subband s. This results in co-channel interference between the collocated users. Therefore, user k applies SIC [8] before demodulating its own signal, resulting in the following achieved rate:

$$R^t_{k,s,a} = B_c \log_2 \left(1 + \frac{p^t_{k,s,a} \left(h^t_{k,s,a}\right)^2}{\sum\limits_{k' \in \mathcal{I}^t_{k,s,a}} p^t_{k',s,a} \left(h^t_{k,s,a}\right)^2 + N_0 B_c} \right). \tag{3.1}$$

In Eq. (3.1), $p^t_{k,s,a}$ and $h^t_{k,s,a}$ are respectively the transmit power and the channel gain of user k over subband s when assigned to antenna a at timeslot t. The noise power spectral density is denoted by N_0. The first term in the denominator reflects the co-channel interference resulting from NOMA and experienced by user k, scheduled on subband s when the latter is assigned to antenna a. Denote by $\mathcal{S}_s$ the set of users scheduled over subband s. The co-channel interference experienced at the level of user k is caused by users in $I^t_{k,s,a} = \left\{ (k' \in \mathcal{S}_s \setminus \{k\}) \cap \left(h^t_{k',s,a} > h^t_{k,s,a} \right) \right\}$, i.e., users scheduled on subband s and having a higher channel gain than k on s.

SIC results in a significant complexity increase at the receiver side, compared to OMA signaling; therefore, in this study, the maximum value of N_s is restricted to 2, for every subband $s \in \mathcal{S}$.

3.2.2 User Characteristics

In this work, we differentiate between two user classes characterized by different requirements: BE users and RT users.

3.2.2.1 BE Users

This category includes users running delay-tolerant, rate-demanding applications such as file download or web browsing. The goal of these users is to maximize both the achieved data rates and the system fairness. The performance measure for BE users is therefore chosen to be:

$$M^t_{BE}(\boldsymbol{x^t}, \boldsymbol{p^t}) = \sum_{k_{BE}=1}^{K_{BE}} \sum_{a=1}^{A} \sum_{s=1}^{S} x^t_{k_{BE},s,a} \, R^t_{k_{BE},s,a} \, f\left(T^t_{k_{BE}}\right), \tag{3.2}$$

where $x^t_{k_{BE},s,a} = 1$ if user k_{BE} is scheduled on subband s when the latter is assigned to antenna a, and 0 otherwise. Variable $\boldsymbol{p^t}$ denotes the power allocation vector and f is a measure of system fairness that depends on the average data rate $T^t_{k_{BE}}$ of each BE user until the beginning of timeslot t. The average data rate of user k_{BE}, $T^t_{k_{BE}}$, is updated at

the beginning of each timeslot according to:

$$T_{k_{BE}}^{t} = \left(1 - \frac{1}{t_c}\right) T_{k_{BE}}^{t-1} + \frac{1}{t_c} R_{k_{BE}}^{t-1}. \tag{3.3}$$

In Eq. (3.3), t_c is the averaging window and $R_{k_{BE}}^{t-1}$ is the total rate achieved by user k_{BE} at timeslot $(t-1)$, i.e.,

$$R_{k_{BE}}^{t-1} = \sum_{a=1}^{A} \sum_{s=1}^{S} R_{k_{BE},s,a}^{t-1}. \tag{3.4}$$

The expression of Eq. (3.2) is a generic form that can enclose a wide range of specific performance metrics. A common trait of these metrics is the combination of the achieved rate and the system fairness in the scheduling decision. If the expression of Eq. (3.2) did not include the achieved throughput term, the scheduler would optimize system fairness, while penalizing the achieved sum rate. In contrast, if this expression did not include the fairness measure, the system throughput would be maximized by only scheduling users with a high channel gain, hence achieving a low system fairness. Therefore, to optimize performance, both the achieved throughput and the fairness measure need to be accounted for in the expression of Eq. (3.2). The maximum of Eq. (3.2) for BE users is reached when the product between their achieved rates and the fairness between them is maximized. Hence, by adopting this measure, a tradeoff between the maximization of the achieved rates and that of user fairness is reached. This tradeoff can be efficiently reached by the well known proportional fairness (PF) scheduler [28], known to achieve the best balance between rate and system fairness. In fact, the PF metric is one of the expressions embodied by Eq. (3.2), and will be adopted later on in the proposed solutions.

3.2.2.2 RT Users

This category includes users running latency-constrained applications. While some applications require only a small rate (e.g., autonomous cars or sensor applications), others are more bandwidth-hungry (e.g., virtual reality). To capture these different requirements, we associate RT users with a strict latency limit $L_{k_{RT}}$ (expressed as an integer number of timeslots) as well as with a requested amount of data bits $D_{k_{RT}}^{\text{req}}$. The satisfaction of RT users depends upon receiving the required amount of data bits, $D_{k_{RT}}^{\text{req}}$, prior to their latency limit. In this work, we propose to take advantage of the fact that the satisfaction of RT users does not depend on the specific timeslots in which they receive the requested data, as long as these slots precede $L_{k_{RT}}$, for every user $k_{RT} \in \mathcal{K}_{RT}$. To illustrate this, suppose RT user k_{RT} requests 10^5 bits within 20 ms, i.e., $L_{k_{RT}} = 20$ with a timeslot duration τ equal to 1 ms. User k_{RT} is considered to be equally satisfied if it receives all the bits by timeslot t, no matter what t is, as long as $t \leq 20$. Benefiting from this fact, each RT user will be given enough resources to reach $D_{k_{RT}}^{\text{req}}$ by the time $t = L_{k_{RT}}$. The total number of requested data bits $D_{k_{RT}}^{\text{req}}$ is therefore equally divided among the timeslots preceding the latency limit. Hence, from the start of the scheduling period till the end of timeslot t, each RT user k_{RT} needs to be allocated a number of bits equal to:

$$D_{k_{RT}}^{\text{req},t} = t\, D_{k_{RT}}^{\text{req}} \,/\, L_{k_{RT}}. \tag{3.5}$$

The number of required bits until timeslot t, $D_{k_{RT}}^{\text{req},t}$, is an increasing function of both the current timeslot index t and $D_{k_{RT}}^{\text{req}}$. Moreover, $D_{k_{RT}}^{\text{req},t}$ is decreasing in $L_{k_{RT}}$. Hence, its

value increases when the latency limit is small, the total required number of bits is large and when the current timeslot index approaches the latency limit. With this definition of $D_{k_{RT}}^{\text{req},t}$, the required rate at timeslot t can be written as:

$$R_{k_{RT}}^{\text{req},t} = \frac{D_{k_{RT}}^{\text{req},t} - D_{k_{RT}}^{\text{ach},t-1}}{\tau}, \tag{3.6}$$

where $D_{k_{RT}}^{\text{ach},t-1}$ denotes the received number of bits by k_{RT} until the end of the previous timeslot $(t-1)$.

Let $\mathbb{I}_{k_{RT}}^{t}(\boldsymbol{x^t}, \boldsymbol{p^t})$ be a measure of the satisfaction of k_{RT} defined by:

$$\mathbb{I}_{k_{RT}}^{t}(\boldsymbol{x^t}, \boldsymbol{p^t}) = \begin{cases} 1 & \text{if } \sum_{a=1}^{A} \sum_{s=1}^{S} x_{k_{RT},s,a}^{t} \, R_{k_{RT},s,a}^{t} \geq R_{k_{RT}}^{\text{req},t}, \\ 0 & \text{otherwise.} \end{cases} \tag{3.7}$$

In Eq. (3.7), $\mathbb{I}_{k_{RT}}^{t}(\boldsymbol{x^t}, \boldsymbol{p^t}) = 1$, and hence user k_{RT} is satisfied if the current resource and power allocation scheme allows user k_{RT} to achieve a sum rate that is at least equal to its requested rate $R_{k_{RT}}^{\text{req},t}$, which is calculated to allow user k_{RT} to reach its requested number of data bits before its latency limit. In the opposite case, user k_{RT} is not satisfied with the current allocation which reflects in having $\mathbb{I}_{k_{RT}}^{t}(\boldsymbol{x^t}, \boldsymbol{p^t}) = 0$.

Having Eq. (3.7) at hand, we propose to formulate the optimization function for all RT users as:

$$M_{RT}^{t}(\boldsymbol{x^t}, \boldsymbol{p^t}) = \sum_{k=1}^{K_{RT}} \mathbb{I}_{k_{RT}}^{t}(\boldsymbol{x^t}, \boldsymbol{p^t}). \tag{3.8}$$

Using the above formulation, Eq. (3.8) measures the number of RT users having received their requested data rate at each timeslot t. Hence, at the end of the latency period of all RT users, Eq. (3.8) determines the number of satisfied RT users, i.e., having received the totality of requested data bits.

For concision purposes, the timeslot index t will be dropped in the rest of this chapter when there is no confusion. Table 3.1 contains the main notations used throughout this chapter.

3.2.3 Optimization Problem

Having defined the performance measures to be maximized for both user types, the following optimization problem must be solved at each timeslot:

$$\max_{\boldsymbol{x},\boldsymbol{p}} \quad (3.2) \tag{3.9}$$

$$\begin{aligned}
\text{such that} \quad & (3.8), && \text{(3.9a)} \\
& \sum_{a \in \mathcal{A}} x_{k,s,a} \leq 1, \ \forall (k,s) \in \mathcal{K} \times \mathcal{S}, && \text{(3.9b)} \\
& \sum_{k \in \mathcal{K}} x_{k,s,a} \leq 2, \ \forall (s,a) \in \mathcal{S} \times \mathcal{A}, && \text{(3.9c)} \\
& \sum_{k \in \mathcal{K}} \sum_{s \in \mathcal{S}} P_{s,a} x_{k,s,a} \leq P, \ \forall a \in \mathcal{A}, && \text{(3.9d)} \\
& p_{k_1,(s,a)} < p_{k_2,(s,a)} \ \forall (s,a), \in \mathcal{S} \times \mathcal{A}, && \text{(3.9e)} \\
& x_{k,s,a} \in \{0,1\}. && \text{(3.9f)}
\end{aligned}$$

Table 3.1 – Notation Table

K	Total number of users	S	Number of subbands
K_{RT}	Number of RT users	K_{BE}	Number of BE users
$\mathcal{S}_s$	Set of users scheduled on subband s	$\mathcal{S}_k$	Set of subbands assigned to user k
A	Number of distributed antennas	$\boldsymbol{N} = \{N_a\}$	Number of subbands assigned to each antenna
P_{CAS}, P_{DAS}	Power per antenna in the CAS and DAS settings resp.	P_a	Power per subband on antenna a
t	Timeslot index	$\mathcal{K}_{\text{active}}$	Set of active users
$\boldsymbol{p}^t = \{p^t_{k,s,a}\}$	Power allocated to user k on subband s and antenna a at timeslot t	$\boldsymbol{h} = \{h^t_{k,s,a}\}$	Channel gain of user k on subband s and antenna a at timeslot t
$\boldsymbol{R}^t = \{R^t_{k,s,a}\}$	Rate achieved by user k on subband s and antenna a at timeslot t	$\boldsymbol{T} = \{T^t_k\}$	Average rate achieved by BE user k before reaching timeslot t
M^t_{RT}, M^t_{BE}	Satisfaction measure of RT and BE users resp. at timeslot t	$\mathcal{U}^t_{RT}(.,s)$, $\mathcal{U}^t_{BE}(.,s)$	Utility of RT and BE users resp. on subband s at timeslot t
$\boldsymbol{L} = \{L_k\}$, $\boldsymbol{D}^{\text{req}} = \{D^{\text{req}}_k\}$	Latency limit and number of requested data bits of each user k resp.	$\boldsymbol{R}^{\text{req},t} = \{R^{\text{req},t}_k\}$, $\boldsymbol{D}^{\text{req},t} = \{D^{\text{req},t}_k\}$ $\boldsymbol{D}^{\text{ach},t} = \{D^{\text{ach},t}_k\}$	Required rate, required number of bits, and achieved number of bits of each user k at timeslot t resp.
Ψ	Matching outcome	$\mathcal{PL}$	Preference list
$\mathcal{M}(s)$, $\mathcal{AS}(s)$	Matching set and applying user set of subband s resp.	$v^{k,a}_p$	Proposing virtual user relative to user k on antenna a

Constraint (3.9b) restricts each subband to be assigned to one antenna only in the DAS case, while (3.9c) limits the maximum number of users per subband to 2. Constraint (3.9d) is the power budget per antenna, where $P = P_{CAS}$ (resp. $P = P_{DAS}$) in the CAS setting (resp. DAS setting). Denoting by k_1 and k_2 the users scheduled on (s, a) such that $h_{k_1,s,a} > h_{k_2,s,a}$, user k_2 must be allocated a higher power value than user k_1 as expressed in constraint (3.9e) to guarantee SIC stability [49, 81], i.e., successful decoding at the user side. Indeed, as shown in [30], the power of the weak user must be strictly greater than that of the strong user. In the opposite case, the outage probabilities of the users will be always one. Note that the optimization problem in (3.9) is formulated for the general case of a NOMA-DAS system, the other system configurations (i.e., OMA-CAS, OMA-DAS,

and NOMA-CAS) being special cases of it.

The optimization problem in (3.9) aims at optimizing the performance of BE users while satisfying the requirements of RT users as expressed in constraint (3.9a). Since the applications of RT users are time sensitive and can be considered as "more urgent" than BE applications, satisfying RT users is given a higher priority in the proposed solutions. Moreover, problem (3.9) is a mixed-integer optimization problem for which an optimal solution is computationally intractable. If equal power repartition between subbands assigned to the same antenna is assumed, solving problem (3.9) consists of finding the optimal subband and antenna assignment. To find this optimal assignment, in Section 3.3, we propose a low-complexity greedy algorithm that aims at both maximizing the percentage of satisfied RT users and maximizing the achieved rates and fairness of BE users for a DAS system with NOMA pairing. The proposed algorithm in Section 3.3 can be easily extended to other system settings. To further optimize system performance, in Section 3.4, we invoke the two-sided matching theory framework to obtain a suboptimal solution for the formulated problem.

3.3 Low-Complexity Greedy Algorithm

As the system consists of users with heterogeneous traffic, the allocation technique must take into account the difference in priority between users. Therefore, whenever RT users are to be scheduled in the current timeslot, these users are first assigned to subbands and antennas in an OMA manner. Once the OMA phase is done, we proceed with the NOMA user pairing phase. When no more RT users need scheduling, BE users are scheduled following the PF principle. In the following, the different steps of the allocation procedure are described.

3.3.1 Phase 1: Assignment of Users and Subbands to Antennas

The goal of this first step in the allocation technique is to assign users and subbands to the distributed antennas. First and foremost, this distribution must be performed in such a way to guarantee the requirements of RT users. It should also maximize the utility of BE users, to the extent of the possible.

Assigning subbands to RRHs is not a straightforward task. Indeed, basing the assignment solely on observed channel gain values may lead to an antenna a being assigned a large number of subbands. In that case, the power per subband on antenna a decreases, while a better performance would have been achieved by assigning the subband to a less congested antenna. Therefore, we start by following the proposal in [23] that relies on large-scale fading $w_{k,a}$ between user k and antenna a to estimate N_a, the potential number of subbands assigned to a:

$$N_a = \left\lfloor \frac{S \times \sum_{k=1}^{K} w_{k,a}}{\sum_{a=1}^{A} \sum_{k=1}^{K} w_{k,a}} \right\rfloor, \; a \in \mathcal{A}. \tag{3.10}$$

The floor operation $\lfloor . \rfloor$ in Eq. (3.10) results in $(S - \sum_a N_a)$ unallocated subbands. This number will be provisionally given to the antennas having the smallest N_a. Following an equal power distribution between subbands assigned to the same antenna, an initial

approximation of the potential power per subband on a is given by:

$$P_{s,a} = P_a = \frac{P_{DAS}}{N_a}. \tag{3.11}$$

However, contrary to [23], in our work, N_a is strictly used for initial power approximation and does not dictate the number of assigned subbands to each antenna a in each timeslot.

With these estimations of power per subband on each antenna, we proceed with the resource allocation for RT users giving them the benefit of being allocated subbands with the best perceived channel coefficient. Since RT users may have different data and latency requirements, we introduce a priority function in order to favor those RT users with stringent requirements. Such a function should be increasing with $D_{k_{RT}}^{\text{req},t}$ and decreasing with $L_{k_{RT}}$ and with the number of bits transferred so far. Therefore, we propose the following priority measure:

$$\text{Priority}(k_{RT}, t) = \frac{\left(D_{k_{RT}}^{\text{req},t} - \left(D_{k_{RT}}^{\text{ach},t-1} + d_{k_{RT}}^{t}\right)\right)/D_{k_{RT}}^{\text{req},t}}{(L_{k_{RT}} - t)/L_{k_{RT}}}, \tag{3.12}$$

where $d_{k_{RT}}^{t}$ is the number of bits transferred so far during timeslot t. Note that the total number of transferred bits at the end of the second phase of the allocation technique (NOMA pairing), at timeslot t, is given by:

$$D_{k_{RT}}^{\text{ach},t} = D_{k_{RT}}^{\text{ach},t-1} + d_{k_{RT}}^{t}. \tag{3.13}$$

Algorithm 2 describes the first phase of the allocation process applied while there are RT users in the system. The set of active RT users $\mathcal{K}_{\text{active}}$, consisting of all RT users $k \in \mathcal{K}_{RT}$ with $R_k^{\text{req},t} \geq 0$, is first found. After updating the priorities of all RT users in $\mathcal{K}_{\text{active}}$, the best resource (i.e., best subband and best antenna) for users having the highest priority is found. This decision is based on the criterion $(P_a h_{k_{RT},s,a}^2)$ rather than just $(h_{k_{RT},s,a}^2)$ in order to account for different power levels on different antennas. When multiple users have the same highest priority (i.e., when the set $\mathcal{MP}$ consists of multiple users), the one having the worst best perceived channel coefficient on a particular subband/antenna pair is prioritized and scheduled on its preferred subband. The choice of prioritizing the user having the worst best perceived channel coefficient when $\mathcal{MP}$ consists of multiple users aims at enhancing the fairness among RT users. The scheduled user is denoted by k^*, and the assigned subband/antenna pair to user k^* are denoted by s_{k^*} and a_{k^*} respectively. If assigning a subband to an antenna a_{k^*} leads to a_{k^*} being assigned more than $N_{a_{k^*}}$ subbands, the power per subband on antenna a_{k^*}, $P_{a_{k^*}}$, is updated. The power values, achieved rates and achieved number of bits of all users scheduled on a_{k^*} are also updated.

Steps 4 till 14 of Algorithm 2 are repeated until $\mathcal{K}_{\text{active}} = \emptyset$ or until no free subbands remain. If the former happens, BE users are scheduled on the remaining subbands, using NOMA, in a way that does not penalize the achieved level of satisfaction of RT users as noted by step 19 of Algorithm 2.

BE users are scheduled according to the PF principle as shown between steps 16 and 29 of Algorithm 2. The PF scheduler [28] is known to achieve a tradeoff between total throughput and fairness maximization, by scheduling on each subband s the user (or user set in the case of NOMA) that maximizes a certain metric, labeled the "PF metric". However, the traditional PF scheduler achieves fairness through the consideration of historical

Algorithm 2 Assignment of subbands, single RT users and NOMA BE users to antennas

Input: $\mathcal{K}_{\text{active}}$, $\mathcal{K}_{BE}$, $\mathcal{S}$, $\boldsymbol{h}$, $\boldsymbol{P}$, $\boldsymbol{N}$, $\boldsymbol{D}^{\text{ach},t-1}$, $\boldsymbol{D}^{\text{req},t}$, $\boldsymbol{L}$, $\boldsymbol{T}$.
Output: $\boldsymbol{A}_{RT}$, $\boldsymbol{A}_{BE}$, $\boldsymbol{R}_{RT}$, $\boldsymbol{R}_{BE}$, $\boldsymbol{P}_{RT}$, $\boldsymbol{P}_{BE}$.

Initialization $\boldsymbol{A}_{RT}, \boldsymbol{R}_{RT}, \boldsymbol{P}_{RT} = \{0\}^{K_{RT}\times S\times A}$, $\boldsymbol{A}_{BE}, \boldsymbol{R}_{BE}, \boldsymbol{P}_{BE} = \{0\}^{K_{BE}\times S\times A}$.
$d^t_{k_{RT}} = 0$, $\forall k_{RT} \in \mathcal{K}_{RT}$.
while $\mathcal{K}_{\text{active}} \neq \emptyset$ & $\mathcal{S} \neq \emptyset$ **do** // Assignment of RT users
  Update priorities for users in $\mathcal{K}_{\text{active}}$ and remove users having negative priorities from the active set.
  $\mathcal{MP} \leftarrow$ users having the highest priority.
  while $\mathcal{MP} \neq \emptyset$ **do**
    $(s_k, a_k) = \underset{s,a}{\text{argmax}} \left(P_a h^2_{k,s,a}\right),\ \forall\, k \in \mathcal{MP}$.
    Schedule $k^* = \underset{k\in\mathcal{MP}}{\text{argmin}} \left(P_{a_k} h^2_{k,s_k,a_k}\right)$ on s_{k*} and assign the latter to antenna a_{k^*}.
    Update $d^t_{k^*}$.
    Let $\mathcal{S} = \mathcal{S} \setminus \{s_{k^*}\}$ and $\mathcal{S}_{a_{k^*}} = \mathcal{S}_{a_{k^*}} \cup \{s_{k^*}\}$. //$\mathcal{S}_{a_{k^*}}$ is the set of subbands assigned to antenna a_{k^*}
    if $|\mathcal{S}_{a_{k^*}}| > N_{a_{k^*}}$ **then**
      Update $P_{a_{k^*}}$, and update the achieved rate, the allocated power and the achieved number of bits $d^t_{k_{RT}}$ of all users scheduled on a_{k^*}.
    end if
    $\mathcal{MP} = \mathcal{MP} \setminus \{k^*\}$.
  end while
end while
if $\mathcal{S} \neq \emptyset$ **then** //Assignment of BE users if all RT users are satisfied
  for $i = 1 : |\mathcal{S}|$ **do**
    $s = \mathcal{S}(i)$.
    Remove from $\mathcal{A}$ the antennas which, if assigned another subband, would cause some RT users to become unsatisfied.
    for $j = 1 : |\mathcal{US}|$ **do**
      Divide P_a, $\forall a \in \mathcal{A}$ between users in us_j using FTPA.
      Find PF_{us_j} on $a^*_{us_j} = \underset{a}{\text{argmax}} \sum\limits_{k_{BE}\in us_j} R_{k_{BE},s,a}$ using Eq. (3.14).
    end for
    Schedule $us^* = \underset{us\in\mathcal{US}}{\text{argmax}}\, \text{PF}_{us}$ on subband s.
    if $|\mathcal{S}_{a_{\mathcal{U}^*_i}}| > N_{a_{\mathcal{U}^*_i}}$ **then**
      Update $P_{a_{\mathcal{U}^*_i}}$, and update the achieved rate, the allocated power and the achieved number of bits $d^t_{k_{RT}}$ for all users scheduled on a_{us^*}.
    end if
  end for
end if

rates up to the last complete allocation slot; i.e., it does not take into consideration the rates achieved during the current slot. That is why we adopt the enhanced PF scheduler from [96] and associate each user set us with the following PF metric when considering the allocation of subband s:

$$PF_{us} = \sum_{k_{BE} \in us} \frac{R^t_{k_{BE},s,a^*_s}}{(1-\frac{1}{t_c})T^t_{k_{BE}} + \frac{1}{t_c}\sum_{i=1}^{s-1} R^t_{k_{BE},i,a^*_i}}. \tag{3.14}$$

In Eq. (3.14), antenna a^*_s is chosen so as to maximize the rate of users in us according to:

$$a^*_s = \underset{a \in \mathcal{A}}{\operatorname{argmax}} \sum_{k_{BE} \in us} R^t_{k_{BE},s,a}. \tag{3.15}$$

The second term in the denominator of Eq. (3.14) accounts for the rate achieved by users in us during the current timeslot (if any) before considering s for allocation, promoting better fairness. Note that Eq. (3.14) is computed for all $|\mathcal{US}| = K_{BE} + P(K_{BE}, 2)$ possible user sets, where K_{BE} accounts for OMA signaling, whereas $P(K_{BE}, 2)$ accounts for NOMA signaling. In the case where the user set us consists of one BE user only, i.e., in the case of OMA signaling, both the PF metric and antenna a^*_s of set us can still be computed using Eq. (3.14) and (3.15) respectively, after discarding the sum term.

To find the antenna satisfying Eq. (3.15) for each user set us and over each subband s, the rates of BE users belonging to us over subband s for each potential antenna $a \in \mathcal{A}$ must be found. To do so, the power repartition between BE users belonging to each user set us must be decided upon. In this work, fractional transmit power allocation (FTPA) [97] is used to partition the power allocated to each subband among users in us. Hence, the power allocated to user $k_{BE} \in us$ over subband s when the latter is assigned to antenna a is given by:

$$p_{k_{BE},s,a} = \frac{P_a \left(\frac{h^2_{k_{BE},s,a}}{N_0 B_c}\right)^{-\alpha_{FTPA}}}{\sum\limits_{k' \in \mathcal{K}_{BE}, x_{k',s,a}=1} \left(\frac{h^2_{k',s,a}}{N_0 B_c}\right)^{-\alpha_{FTPA}}}, \tag{3.16}$$

where α_{FTPA} is a decay factor. Eq. (3.16) ensures a higher power for the user with a lower channel gain, guaranteeing SIC stability and hence satisfying constraint (3.9e). After dividing the power between users in each candidate set us, the antenna maximizing the sum rate of users in us is found according to Eq. (3.15), enabling the computation of the PF metric associated with user set us. In the case of a user set us consisting of one BE user only, the total power P_a is allocated to the user in us before finding antenna a^*_s and computing the PF metric for user set us.

Upon computing Eq. (3.14) for all possible user sets, the scheduled user set us^* on subband s is chosen according to:

$$us^* = \underset{us \in \mathcal{US}}{\operatorname{argmax}}\, PF_{us}. \tag{3.17}$$

Remark 1. *If at the end of Algorithm 2, an antenna a is assigned less than N_a subbands, P_a is re-calculated and the power and rate of users scheduled on antenna a are updated.*

3.3.2 Phase 2: NOMA pairing on subbands assigned to RT users

Assuming equal inter-subband power on each antenna may lead to some RT users having more power than needed for achieving their target throughput. Also, some RT users may not be allocated enough resources, either because the system is congested, or because of their bad channel conditions. That is why we perform a NOMA pairing phase in which we assign second users to subbands assigned to RT users, while keeping the subband-antenna assignment unvaried and guaranteeing the required rates to RT users already scheduled.

First, the set of unsatisfied RT users must be determined. The amount of excess power allocated to satisfied RT users must also be found. For this purpose, we start by estimating the amount of required throughput in timeslot t, for each RT user, according to Eq. (3.6). For each user k_{RT} exceeding its required rate, the amount of power needed to reach the required rate of user k_{RT}, $R_{k_{RT}}^{\text{req},t}$, on its assigned set of subbands, $\mathcal{S}_{k_{RT}}$, is re-calculated by solving the following power minimization problem:

$$\min_{\boldsymbol{p}_{k_{RT}}} \sum_{s \in \mathcal{S}_{k_{RT}}} p_{k_{RT},s,a_s} \tag{3.18}$$

$$\text{such that} \sum_{s \in \mathcal{S}_{k_{RT}}} R_{k_{RT},s,a_s} = R_{k_{RT}}^{\text{req},t}, \tag{3.18a}$$

$$0 \leq p_{k_{RT},s,a_s} \leq P_{a_s}. \tag{3.18b}$$

Constraint (3.18b) is imposed to enforce that no subband is allocated a power larger than the one it was initially allocated so as to not exceed the power budget of each antenna.

Solving the above optimization problem leads to the well-known waterfilling solution where the power value allocated to user k_{RT} on subband s when the latter is assigned to antenna a_s, p_{k_{RT},s,a_s}, is given by:

$$p_{k_{RT},s,a_s} = \left[\frac{\lambda_{k_{RT}} B_c}{\log(2)} - \frac{N_0 B_c}{h_{k_{RT},s,a_s}^2} \right]_0^{P_{a_s}}. \tag{3.19}$$

In Eq. (3.19), $\lambda_{k_{RT}}$ is the Lagrange multiplier given by:

$$\lambda_{k_{RT}} = 2^{\frac{1}{S_{k_{RT}}} \left[\frac{R_{k_{RT}}^{\text{req},t}}{B_c} - \sum\limits_{s \in \mathcal{S}_{k_{RT}}} \log_2 \left(\frac{h_{k_{RT},s,a_s}^2}{\log(2) N_0} \right) \right]}. \tag{3.20}$$

The required rate of user k_{RT} on subband $s \in \mathcal{S}_{k_{RT}}$, $R_{k_{RT},s,a_s}^{\text{req,t}}$, is found by replacing Eq. (3.19) into Eq. (3.1). This rate will be guaranteed for all satisfied RT users during the NOMA pairing step.

Since BE users are scheduled directly using NOMA in Algorithm 2 (if scheduled), only subbands assigned to RT users, $\mathcal{S}_{RT}$, are considered in the second phase of the allocation technique. For that purpose, we start by checking if the achieved rate of each scheduled RT user, k_{RT}, in Algorithm 2 over its assigned subband s exceeds the required one. In case the achieved rate of user k_{RT} on subband s does not exceed the required one, subband s is removed from the set of subbands available for NOMA pairing, $\mathcal{S}_{RT}$, and user k_{RT} will remain its sole occupier. In the opposite case, k_{RT} can share subband s and two scenarios are considered for each candidate user $k' \in \mathcal{K}$ for NOMA pairing in order to guarantee $R_{k_{RT},s,a_s}^{\text{req,t}}$.

3.3.2.1 Scenario 1: $h_{k_{RT},s,a_s} > h_{k',s,a_s}$

In this case, user k_{RT} is paired as first user in NOMA on subband s and its required rate is given by:

$$R^{\text{req},t}_{k_{RT},s,a_s} = B_c \log_2 \left(1 + \frac{p^{1,\text{req}}_{k_{RT},s,a_s} h^2_{k_{RT},s,a_s}}{N_0 B_c}\right), \tag{3.21}$$

where the necessary power value $p^{1,\text{req}}_{k_{RT},s,a_s}$ is given by Eq. (3.19). To guarantee SIC stability, the value of $p^{1,\text{req}}_{k_{RT},s,a_s}$ must be less than the power value allocated to each candidate second user, k', multiplexed on subband s. This translates into considering subband s for NOMA pairing, with user k_{RT} as first user, if and only if $p^{1,\text{req}}_{k_{RT},s,a_s} < P_{a_s}/2$. If this condition is verified, the power available for a candidate user k' on subband s is given by:

$$p^{2,\text{av}}_{s,a_s} = P_{a_s} - p^{1,\text{req}}_{k_{RT},s,a_s}. \tag{3.22}$$

In the opposite case, subband s is removed from the set of subbands available for NOMA pairing $\mathcal{S}_{RT}$.

3.3.2.2 Scenario 2: $h_{k_{RT},s,a_s} < h_{k',s,a_s}$

In this case, user k_{RT} is paired as second user on subband s via NOMA with a required rate given by:

$$R^{\text{req},t}_{k_{RT},s,a_s} = B_c \log_2 \left(1 + \frac{p^{2,\text{req}}_{k_{RT},s,a_s} h^2_{k_{RT},s,a_s}}{p^{1,\text{av}}_{s,a_s} h^2_{k_{RT},s,a_s} + N_0 B_c}\right), \tag{3.23}$$

where $p^{1,\text{av}}_{s,a_s}$ is the available power on s for a candidate first user, k':

$$p^{1,\text{av}}_{s,a_s} = P_{a_s} - p^{2,\text{req}}_{k_{RT},s,a_s}. \tag{3.24}$$

After substituting Eq. (3.24) in Eq. (3.23), $p^{2,\text{req}}_{k_{RT},s,a_s}$ can be expressed as:

$$p^{2,\text{req}}_{k_{RT},s,a_s} = \frac{(b-1)\left(P_{a_s} h^2_{k_{RT},s,a_s} + N_0 B_c\right)}{b h^2_{k_{RT},s,a_s}}, \tag{3.25}$$

where $b = 2^{R^{\text{req},t}_{k_{RT},s}/B_c}$. SIC stability is guaranteed if $p^{2,\text{req}}_{k_{RT},s,a_s} > P_{a_s}/2$, which results in ensuring:

$$P_{a_s} h^2_{k_{RT},s,a_s} (0.5b - 1) + (b-1) N_0 B_c > 0. \tag{3.26}$$

Eq. (3.26) is guaranteed if $b \geq 2$. In other words, if being scheduled as a second user, the required rate of user k_{RT} on subband s must satisfy:

$$R^{\text{req},t}_{k_{RT},s,a_s} = \max\left(R^{\text{req},t}_{k_{RT},s,a_s}, B_c\right). \tag{3.27}$$

Note that guaranteeing SIC stability by enforcing Eq. (3.27) does not result in any degradation regarding the rate of user k_{RT}.

In Algorithm 3, we describe the steps followed for NOMA pairing on subbands assigned to RT users. First, the amount of excess power assigned to satisfied RT users is calculated. Moreover, the set of available subbands for NOMA pairing, $\mathcal{S}_{RT}$, is found. The set of unsatisfied RT users, $\mathcal{K}_{\text{active}}$, is also initialized. If the set $\mathcal{K}_{\text{active}}$ is not empty, we start

Algorithm 3 NOMA pairing on subbands assigned to RT users

Input: $\mathcal{K}_{RT}$, $\mathcal{K}_{BE}$, $\mathcal{S}$, $\boldsymbol{h}$, $\boldsymbol{A}_{RT}$, $\boldsymbol{A}_{BE}$, $\boldsymbol{R}_{RT}$, $\boldsymbol{R}_{BE}$, $\boldsymbol{P}_{RT}$, $\boldsymbol{P}_{BE}$, $\boldsymbol{D}^{\text{ach},t-1}$, $\boldsymbol{D}^{\text{req},t}$, $\boldsymbol{L}$, $\boldsymbol{T}$.
Output: $\boldsymbol{A}_{RT}$, $\boldsymbol{A}_{BE}$, $\boldsymbol{R}_{RT}$, $\boldsymbol{R}_{BE}$, $\boldsymbol{P}_{RT}$, P_{BE}.

Solve problem (3.18) for satisfied RT users.
Find the set of subbands S_{RT} on which NOMA pairing can take place.
Add unsatisfied users to $\mathcal{K}_{\text{active}}$.
while $\mathcal{K}_{\text{active}} \neq \emptyset$ & $\mathcal{S}_{RT} \neq \emptyset$ **do**
  $R_k^{\text{lack},t} = R_k^{\text{req},t} - \sum_{s\in\mathcal{S}_k} R_{k,s,a_s},\ \forall\, k \in \mathcal{K}_{\text{active}}$.
  Compute priorities for users in $\mathcal{K}_{\text{active}}$ and remove those having negative priorities.
  $\mathcal{MP} \leftarrow$ Users with highest priorities.
  while $\mathcal{MP} \neq \emptyset$ **do**
    for $i = 1 : |\mathcal{MP}|$ **do**
      $k' = \mathcal{MP}(i)$.
      for $j = 1 : |\mathcal{S}_{RT}|$ **do**
        $s = \mathcal{S}_{RT}(j)$.
        Find k_s, the RT user scheduled on subband s.
        Let a_s be the antenna to which subband s is assigned.
        if $h(k_s, s, a_s) > h(k', s, a_s)$ **then**
          Calculate $R^{\text{temp}}_{k',s,a_s}$ with the power calculated using Eq. (3.22).
        else
          Calculate $R^{\text{temp}}_{k',s,a_s}$ with the power calculated using Eq. (3.24).
        end if
      end for
      $\text{dist}_{k',s} = \left(R^{\text{temp}}_{k',s,a_s} - R^{\text{lack},t}_{k'}\right),\ \forall\, s \in \mathcal{S}_{RT}$.
      if $\exists\, s \in \mathcal{S}_{RT}$ s.t. $\text{dist}_{k',s} > 0$ **then**
        $s^*_{k'} = \underset{s\in\mathcal{S}_{RT} \text{ s.t. } \text{dist}_{k',s}>0}{\text{argmin}} (\text{dist}_{k',s})$.
      else
        $s^*_{k'} = \underset{s\in\mathcal{S}_{RT}}{\text{argmax}} (\text{dist}_{k',s})$.
      end if
    end for
    Schedule user $k'^* = \underset{k'\in\mathcal{MP}}{\text{argmin}} \left(\text{dist}_{k',s^*_{k'}}\right)$ on its preferred subband $s^*_{k'}$.
    $\mathcal{S}_{RT} = \mathcal{S}_{RT} \setminus (s^*_{k'^*})$.
  end while
end while
if $\mathcal{S}_{RT} \neq \emptyset$ **then** //Assignment of BE users
  for $j = 1 : |\mathcal{S}_{RT}|$ **do**
    $s = \mathcal{S}_{RT}(j)$.
    for $k_{BE} = 1 : K_{BE}$ **do**
      Find k_s, the RT user scheduled on subband s.
      if $h(k_s, s, a_s) > h(k_{BE}, s, a_s)$ **then**
        Find the available power to user k_{BE} from Eq. (3.22).
        Calculate the PF metric associated to user k_{BE}, $PF_{k_{BE}}$ using Eq. (3.14).
      else
        Find the available power to user k_{BE} from Eq. (3.24).
        Calculate the PF metric associated to user k_{BE}, $PF_{k_{BE}}$ using Eq. (3.14).
      end if
    end for

42: Schedule $k^*_{BE} = \underset{k_{BE} \in \mathcal{K}_{BE}}{\text{argmax}}\ PF_{k_{BE}}$ on subband s.
43: **end for**
44: **end if**

by scheduling each user in $\mathcal{K}_{\text{active}}$ on the subbands that provide a total rate closest to its requested rate. To do that, the rate each user $k \in \mathcal{K}_{\text{active}}$ lacks in order to reach is total requested rate is computed as per step 3 of Algorithm 3. The priorities of active RT users are also computed using Eq. (3.12), before removing from the active user set those RT users having a negative priority. A negative priority for RT user k indicates that user k received its requested data rate and is hence satisfied with the current allocation. For users having the same highest priority (users in the set $\mathcal{MP}$), the achievable rate for each user $k' \in \mathcal{MP}$, $R^{\text{temp}}_{k',s,a_s}$, on every subband $s \in \mathcal{S}_{RT}$ is found, as shown between steps 9 and 17 of Algorithm 3. Then, for each user k', the distance between the achieved rate, $R^{\text{temp}}_{k',s,a_s}$, and the lacking rate, $R^{\text{lack,t}}_{k'}$, is found, as per step 18. If the achieved rate on subband s $R^{\text{temp}}_{k',s,a_s}$ is larger than the rate user k' lacks, this distance is positive. In the opposite case, $\text{dist}_{k',s}$ is negative. For each user k', a certain subband $s^*_{k'}$ is then retained according to the following dynamics: if there exists at least one subband $s \in \mathcal{S}_{RT}$ for which $\text{dist}_{k',s}$ is positive, i.e., subband s allows user k' to reach its required rate, subband $s^*_{k'}$ ensuring an achieved rate closest to the rate user k' lacks is retained as per step 20. In the opposite case, i.e., if all subbands $s \in \mathcal{S}_{RT}$ do not ensure a positive distance $\text{dist}_{k',s}$, subband $s^*_{k'}$ maximizing this distance, i.e., maximizing the achieved rate, is retained as per step 22.

Having determined the best subband for each user $k' \in \mathcal{MP}$, user k'^* having the smallest distance to its required rate is then prioritized as shown in step 25. The rational of choosing to prioritize the user with the smallest distance to its required rate is justified as follows: when there is at least one user having a negative distance, such a user is prioritized and given its best subband; when all users have positive distances, the metric allows us to satisfy the one for whom the distance is the smallest, i.e. it promotes the efficient use of the available spectrum.

Step 2 to step 28 of Algorithm 3 are repeated until either all RT users are satisfied or no subbands remain for NOMA pairing. In the former case, on the remaining available subbands for NOMA pairing, BE users are paired via NOMA using the PF scheduler to enhance their performance. The BE users pairing technique is shown between steps 29 and 41. Note that at this stage, only one BE user is scheduled over each available subband $s \in \mathcal{S}_{RT}$ via NOMA, as an RT user already occupies subband s. Hence, the PF metric is computed using Eq. (3.14) for each possible BE user, i.e., by dropping the sum term.

3.3.3 Global Resource Allocation Technique

The proposed allocation technique is summarized in Algorithm 4.

3.3.4 Numerical Results

To evaluate the performance of the proposed solution, we consider a single cell having a radius of $R_d = 500$ m with $A = 4$ RRHs. One antenna is located at the cell center while the others are equally distanced and positioned on a circle of radius $2R_d/3$ with an angular separation of 120°. In the DAS setting, each RRH or distributed antenna has a

Algorithm 4 Proposed Allocation Technique

Input: $\mathcal{K}_{RT}$, $\mathcal{K}_{BE}$, $\mathcal{S}$, $\boldsymbol{h}$, $\boldsymbol{P}$, $\boldsymbol{N}$, $\boldsymbol{D}^{\text{ach},t-1}$, $\boldsymbol{D}^{\text{req},t}$, $\boldsymbol{L}$, $\boldsymbol{T}$.
Output: $\boldsymbol{A}_{RT}$, $\boldsymbol{A}_{BE}$, $\boldsymbol{R}_{RT}$, $\boldsymbol{R}_{BE}$, $\boldsymbol{P}_{RT}$, $\boldsymbol{P}_{\boldsymbol{BE}}$.
1: **if** there are RT users to be scheduled **then**
2: Find the assignment of subbands to the antennas as well as the assignment of single RT users and NOMA BE users using Algorithm 2.
3: Perform NOMA pairing on subbands assigned to RT users using Algorithm 3.
4: **else** // The system consists of BE users only
5: **for** $s = 1 : |\mathcal{S}|$ **do**
6: Schedule BE users using the PF scheduler as in steps 20 to 24 of Algorithm 2.
7: **end for**
8: **end if**

power budget $P_{DAS} = 10$ W. For fair comparison, in case of a CAS setting, the power budget of the antenna is chosen to be $P_{CAS} = 40$ W. The system bandwidth $B = 10$ MHz is divided into $S = 32$ subbands. Signals undergo frequency-selective Rayleigh fading with a root mean square delay spread of 500 ns and a distance-dependent path loss with a decay factor of 3.76. The noise power spectral density is $N_0 = 4.10^{-18}$ mW/Hz and the decay factor for FTPA is $\alpha = 0.5$. Unless otherwise stated, the simulated system consists of $K_{RT} = 20$ RT users and $K_{BE} = 20$ BE users. In this study, we assume perfect channel estimation. The simulation parameters are summarized in Table 3.2.

Table 3.2 – Simulation Parameters

Cell Radius R_d	500 m
Number of antennas A in the DAS setting	4
Overall Transmission Bandwidth B	10 MHz
Number of subbands S	32
Distance Dependent Path Loss	$128.1 + 37.6 \log_{10}(d)$(dB), d in Km
Receiver Noise Density N_0	4.10^{-18} mW/Hz
FTPA decay factor α_{FTPA}	0.5
Power budget per RRH in the DAS setting P_{DAS}	10 Watts (40 dBm)
Power budget per RRH in the CAS setting P_{CAS}	40 Watts (46 dBm)
Number of RT users in the cell K_{RT}	20
Number of BE users in the cell K_{BE}	20

The performance of the allocation technique proposed in this section for the DAS setting with NOMA signaling is denoted by Prop-NOMA-DAS. Its performance is compared with a technique adopting the enhanced PF scheduler [96] to schedule all users, denoted by PF-DAS. The PF-DAS technique does not differentiate between users, i.e., it does not take into account the presence of high-priority RT users in the allocation process. However, for fair comparison, when an RT user k_{RT} reaches its latency limit or its requested number of bits, user k_{RT} is removed from the set of users to be scheduled. The performance of the proposed technique in different system settings is also considered for comparison. More

specifically, to evaluate the performance of NOMA signaling, a version of the proposed technique relying on OMA signaling, denoted as Prop-OMA-DAS, is tested. To evaluate the performance of adopting a DAS setting, a CAS version of the proposed technique, denoted as Prop-NOMA-CAS, is also simulated. Finally, to show the gain from allowing a flexible number of subbands per distributed antenna, the performance of a version of our technique, that restricts the number of subbands per antenna as done in [23], denoted as Prop-NOMA-DAS-F, is presented. In Prop-NOMA-DAS-F, the estimated number of subbands per antenna, given by Eq. (3.10), is not updated throughout the allocation process.

3.3.4.1 Evaluation of the Performance of RT Users

First, the performance of the proposed technique in terms of satisfying RT users is evaluated. To reflect the requirements of different services and applications, RT users are partitioned into 3 classes, denoted by C1, C2 and C3. While all users request $D^{\text{req}} = 10^5$ bits, a user k_{RT} belonging to class C1 has a latency limit of $L_{k_{RT}} = 6$ ms, while a user k_{RT} in class C2 (resp. class C3) has a latency limit of $L_{k_{RT}} = 10$ ms (resp. $L_{k_{RT}} = 15$ ms). Moreover, to test different system congestion levels, three scenarios are simulated. The first one, S1, being the least strict, consists of only $K_{RT} = 5$ RT users belonging to class C3, i.e., the class with the most relaxed latency limit among the considered classes. The second scenario, S2 consists of 2 RT users in class C1, and 9 RT users in each of C2 and C3, leading to a system with a total of $K_{RT} = 20$ RT users. Finally, in the third scenario, S3, 5 RT users belong to each of classes C1 and C2, while 10 RT users belong to class C3, leading once again to a system with a total of $K_{RT} = 20$ RT users. Hence, this third scenario in the most strict among the simulated scenarios. In all three scenarios, the number of BE users is $K_{BE} = 20$.

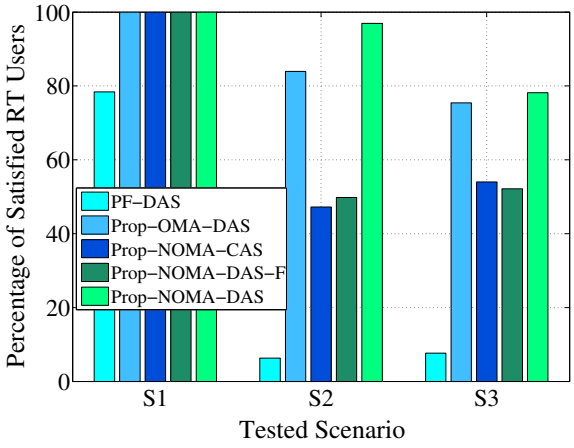


Figure 3.2 – Percentage of Satisfied RT Users

In Fig. 3.2, the performance of the considered techniques in terms of the percentage of satisfied RT users in different scenarios is plotted. As expected, scenario S1, the least strict, has the largest percentage of satisfied users for all techniques. For scenario S1, our technique (in its different versions) guarantees the satisfaction of all RT users. However, PF-DAS can reach satisfaction in only 78% of the cases, which proves the importance of a traffic-aware scheduling technique. As the simulated conditions become harder, the performance of PF-DAS rapidly degrades, reaching less than 10% in both of scenarios S2 and S3. Among the different versions of our proposed technique, the one based on NOMA

signaling in a DAS setting, Prop-NOMA-DAS, has the best performance. For example, Prop-NOMA-DAS guarantees the satisfaction of RT users in 97% of the cases in scenario S2, exceeding the level of satisfaction of both Prop-NOMA-CAS and Prop-NOMA-DAS-F by 45%. In scenario S3, Prop-NOMA-DAS achieves the highest performance once again with a 78% satisfaction level, outperforming the satisfaction level of both Prop-NOMA-CAS and Prop-NOMA-DAS-F by close to 25%. In comparison with the OMA version of the proposed technique, Prop-NOMA-DAS outperforms Prop-OMA-DAS by 13% in terms of RT users satisfaction in scenario S2. In fact, in S2, some RT users achieve rates that are higher than the required ones in the OMA step. Thus, these users can be paired with other RT users in the NOMA step, increasing the satisfaction percentage for NOMA signaling. This is not the case for scenario S3 where the requirements of RT users are more strict, increasing the number of RT users that cannot share their subbands. This leads to a close performance for OMA and NOMA signaling, with the NOMA version slightly outperforming its OMA counterpart.

For unsatisfied users in scenarios S2 and S3, the measured average number of received data bits for Prop-NOMA-DAS is 9.7×10^4. This number decreases to 5.75×10^4 for PF-DAS. Therefore, while Prop-NOMA-DAS cannot satisfy all users, it allows unsatisfied users to get closer to their requirement ($D^{\text{req}} = 10^5$ bits).

To conclude, Fig. 3.2 shows the performance gain achieved by adopting a traffic-aware solution to schedule users. Moreover, the performance gain from relying on NOMA signaling in a DAS setting is also validated. Fig. 3.2 also shows that the performance gain obtained by adopting a varying number of subbands per antenna throughout the allocation process in the distributed setting is significant with respect to a non-adaptive approach as in [23].

3.3.4.2 Evaluation of the Performance of BE Users

To evaluate the performance of BE users, we consider two system-level performance indicators: the achieved system throughput and the user fairness. The latter is assessed through Jain's fairness index [24]:

$$J = \frac{\left(\sum\limits_{k \in \mathcal{K}_{BE}} R_k\right)^2}{K_{BE} \sum\limits_{k \in \mathcal{K}_{BE}} R_k^2}, \tag{3.28}$$

where R_k is the total achieved throughput by BE user k within a timeslot. Jain's fairness index ranges between 0 and 1 with the maximum achieved in the case of absolute fairness.

Fig. 3.3 plots the fairness achieved by the different techniques in the simulated settings. Fig. 3.3 shows that PF-DAS and Prop-NOMA-DAS have a similar performance in terms of user fairness. This similar performance shows that the proposed allocation technique does not jeopardize the fairness between BE users. Moreover, Prop-NOMA-DAS outperforms both Prop-OMA-DAS and Prop-NOMA-CAS, further validating the performance gain of adopting NOMA signaling in a DAS setting. Fig. 3.3 also shows that Prop-NOMA-DAS greatly outperforms Prop-NOMA-DAS-F in terms of user fairness. In fact, by enforcing a certain number of subbands per antenna in a DAS setting as in Eq. (3.10), Prop-NOMA-DAS-F does not guarantee BE users are served on their best perceived subbands. This

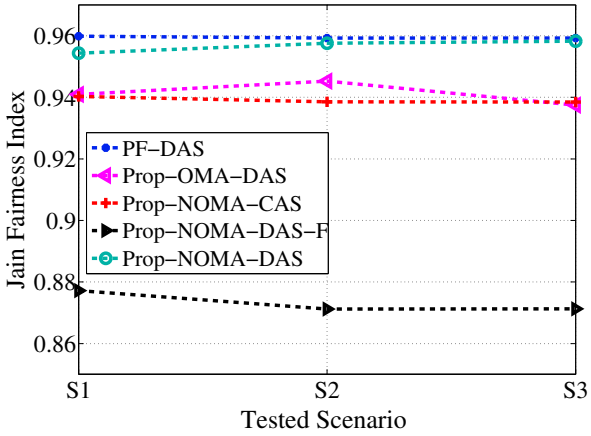
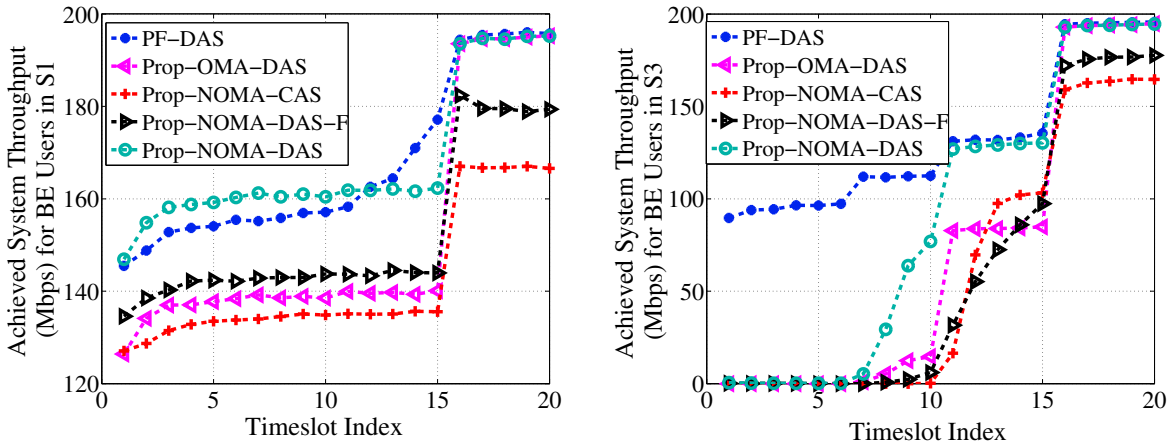


Figure 3.3 – Fairness Achieved in S1, S2 and S3 by the Different Techniques

is not the case for Prop-NOMA-DAS where a flexible number of subbands per antenna is allowed.



(a) Scenario 1 (b) Scenario 3

Figure 3.4 – Evolution of the achieved system throughput in Mbps for BE users per timeslot for two different scenarios

Fig. 3.4 shows the evolution of the system throughput as time progresses, for scenarios S1 and S3, encompassing systems with relaxed and strict requirements. For both scenarios, Prop-NOMA-DAS outperforms Prop-OMA-DAS, Prop-NOMA-CAS and Prop-NOMA-DAS-F. Hence, our decision to use NOMA and DAS for a mixed traffic system with a flexible number of subbands per antenna in the DAS setting is justified.

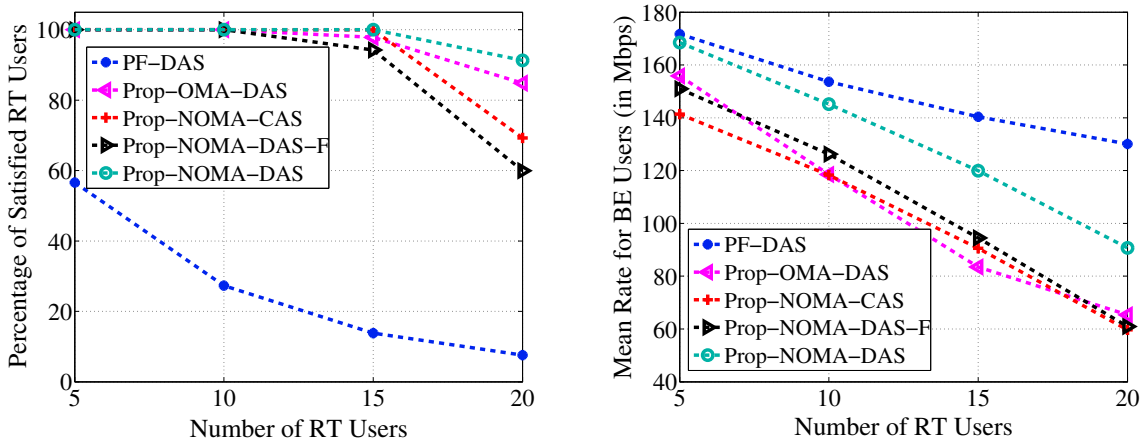
Moving to the comparison of Prop-NOMA-DAS with PF-DAS, for scenario S1, Prop-NOMA-DAS outperforms PF-DAS in the first 12 timeslots, when most of the RT users are still in the active set of the allocation process. This is because Prop-NOMA-DAS gives the minimum necessary amount of resources to RT users in order to maximize the performance of BE users. Starting from timeslot 13, PF-DAS achieves a higher data rate than Prop-NOMA-DAS. The higher data rate is due to the fact that the PF scheduler aims at maximizing system throughput and user fairness. Therefore, some RT users might receive their required number of bits before reaching their latency limit and exit the system, leaving more resources for BE users. Once all RT users have exited the system

(at the end of timeslot 15), Prop-NOMA-DAS achieves a similar performance to PF-DAS. However, when averaging system throughput over time, we find that PF-DAS achieves a mean BE throughput of 167 Mbps while our technique achieves 168 Mbps, hence has a slightly superior performance. Also, recall that PF-DAS satisfies RT users in only 78% of the cases in scenario S1 compared to a 100% satisfaction level for Prop-NOMA-DAS as shown in Subsection 3.3.4.1.

Moving to scenario S3, we notice that Prop-NOMA-DAS does not give any resources to BE users prior to timeslot 6. In fact, until timeslot 6, all 20 RT users are awaiting scheduling, and Prop-NOMA-DAS prioritizes them in the resource allocation, which is not the case for PF-DAS. Starting from timeslot 6, Prop-NOMA-DAS starts giving more and more resources to BE users as the system becomes less congested. After timeslot 15, all RT users exit the system, resulting in Prop-NOMA-DAS and PF-DAS achieving similar performance.

Finally, in Fig. 3.5, the performance of the different algorithms is shown for a varying number of RT users. The following cases are studied:

1. $K_{RT} = 5$ users (1 user in C1, 1 user in C2 and 3 users in C3),
2. $K_{RT} = 10$ users (2 users in C1, 2 users in C2 and 6 users in C3),
3. $K_{RT} = 15$ users (3 users in C1, 3 users in C2 and 9 users in C3),
4. $K_{RT} = 20$ users (4 users in C1, 4 users in C2 and 12 users in C3).



(a) Percentage of satisfied RT users (b) Mean rate for BE users

Figure 3.5 – Evolution of the system performance for RT and BE users in terms of the number of RT users K_{RT}

Fig. 3.5a shows the superiority of the proposed technique in terms of satisfying RT users. In fact, Prop-NOMA-DAS satisfies RT users in over 90% of the cases, regardless of their number, outperforming the OMA version and the CAS version by up to 6% and 22% respectively when $K_{RT} = 20$ users. Moreover, Prop-NOMA-DAS outperforms Prop-NOMA-DAS-F by up to 31% in terms of the satisfaction level of RT users. In comparison, PF-DAS satisfies RT users in 57% of the cases when $K_{RT} = 5$, and 9% of the cases when $K_{RT} = 20$.

Fig. 3.5b shows that the mean rate achieved by BE users is a decreasing function of the number of RT users for the different techniques. In fact, as K_{RT} increases, more resources are allocated to RT users, decreasing the amount of available resources for, and hence the performance of, BE users. As shown in Fig. 3.5b, PF-DAS outperforms the proposed technique in terms of maximizing the mean rate of BE users as the latter reserves more resources for RT users. Moreover, Prop-NOMA-DAS outperforms Prop-OMA-DAS, Prop-NOMA-CAS and Prop-NOMA-DAS-F by up to 25 Mbps.

3.4 Matching Theory-Based Solution

In this section, we invoke the two-sided matching theory framework [98] to obtain a suboptimal solution for the formulated problem in (3.9). To test the different studied system settings, i.e., OMA-CAS, OMA-DAS, NOMA-CAS and NOMA-DAS, a solution based on the matching theory framework is formulated for each setting.

3.4.1 Matching Theory-Based Solution in the OMA-CAS setting

In an OMA-CAS, a single central base station is deployed. Moreover, each subband can be assigned to one user only since OMA signaling is adopted. Hence, the antenna index a can be dropped from all involved variables and the optimization problem (3.9) can be re-formulated as:

$$\max_{\boldsymbol{x},\boldsymbol{p}} \quad (3.2) \tag{3.29}$$

$$\begin{aligned}
\text{such that} \quad & (3.8), && \text{(3.29a)}\\
& \sum_{k\in\mathcal{K}} x_{k,s} \leq 1, \ \forall\, s \in \mathcal{S}, && \text{(3.29b)}\\
& \sum_{k\in\mathcal{K}} \sum_{s\in\mathcal{S}} P_s x_{k,s} \leq P_{CAS}, && \text{(3.29c)}\\
& x_{k,s} \in \{0,1\}. && \text{(3.29d)}
\end{aligned}$$

In problem (3.29), variable $x_{k,s}$ now indicates the assignment of users to subbands, i.e., $x_{k,s} = 1$ if user k is scheduled on subband s and 0 otherwise. Moreover, constraints (3.9c) and (3.9d) are adapted to the OMA-CAS setting. Constraints (3.9b) and (3.9e) are discarded as they are inapplicable in an OMA-CAS context.

Assuming an equal inter-subband power allocation, solving problem (3.29) consists of finding the user-subband assignment optimizing the system performance.

3.4.1.1 Subband Allocation Problem as a Matching Game

To develop a low-complexity solution for (3.29), the subband assignment problem can be modeled as a two-sided one-to-many matching game. In this model, the set of subbands $\mathcal{S}$ and the set of users $\mathcal{K} = \mathcal{K}_{RT} \cup \mathcal{K}_{BE}$ form two disjoint sets of selfish and rational agents. In this first system setting, a subband s can be assigned to at most one user while a user k can be matched with more than one subband. If a user k is scheduled on subband s, the pair (k, s) forms a *matching pair*.

A one-to-many matching Ψ is defined as a mapping from the set $\mathcal{K} \cup \mathcal{S}$ into the set of all subsets of $\mathcal{K} \cup \mathcal{S}$ such that for each $k \in \mathcal{K}$ and $s \in \mathcal{S}$:

1. $\Psi(k) \subseteq \mathcal{S}, \forall k \in \mathcal{K}$,
2. $\Psi(s) \subseteq \mathcal{K}, \forall s \in \mathcal{S}$,
3. $|\Psi(s)| = 1, \forall s \in \mathcal{S}$,
4. $s \in \Psi(k) \Leftrightarrow k \in \Psi(s)$.

To reach a final matching Ψ, each player p builds a preference relation $\succ_p$ over the players from the other set. Using these predefined preference relations, players dynamically interact with each other to reach a stable matching. Matching Ψ is stable when there are no user k and subband s that are not matched together, but prefer each other over their partners $\Psi(k)$ and $\Psi(s)$, respectively. Hence, the subband assignment problem can be represented by the tuple $(\mathcal{S}, \mathcal{K}, \succ_s, \succ_k)$.

3.4.1.2 Preference Lists

To decide on the outcome of the game, each user k builds a preference list $\mathcal{PL}_k$ over the set of subbands $\mathcal{S}$. The preference list $\mathcal{PL}_k$ is sorted in a descending order based on the channel gain experienced by user k over all subbands in the set $\mathcal{S}$. In other words, $\mathcal{PL}_k$ is based on Definition 18.

Definition 18. *Assuming equal inter-subband power allocation, users base their preferences on the channel gains over different subbands. Put differently, user k prefers subband s_i over s_j, i.e., $s_i \succ_k s_j$, where $s_i, s_j \in \mathcal{S}$, if $h_{k,s_i} > h_{k,s_j}$.*

Similarly, each subband s_i bases its preferences over the set of users according to Definition 19.

Definition 19. *Subbands must account for the heterogeneity of users while building their preference lists by always preferring RT users over BE users, since RT users have the highest priority. Therefore, we define the preference relation of subband s_i as:*

$$k_{RT} \succ_{s_i} k_{BE}, \forall\, s_i \in \mathcal{S},\ k_{RT} \in \mathcal{K}_{RT} \text{ and } k_{BE} \in \mathcal{K}_{BE}.$$

In addition to preferring RT users, subbands must also be able to separately prioritize among the set of RT users and that of BE users. Therefore, at timeslot t, the following utility metric, inspired by the proposed metrics of [85, 86], is introduced for RT users:

$$\mathcal{U}^t_{RT}(k^l_{RT}, s_i) = \frac{R^t_{k^l_{RT}, s_i}}{R^t_{k^l_{RT}}} \times \frac{t}{L_{k^l_{RT}}}. \tag{3.30}$$

In Eq. (3.30), $R^t_{k^l_{RT}, s_i}$ denotes the achievable rate by user k^l_{RT} over subband s_i if matched together at timeslot t and $R^t_{k^l_{RT}}$ is the rate already achieved by k^l_{RT} before reaching timeslot t. Eq. (3.30) is proportional to the achieved rate of user k_{RT} over subband s_i, and to the timeslot index t. Therefore, the users with the highest priorities are those who would benefit the most from subband s_i in terms of rate, and who have most approached their

latency limit. Moreover, Eq. (3.30) is inversely proportional to the already achieved rate of user k_{RT}, $R^t_{k^l_{RT}}$, as well as its the latency limit. By considering the timeslot index t, the utility metric for RT user k^l_{RT}, $\mathcal{U}^t_{RT}(k^l_{RT}, s_i)$, grows larger as t increases. Also, by accounting for the latency limit $L_{k^t_{RT}}$, $\mathcal{U}^t_{RT}(k^l_{RT}, s_i)$ is increased for a more stringent latency requirement. In addition to considering the timeslot index and the latency limit, Eq. (3.30) accounts for the transmission rate of k^l_{RT} if it were scheduled on s_i. This enables users with a better channel quality to have a higher priority during scheduling, therefore increasing spectral efficiency. Last, by also accounting for the achieved data rate of k^l_{RT} prior to reaching timeslot t, Eq. (3.30) achieves a certain fairness level between RT users by prioritizing users that have not previously achieved a large enough throughput. Each subband s_i bases its preferences over RT users according to the following definition.

Definition 20. *At timeslot t, subband s_i prefers user k^l_{RT} over user k^m_{RT}, i.e., $k^l_{RT} \succ_{s_i} k^m_{RT}$, if $\mathcal{U}^t_{RT}\left(k^l_{RT}, s_i\right) > \mathcal{U}^t_{RT}\left(k^m_{RT}, s_i\right)$.*

Similarly, subbands must differentiate between BE users, and the utility metric should strike a tradeoff between fairness and rate maximization. Hence, the PF scheduler metric [28] is adopted:

$$\mathcal{U}^t_{BE}\left(k^l_{BE}, s_i\right) = \frac{R^t_{k^l_{BE}, s_i}}{T^t_{k^l_{BE}}}. \tag{3.31}$$

It should be noted that Eq. (3.31) aims to maximize the performance measure for BE users formulated in Eq. (3.2), as the fairness measure $f(T^t_{k_{BE}})$ in Eq. (3.2) is represented by the weight $1/T^t_{k_{BE}}$ in Eq. (3.31). It should also be noted that the enhanced PF scheduler [96] cannot be adopted as the utility function of BE users when matching theory is used to assign users to subbands. In fact, as shown in Section 3.4.1.3, the assignment of all the subbands is determined at once. Hence, the rate achieved during the current allocation slot t cannot be accounted for in Eq. (3.31). Each subband s_i bases its preferences over BE users at timeslot t according to Definition 21.

Definition 21. *At timeslot t, subband s_i prefers user k^l_{BE} over user k^m_{BE}, i.e., $k^l_{BE} \succ_{s_i} k^m_{BE}$, if $\mathcal{U}^t_{BE}\left(k^l_{BE}, s_i\right) > \mathcal{U}^t_{BE}\left(k^m_{BE}, s_i\right)$.*

Having defined the preference relations for both sets of players, the proposed algorithm to solve the formulated matching game in the OMA-CAS setting is described next.

3.4.1.3 Proposed OMA-CAS DA Algorithm

Since users have different priority levels, the classical DA algorithm [98] cannot be directly used to solve the considered matching game. That is why, in Algorithm 5, a priority-aware version of the DA algorithm suitable for the studied context in the OMA-CAS setting is proposed.

Initially, the set of active users $\mathcal{K}_{\text{active}}$ is built; it includes all RT users that have not yet received their requested data rate in the current timeslot t and all BE users. Each user $k \in \mathcal{K}_{\text{active}}$ builds its preference list over the set of subbands, $\mathcal{PL}_k$. In its turn, each subband s_i initializes its matching set, $\mathcal{M}(s_i)$, consisting of the user to which it is assigned throughout the algorithm. At the first iteration of the algorithm, $\mathcal{M}(s_i)$ is initialized to the empty set for every subband $s_i \in \mathcal{S}$.

In the first phase of the algorithm, each active user k applies to its most preferred subband, i.e., the very first element in $\mathcal{PL}_k$, as shown in step 9 of Algorithm 5. Each subband s_i receiving user proposals forms an applying user set denoted by $\mathcal{AS}(s_i)$. The proposing set $\mathcal{AS}(s_i)$ of subband s_i consists of users having proposed to s_i during the current iteration, in addition to the user matched to subband s_i at the previous iteration, i.e., the user in set $\mathcal{M}(s_i)$. The update of the proposing set $\mathcal{AS}(s_i)$ for each subband $s_i \in \mathcal{S}$ is presented in step 13 of the algorithm. Note that the proposing set $\mathcal{AS}(s_i)$ can be empty if none of the users apply to subband s_i at the current iteration and subband s_i was not assigned to any user at the previous iteration of the algorithm, i.e., $\mathcal{M}(s_i)$ is empty.

Having a system with heterogeneous users, subband s_i must prioritize RT users in the decision phase, i.e., in Phase 2 of Algorithm 5. Therefore, for each subband s_i receiving user proposals, i.e., having $\mathcal{AS}(s_i) \neq \emptyset$, two cases are considered:

1. if RT users are among the applicants, the most preferred RT user according to Eq. (3.30), k^*_{RT}, is accepted by subband s_i, as per step 15. All users in $\mathcal{AS}(s_i) \setminus \{k^*_{RT}\}$ are rejected by subband s_i and the active user set, $\mathcal{K}_{\text{active}}$, is updated to reflect the resulting rate requirement changes. In other words, if scheduling the most preferred RT user k^*_{RT} results in k^*_{RT} receiving its total rate requirement, user k^*_{RT} is removed from the active user set.

2. if a subband receives applications from BE users only, it accepts the most preferred BE user according to Eq. (3.31), as per step 18 of the algorithm. All remaining users in $\mathcal{AS}(s_i) \setminus \{k^*_{BE}\}$ are rejected by subband s_i.

At the end of the second phase of the algorithm, every user removes the subband that it proposed to at the current iteration from its preference list, as shown in Phase 3 of the algorithm.

Phases 1, 2 and 3 of Algorithm 5 are repeated until the preference lists of all active users are empty. Upon termination, the optimal matching result, i.e., the optimal user-subband assignment, is obtained.

3.4.2 Matching Theory-Based Solution in the OMA-DAS Context

In an OMA-DAS system, A antennas are deployed in the cell. Furthermore, each subband can be assigned to one user only due to OMA signaling. Hence, the optimization problem (3.9) can be re-formulated as:

$$\max_{\boldsymbol{x},\boldsymbol{p}} \quad (3.2) \tag{3.32}$$

$$\begin{aligned}
\text{such that} \quad & (3.8), && \text{(3.32a)}\\
& \sum_{a\in\mathcal{A}} x_{k,s,a} \leq 1, \ \forall\,(k,s) \in \mathcal{K}\times\mathcal{S}, && \text{(3.32b)}\\
& \sum_{k\in\mathcal{K}} x_{k,s,a} \leq 1, \ \forall\,(s,a) \in \mathcal{S}\times\mathcal{A}, && \text{(3.32c)}\\
& \sum_{k\in\mathcal{K}}\sum_{s\in\mathcal{S}} P_{s,a}x_{k,s,a} \leq P_{DAS}, \ \forall\, a \in \mathcal{A}, && \text{(3.32d)}\\
& x_{k,s,a} \in \{0,1\}, && \text{(3.32e)}
\end{aligned}$$

where constraint (3.9c) is adapted to the OMA setting. Moreover, constraint (3.9e) is discarded as it is inapplicable in an OMA context.

Assuming an equal inter-subband power repartition on each distributed antenna, a solution for the user assignment to subbands and antennas is next found using the matching theory framework.

Algorithm 5 Priority-Aware OMA DA Algorithm

Input: $\mathcal{K}_{RT}, \mathcal{K}_{BE}, \boldsymbol{H_{RT}}, \boldsymbol{H_{BE}}, \boldsymbol{R}^{\text{req}}_{\boldsymbol{RT}}, \boldsymbol{L_{RT}}, \boldsymbol{T_{BE}}, t, \boldsymbol{N}$.

Output: $\boldsymbol{A_{RT}}, \boldsymbol{A_{BE}}, \boldsymbol{R_{RT}}, \boldsymbol{R_{BE}}$. // $\boldsymbol{A_{RT}}$ (resp. $\boldsymbol{A_{BE}}$) is the assignment matrix of RT users (resp. BE users) to subbands and antennas while $\boldsymbol{R_{RT}}$ (resp. $\boldsymbol{R_{BE}}$) denotes their achieved rates over each (subband, antenna) pair

Initialization:

if CAS setting **then**

$\mathcal{K}^{RT}_{\text{active}} = \left\{k_{RT} \in \mathcal{K}_{RT} / R^{\text{req}}_{k_{RT}} > 0\right\}$,

$\mathcal{K}_{\text{active}} = \mathcal{K}_{BE} \cup \mathcal{K}^{RT}_{\text{active}}$.

else if DAS setting **then**

Construct virtual user sets $\mathcal{K}^{v}_{RT}$ and $\mathcal{K}^{v}_{BE}$.

$\mathcal{K}^{RT}_{\text{active}} = \left\{k \in \mathcal{K}_{RT} \,/\, R^{\text{req}}_{k_{RT}} > 0\right\}$, $\mathcal{K}^{RT,v}_{\text{active}} = \left\{v^{k_{RT},a}, a = 1, \ldots, A \,/\, R^{\text{req}}_{k_{RT}} > 0\right\}$,

$\mathcal{K}^{r}_{\text{active}} = \mathcal{K}_{BE} \cup \mathcal{K}^{RT}_{\text{active}}$, $\mathcal{K}_{\text{active}} = \mathcal{K}^{v}_{BE} \cup \mathcal{K}^{RT,v}_{\text{active}}$.

end if

Build the preference lists $\mathcal{PL}_k$ of users in $\mathcal{K}_{\text{active}}$.

Set $\mathcal{M}(s_i) = \emptyset, \forall\, s_i \in \mathcal{S}$.

Phase 1: Active users apply to subbands

if CAS setting **then**

Each user $k \in \mathcal{K}_{\text{active}}$ proposes to the first subband in $\mathcal{PL}_k$.

else if DAS setting **then**

Each real user $k \in \mathcal{K}^{r}_{\text{active}}$ chooses its proposing virtual user $v^{k,a}_{p}$, according to Eq. (3.33), and proposes to the first subband in $\mathcal{PL}_{v^{k,a}_{p}}$.

end if

Find the applicant set $\mathcal{AS}(s_i)$ for each subband $s_i \in \mathcal{S}$, i.e., $\mathcal{AS}(s_i) = \mathcal{AS}(s_i) \cup \mathcal{M}(s_i),\ \forall s_i \in \mathcal{S}$.

Phase 2: Subbands make decisions

if $\mathcal{AS}(s_i) \neq \emptyset$ and $\mathcal{AS}(s_i) \cap \mathcal{K}^{RT}_{\text{active}} \neq \emptyset$ **then**

$\mathcal{M}(s_i) = \{k^{*}_{RT}\}$, where $k^{*}_{RT} = \underset{k^{l}_{RT} \in \mathcal{AS}(s_i) \cap \mathcal{K}^{RT}_{\text{active}}}{\text{argmax}} \mathcal{U}^{t}_{RT}(k^{l}_{RT}, s_i)$.

Update $R^{\text{req}}_{k^{*}_{RT}}$, $\mathcal{K}^{RT}_{\text{active}}$, $\mathcal{K}^{RT,v}_{\text{active}}$ and $\mathcal{K}_{\text{active}}$.

else if $\mathcal{AS}(s_i) \neq \emptyset$ **then**

$\mathcal{M}(s_i) = \{k^{*}_{BE}\}$, where $k^{*}_{BE} = \underset{k^{l}_{BE} \in \mathcal{AS}(s_i)}{\text{argmax}} \mathcal{U}^{t}_{BE}(k^{l}_{BE}, s_i)$.

end if

Phase 3: Preference lists update

Each user k that proposed to subband $s_i, \forall\, s_i \in \mathcal{S}$, removes s_i from $\mathcal{PL}_k$.

Repeat Phases 1, 2 and 3 until $\mathcal{PL}_k = \emptyset,\ \forall\, k \in \mathcal{K}_{\text{active}}$

3.4.2.1 DAS Matching Game Model and Algorithm

In this section, we aim to apply matching theory to solve the subband allocation problem in the DAS context. However, the DAS layout brings a new dimension into the resource allocation problem: antenna association. In addition to the user-subband assignment of the CAS context, it is also necessary to decide on the serving antenna for each subband assigned to a user. This new dimension complicates the problem considerably and renders the application of matching theory challenging. In fact, since we have to associate each user with a subband and an antenna, we are faced with a three-dimensional matching problem for which a stable solution is not guaranteed to exist [99].

Most previous studies on resource allocation in distributed settings make some assumptions with the aim of making the problem tractable. For example, [37] focused on antenna selection and power allocation to maximize the EE of a DAS setting, where the spectrum was assumed to consist of one subband only. However, in practical systems, this assumption does not hold. Maximizing the EE was also the purpose of [100], where the authors associated the user with the antenna providing the best average signal-to-interference-plus-noise-ratio (SINR), before proceeding with the subband and power allocation steps. Although this antenna selection scheme may seem logical, it might result in an antenna being associated a large number of users. This decreases the power available to each user on that antenna and some users may benefit from being assigned to other, less congested antennas. Moreover, in [100], a user was constrained to be associated with one antenna only, and all RRHs have access to the whole spectrum which increases the interference. In [23], the subband and power allocation steps were separated. For the subband assignment, the number of subbands per antenna was estimated based on the average path-loss of the users, and the actual subband assignment was performed with the aim of maximizing proportional fairness.

In the current solution, a user is not restricted to be assigned to a unique antenna. Moreover, a subband can be assigned to one antenna only to limit interference. In addition to that, no *a priori* information about the distribution of subbands among RRHs is assumed. To overcome these challenges, the concept of virtual users is introduced:

Definition 22. *A user k is duplicated into A virtual users, A being the number of antennas in the cell. Each virtual user $v^{k,a}$ represents the potential association of a real user k with a serving antenna a.*

Introducing the concept of virtual users leads to a total of $K \times A$ virtual users in the system. However, this transformation recovers the two-dimensional aspect of the resource allocation problem, which makes it possible to find a solution using matching theory.

As in the CAS case, the sets of players in the matching game as well as the preference lists must be defined. In the DAS context, the sets participating in the matching game are the set of virtual users $\mathcal{KV}$, consisting of RT and BE virtual users, and the set of subbands $\mathcal{S}$. Virtual users and subbands also build their preference relations according to Definitions 18 and 19 respectively. However, the algorithm conceived for the CAS case must be revisited.

In fact, according to the algorithm conceived for the OMA-CAS setting, each active user proposes to its preferred subband during each iteration. In the DAS setting, the user set consists of virtual users, representing the different associations of real users to serving antennas. Allowing each active virtual user to propose to its most preferred subband might result in a real user being allocated more than one subband at each iteration. This

variation of the algorithm (i.e., allowing each virtual user to propose at each iteration) was compared with a second one that restricts each real user to apply through one virtual user only. Simulations showed that the second variation yields better results. In fact, by restricting a real user to apply to subbands through one, well chosen, virtual user at each iteration, a better spectrum usage is achieved, enhancing the performance of all users in the system. Therefore, in the proposed solution for the OMA-DAS case, this second variation is retained. In other words, at each iteration, among virtual users pertaining to the same real user, only one is allowed to propose to its most preferred subband. This proposing virtual user is selected to guarantee the best performance among virtual users relative to the same real user. Put differently, each real user must aim to be assigned to the antenna guaranteeing the best performance. However, the choice of the best antenna should not only take into consideration the channel gains of the users. In fact, the power levels of the subbands generally differ between antennas, depending on the respective congestion levels of the antennas. Consequently, the power level per subband on each antenna should also be considered in the decision process. Assuming equal inter-subband power allocation on each antenna, the power allocated to each subband assigned to antenna a (hence to each virtual user associated with antenna a) is given by Eq. (3.11), with the derivation of N_a, the number of subbands assigned to antenna a, detailed in Section 3.4.2.2. Then, each real user chooses the proposing virtual user according to Definition 23.

Definition 23. *Proposing virtual user $v_p^{k,a}$, representing the association of real user k with antenna a, is the one satisfying:*

$$v_p^{k,a} = \underset{\substack{v^{k,a'},\\ a'=1,\cdots,A}}{\operatorname{argmax}} \left(P_{a'} \times h^2_{v^{k,a'},s^*_{v^{k,a'}}} \right). \tag{3.33}$$

In Eq. (3.33), $v^{k,a'}$ represents the virtual user associated to antenna a' and relating to real user k. Subband $s^*_{v^{k,a'}}$ is the preferred subband of virtual user $v^{k,a'}$, i.e., the very first one in its preference list $\mathcal{PL}_{v^{k,a'}}$. Choosing the proposing user following Eq. (3.33) ensures that real users are matched with their best subbands and antennas, in terms of rate maximization.

Algorithm 5 summarizes the proposed DA resource allocation algorithm in the OMA-DAS context, when the number of subband per antenna, N_a, is known for every antenna $a \in \mathcal{A}$.

As in the OMA-CAS case, each active user $k \in \mathcal{K}_{\text{active}}$ builds a preference list over the set of subbands denoted by $\mathcal{PL}_k$. Also, each subband $s_i \in \mathcal{S}$ initializes its matching set $\mathcal{M}(s_i)$.

In the first phase of the algorithm, each active real user proposes to its preferred subband by choosing its proposing virtual user as in step 11 of the algorithm. Each subband s_i receiving user proposals updates its proposing user set $\mathcal{AS}(s_i)$ by adding both the proposing users at the current iteration, and the matched user at the previous iteration.

In the second phase of the algorithm, subbands decide on their current matching in the same manner as in the OMA-CAS case. Once decisions are made, each proposing virtual user removes the subband to which it proposed from its preference list. Note that since at each iteration, only the proposing virtual user is allowed to apply to its preferred subband, the preference list of the proposing virtual user is the only one to get updated.

The preference lists of the rest of the virtual users representing the same real user are kept unchanged.

Phases 1, 2 and 3 of Algorithm 5 are repeated until the preference lists of all active users become empty. Upon termination, the assignment of users to the subbands and the antennas is simultaneously obtained.

3.4.2.2 Estimation of the number of subbands per antenna

In this work, we do not assume any *a priori* information about the assignment of subbands to the different antennas, which leads us to introduce an algorithm to find this assignment in this section.

As in Section 3.3.1, for every antenna $a \in \mathcal{A}$, the initial estimation of the number of assigned subbands per antenna, N_a, is found according to Eq. (3.10). Using this initial estimation, Algorithm 6 provides the final assignment of users and subbands to antennas. In each iteration of Algorithm 6, each antenna performs equal-power distribution using its estimated number of assigned subbands. Then, subband and user assignment is determined using Algorithm 5. The number of subbands per antenna is then updated, as well as the power per subband on each antenna, P_a, according to the assignment yielded by Algorithm 5.

These steps are repeated until convergence, i.e., until the number of subbands per antenna remains unchanged between two successive iterations. Upon convergence, the best assignment of user-subband pairs to the different antennas, in terms of rate maximization, is obtained.

Algorithm 6 OMA-DAS Resource Allocation

Input: $\boldsymbol{N}, \mathcal{K}_{RT}, \mathcal{K}_{BE}, \boldsymbol{H_{RT}}, \boldsymbol{H_{BE}}, \boldsymbol{R}_{\boldsymbol{RT}}^{\text{req}}, \boldsymbol{L_{RT}}, \boldsymbol{T_{BE}}, t$.
Output: $\boldsymbol{A_{RT}}, \boldsymbol{A_{BE}}, \boldsymbol{R_{RT}}, \boldsymbol{R_{BE}}, \boldsymbol{N}$. // $\boldsymbol{N}$ is the number of subbands per antenna

1: ***Repeat:***
2: $N^{\text{old}} = N$.
3: $P_a = P_{DAS}/N_a$.
4: Find the assignment of users and subbands to antennas according to Algorithm 5.
5: Using the resulting $\boldsymbol{A_{RT}}$ and $\boldsymbol{A_{BE}}$, re-calculate $\boldsymbol{N} \in \mathbb{N}^{A\times 1}$ and $\boldsymbol{P} \in \mathbb{R}_{+}^{A\times 1}$ as well as $\boldsymbol{R_{RT}}$ and $\boldsymbol{R_{BE}}$.
6: ***Until convergence***

3.4.3 Matching Theory-Based Solution in the NOMA-CAS Context

Solving the resource allocation problem in the NOMA context using matching theory is not straightforward. On the one hand, the power multiplexing constraint, formulated in (3.9e) and neglected in previous works like [20], must be respected to guarantee SIC stability. On the other hand, applying the methods proposed in previous works like [20] and [101] does not guarantee the rate requirements of RT users. Nor does applying the same algorithms proposed in Sections 3.4.1 and 3.4.2, while allowing multiple users to be scheduled on the same subband. Moreover, because of the interdependencies between users' preferences, due to the inter-user interference between paired users on NOMA

subbands, the outcome of these algorithms is not guaranteed to be optimal. For these reasons, in this section, we generalize the resource allocation techniques proposed in the previous sections to encompass the NOMA case, starting with the NOMA-CAS context.

In a NOMA-CAS, one antenna is deployed in the cell and each subband can be assigned to at most two users. Hence, the optimization problem (3.9) can be re-formulated as:

$$\max_{\boldsymbol{x},\boldsymbol{p}} \quad (3.2) \tag{3.34}$$

$$\text{such that} \quad (3.8), \tag{3.34a}$$

$$\sum_{k\in\mathcal{K}} x_{k,s} \leq 2, \ \forall\, s \in \mathcal{S}, \tag{3.34b}$$

$$\sum_{k\in\mathcal{K}} \sum_{s\in\mathcal{S}} P_s x_{k,s} \leq P_{CAS}, \tag{3.34c}$$

$$p_{k_1,s} < p_{k_2,s}, \ \forall\, s \in \mathcal{S}, \tag{3.34d}$$

$$x_{k,s} \in \{0,1\}. \tag{3.34e}$$

Once again, variable $\boldsymbol{x}$ in problem (3.34) denotes the assignment of users to the different subbands, where $x_{k,s} = 1$ if user k is assigned on subband s, and 0 otherwise. Constraint (3.9b) is discarded since one antenna is considered in a CAS. Moreover, the subscript a is dropped from all involved variables.

Solving problem (3.34) consists of assigning users to subbands in a NOMA manner to maximize the system performance, i.e., maximize expressions (3.2) and (3.8).

The proposed solution to the resource allocation problem in the NOMA-CAS context is divided into two stages. The first one, consisting of the assignment of single RT users, i.e., OMA RT users, and NOMA BE users, aims at maximizing the number of satisfied RT users, as well as boosting the performance of BE users when possible. The second stage, where user pairings are performed on the subbands allocated to RT users, aims at further enhancing system performance.

3.4.3.1 Assignment of Subbands to Single RT Users and NOMA BE Users

To ensure that rate requirements of RT users are met, the allocation process starts by scheduling OMA RT users, as done in section 3.4.1. The preference relations of RT users and subbands are formulated in the same way as in section 3.4.1, i.e., according to Definitions 18 and 19. If all RT users reach their rate requirements and free subbands remain in the system, BE users are scheduled on these subbands directly via NOMA.

NOMA Matching Game for BE Users

In [20], matching theory was used to solve the subband allocation problem in NOMA and a technique based on the DA algorithm was proposed. In the presented algorithm, each user proposes to its preferred subband. The latter accepts the proposals of the N_s users achieving the highest performance, N_s being the maximum number of NOMA paired users. However, interdependencies between users' preferences exist in the NOMA case, because of the interference brought by non-orthogonally scheduling different users on the same subband. This is why previous work as in [21] and [101] performed a swapping step at the end of the matching step to further enhance system performance. This step

allows different users to swap their assigned subbands, conditioned by the approval of all involved players, and generally requires a significant number of additional iterations.

In this study, we follow a different approach by allowing single users as well as pairs of users to make proposals. By doing so, interdependencies between users are directly taken into account, without the need for an additional swap phase. Also, thanks to this new idea, a hybrid-NOMA system (where subbands are either allocated to sole or paired users) is enabled, which achieves a better performance than that of non-hybrid NOMA, as was shown in [102].

Similarly to the previous sections, the sets of players for scheduling BE users in the NOMA-CAS case are subbands and BE users. However, the user set, denoted by $\mathcal{US}$, now consists of both single BE users and pairs of BE users. In other words, the cardinality of set $\mathcal{US}$ is given by:

$$|\mathcal{US}| = |K_{BE}| + \binom{K_{BE}}{2}, \tag{3.35}$$

where the first term accounts for single BE users and the second for BE user pairs.

Each user combination in $\mathcal{US}$ complies with the following definition to form its preference list.

Definition 24. *Each user combination $us_n \in \mathcal{US}$ divides the available power per subband among its members according to the FTPA [75], i.e., according to Eq. (3.16) (Note that if $|us_n| = 1$, the available power per subband is entirely allocated to the sole user in us_n). Then, us_n bases its preferences on:*

$$s_i \succ_{us_n} s_j \;\; \textit{if} \;\; R_{us_n,s_i} > R_{us_n,s_j}, \tag{3.36}$$

where R_{us_n,s_i} is the rate achieved by BE users belonging to the set us_n on subband s_i, i.e.,

$$R_{us_n,s_i} = \sum_{k_{BE} \in us_n} R_{k_{BE},s_i}. \tag{3.37}$$

For subbands, the utility achieved by scheduling us_n is given by:

$$\mathcal{U}^t_{BE}(us_n, s_i) = \sum_{k \in us_n} \mathcal{U}^t_{BE}(k, s_i), \tag{3.38}$$

where $\mathcal{U}^t_{BE}(k, s_i)$ is given by Eq. (3.31).

Subbands then base their preferences on the following definition:

Definition 25. *Subband $s_i \in \mathcal{S}$ bases its preference list over the set of users $\mathcal{US}$ according to:*

$$us_n \succ_{s_i} us_m \quad \textit{if} \quad \mathcal{U}^t_{BE}(us_n, s_i) > \mathcal{U}^t_{BE}(us_m, s_i). \tag{3.39}$$

Phases 1 and 2 of Algorithm 7 describe the steps involved in the first stage of the resource allocation technique in the NOMA-CAS case.

In the first phase, RT users are scheduled in an OMA manner using the CAS version of Algorithm 5. Phase 1 is concluded when either the preference lists of all active RT users are empty, or the set of active RT users is empty, i.e., all RT users received their required rates.

If free subbands remain, Phase 2 of Algorithm 7 schedules BE users in a NOMA manner over the set of remaining subbands, $\mathcal{S}_{\text{rem}}$. First, the user set $\mathcal{US}$ is constructed

and each user set $us_n \in \mathcal{US}$ builds its preference list, $\mathcal{PL}_{us_n}$ over the set of available subbands. Each subband $s_i \in \mathcal{S}_{\text{rem}}$ initializes its matching set $\mathcal{M}(s_i)$.

In Phase 2.1 of Algorithm 7, user sets apply to their preferred subbands. Each subband s_i receiving proposals adds both the proposing user sets as well as the previously matched user set, $\mathcal{M}(s_i)$, to its applying user set $AS(s_i)$. Then, in phase 2.2, each subband receiving user proposals accepts the user set maximizing its utility, as per step 28 of Algorithm 7. Each user set then removes the subband that it proposed to from its preference list, as stated in Phase 2.3 of the algorithm.

Phases 2.1, 2.2 and 2.3 are repeated until the preference lists of all user sets are empty. Upon termination of Phase 2, the hybrid-NOMA assignment of BE users to the remaining subbands is obtained.

3.4.3.2 NOMA Pairing on Subbands Assigned to RT Users

As in the greedy solution (Section 3.3.2), the first stage of the proposed matching solution in the NOMA-CAS context may end up with some scheduled RT users exceeding their required rates, with possibly other RT users not satisfying their requirements. Consequently, in the second stage of the allocation process, NOMA pairing is performed on the subbands assigned to RT users in the first stage.

This second stage starts by finding the amount of power that satisfied RT users can spare, without jeopardizing their satisfaction, as was done in Section 3.3.2. Therefore, for each RT user $k_{RT} \in \mathcal{K}_{RT}$ having exceeded its required rate, the optimization problem in (3.18) is solved. As a result, the amount of power needed by user k_{RT} on each assigned subband $s \in \mathcal{S}_{RT}$ are computed using Eq. (3.19). Moreover, the required rate of user k_{RT} on subband $s \in \mathcal{S}_{k_{RT}}$ is given by:

$$R^{\text{req}}_{k_{RT},s} = B_c \log_2 \left(1 + \frac{p_{k_{RT},s} h^2_{k_{RT},s}}{N_0 B_c}\right). \tag{3.40}$$

This rate must be guaranteed regardless of the NOMA pairing order on subband s. Note that if the rate $R_{k_{RT},s}$ obtained on subband s at the end of the first allocation stage is lower than $R^{\text{req}}_{k_{RT},s}$, subband s is removed from $\mathcal{S}_{RT}$, the set of subbands available for NOMA pairing (given by step 12 of Algorithm 7). In the opposite case, subband s can be shared with another user. To guarantee $R^{\text{req}}_{k_{RT},s}$, two separate cases are considered concerning the channel gain of the candidate user for pairing, user k'.

3.4.3.2.1 Case 1: $h_{k_{RT},s} > h_{k',s}$

In this case, the already scheduled user k_{RT} is paired as first user on subband s with a required rate given by Eq. (3.40). However, to guarantee SIC stability, $p_{k_{RT},s} < p_{k',s}$ must hold. This translates into considering subband s for NOMA pairing if $p_{k_{RT},s} < P_s/2$. If this condition is verified, the available power for user k' on s is given by:

$$p^{2,\text{av}}_s = P_s - p_{k_{RT},s}. \tag{3.41}$$

3.4.3.2.2 Case 2: $h_{k_{RT},s} < h_{k',s}$

In this case, user k_{RT} is paired as second user on subband s. To guarantee its required rate $R^{\text{req}}_{k_{req},s}$, user k_{RT} needs to be assigned a power given by Eq. (3.25). As in Section 3.3.2,

Algorithm 7 Priority-Aware NOMA DA Algorithm

Input: $\mathcal{K}_{RT}, \mathcal{K}_{BE}, \boldsymbol{H_{RT}}, \boldsymbol{H_{BE}}, \boldsymbol{R}^{\text{req}}_{\boldsymbol{RT}}, \boldsymbol{L_{RT}}, \boldsymbol{T_{BE}}, t$.
Output: $\boldsymbol{A_{RT}}, \boldsymbol{A_{BE}}, \boldsymbol{R_{RT}}, \boldsymbol{R_{BE}}$.

Initialization:

if CAS setting **then**
  $\mathcal{K}_{\text{active}} = \left\{k_{RT} \in \mathcal{K}_{RT} / R^{\text{req}}_{k_{RT}} > 0\right\}$.
else if DAS setting **then**
  Construct the virtual RT user set $\mathcal{K}^{v}_{RT}$.
  $\mathcal{K}^{r}_{\text{active}} = \left\{k_{RT} \in \mathcal{K}_{RT}, / R^{\text{req}}_{k_{RT}} > 0\right\}$, $\mathcal{K}_{\text{active}} = \left\{v^{k_{RT},a}, a = 1, \ldots, A \ / R^{\text{req}}_{k_{RT}} > 0\right\}$.
  $P_a = P_{DAS} \,/\, N_a, \ \forall\, a \in \mathcal{A}$.
end if
Build the preference lists $\mathcal{PL}_k$ of users in $\mathcal{K}_{\text{active}}$.
Set $\mathcal{M}(s_i) = \emptyset, \ \forall s_i \in \mathcal{S}$.

Repeat steps 10 to 40 in the DAS setting

$N^{\text{old}} = N$.

Phase 1: Scheduling RT users

Perform phases 1, 2 and 3 from Algorithm 5.

Repeat Phase 1 until $\mathcal{PL}_k = \emptyset, \ \forall\, k \in \mathcal{K}_{\text{active}}$ or until $\mathcal{K}_{\text{active}} = \emptyset$

$\mathcal{S}_{RT} \leftarrow$ Subbands used by RT users, $\mathcal{S}_{\text{rem}} \leftarrow \mathcal{S} \setminus \mathcal{S}_{RT}$.

Phase 2: Scheduling BE users

if $\mathcal{S}_{\text{rem}} \neq \emptyset$ **then**

Initialization:

  if CAS setting **then**
    Construct user set $\mathcal{US}$ consisting of both single users and user pairs.
  else if DAS setting **then**
    Construct the virtual user set $\mathcal{US}$ consisting of both virtual users and virtual user pairs.
  end if
  Build the preference lists $\mathcal{PL}_{us_n}$ of user sets $us_n \in \mathcal{US}$.
  Set $\mathcal{M}(s_i) = \emptyset, \ \forall s_i \in \mathcal{S}_{\text{rem}}$.

Phase 2.1: BE users and BE pairs apply to subbands

  if CAS setting **then**
    Each user set $us_n \in \mathcal{US}$ proposes to the first subband in $\mathcal{PL}_{us_n}$.
  else if DAS setting **then**
    Each real user chooses its proposing virtual user $v_p^{us_n,a}$ according to Eq. (3.33), while each real user pair chooses its proposing virtual user pair $v_p^{us_n,a}$ according to Eq. (3.45). The real user (resp. real user pair) then proposes to the first subband in $\mathcal{PL}_{v_p^{us_n,a}}$.
  end if
  Construct the applicant set $\mathcal{AS}(s_i)$ for each subband $s_i \in \mathcal{S}_{\text{rem}}$, $\mathcal{AS}(s_i) = \mathcal{AS}(s_i) \cup \mathcal{M}(s_i), \ \forall\, s_i \in \mathcal{S}_{\text{rem}}$.

Phase 2.2: Subbands make decisions

  if $\mathcal{AS}(s_i) \neq \emptyset$ **then**
    $\mathcal{M}(s_i) = \{us_n^*\}$, where $us_n^* = \underset{us_n \in \mathcal{A}(s_i)}{\text{argmax}}\ \mathcal{U}^t_{BE}(us_n, s_i)$.
  end if

Phase 2.3: Preference lists update

Each user set us_n (respectively virtual user set $v_p^{us_n,a}$) that proposed to subband s_i, $\forall\, s_i \in \mathcal{S}_{\text{rem}}$, removes s_i from $\mathcal{PL}_{us_n}$ (respectively $\mathcal{PL}_{v_p^{us_n,a}}$).

Repeat Phases 2.1, 2.2 and 2.3 until $\mathcal{PL}_{us_n} = \emptyset,\ \forall\ us_n \in \mathcal{US}$

end if

Using the resulting $\boldsymbol{A_{RT}}$ and $\boldsymbol{A_{BE}}$, re-calculate $\boldsymbol{N} \in \mathbb{N}^{A\times 1}$ and $\boldsymbol{P} \in \mathbb{R}_+^{A\times 1}$ as well as $\boldsymbol{R_{RT}}$ and $\boldsymbol{R_{BE}}$.

Phase 3: NOMA pairing on subbands assigned to RT users

if $\mathcal{S}_{RT} \neq \emptyset$ **then**

Find RT users that exceed their required rates $\mathcal{K}_{\text{RT}}^{\text{excess}}$.

Solve problem (3.18) for $k \in \mathcal{K}_{\text{RT}}^{\text{excess}}$.

Find the available power on each subband $s \in \mathcal{S}_{RT}$.

$\mathcal{K}_{\text{active}} = \{\mathcal{K}_{BE} \cup k_{RT} \in \mathcal{K}_{RT} \,/\, R_{k_{RT}}^{\text{req}} > 0\}$.

Build preference lists for users in $\mathcal{K}_{\text{active}}$.

Use Algorithm 5 to schedule additional users on subbands belonging to $\mathcal{S}_{RT}$.

end if

Until convergence

SIC stability is guaranteed if $p_{k_{RT},s}^{2,\text{req}} > P_s/2$. Hence, if being scheduled as second user, the required rate of user k_{RT} to ensure SIC stability is chosen to be $R_{k_{RT},s}^{\text{req}} = \max\left(R_{k_{RT},s}^{\text{req}}, B_c\right)$. In the case where $R_{k_{RT},s}^{\text{req}}$ takes the value B_c, the value of $p_{k_{RT},s}^{2,\text{req}}$ is recalculated to reflect the rate change. Then, the power available for a potential first user k' can be given by:

$$p_s^{1,\text{av}} = P_s - p_{k_{RT},s}^{2,\text{req}}. \tag{3.42}$$

Having determined the amounts of power available for NOMA pairing, the matching algorithm proposed for the second allocation stage can now be described.

The active users in this second stage are the unsatisfied RT users and the BE users, hence:

$$\mathcal{K}_{\text{active}} = \{\mathcal{K}_{BE} \cup k_{RT} \in \mathcal{K}_{RT} \,/\, R_{k_{RT}}^{\text{req}} > 0\}. \tag{3.43}$$

Let k_{RT} be the RT user already scheduled on subband $s \in \mathcal{S}_{\text{RT}}$ considered for NOMA pairing. First, the achievable rate of each candidate user $k \in \mathcal{K}_{\text{active}}$ over each subband $s \in \mathcal{S}_{\text{RT}}$ is calculated:

$$R_{k,s}^{\text{av}} = \begin{cases} B_c \log_2\left(1 + \frac{p_s^{1,\text{av}} h_{k,s}^2}{N_0 B_c}\right), \text{ if } h_{k,s} > h_{k_{RT},s}, \\ B_c \log_2\left(1 + \frac{p_s^{2,\text{av}} h_{k,s}^2}{p_{k_{RT},s} h_{k,s}^2 + N_0 B_c}\right), \text{ otherwise.} \end{cases} \tag{3.44}$$

Then, each active user builds its preference list over the set of available subbands according to the decreasing order of achievable rates.

Since only one additional user is to be scheduled on each subband allocated to RT users, Algorithm 5 can be used to solve the second stage of the allocation process. However, instead of considering all subbands in set $\mathcal{S}$ as done in Algorithm 5, only those assigned to RT users in the previous stage must be considered (i.e., subbands in the set $\mathcal{S}_{RT}$).

The complete algorithm used to solve the allocation problem in the NOMA-CAS setting is given in Algorithm 7.

3.4.4 Matching Technique in the NOMA-DAS Context

Resource allocation through matching theory in the NOMA-DAS case is conducted by incorporating the concepts from Sections 3.4.2 and 3.4.3.

As in the NOMA-CAS case, the subband and antenna assignment in the NOMA-DAS system is divided into two stages, the first one aiming at satisfying RT users and the second at enhancing system performance.

To solve the resource allocation problem using matching theory, the set of virtual active RT users is first found. As a DAS system is considered, the available power per antenna is equally divided between subbands, using the estimated number of subbands per antenna $N_a, \forall\, a \in \mathcal{A}$, as per step 6 of the algorithm. Active users then construct their preference lists over the set of subbands, and the latter initialize their matching set.

During the first phase of the algorithm, RT users are scheduled using OMA on their preferred subbands. As OMA scheduling is performed, RT users are assigned to their preferred subbands using the DAS version of Algorithm 5. Phase 1 concludes when either the preference lists of all active RT users are empty, or all RT users receive their required rates. Upon termination of Phase 1, the set of subbands assigned to RT users, $\mathcal{S}_{RT}$, in addition to the set of available subbands, $\mathcal{S}_{\text{rem}}$, are found.

If the set of available subbands $\mathcal{S}_{\text{rem}}$ is not empty, BE users are directly scheduled using NOMA on these subbands in Phase 2 of the algorithm. Since a DAS system is considered, the virtual user set $\mathcal{US}$ is first constructed in step 17. Set $\mathcal{US}$ consists of virtual users and virtual user pairs, where each virtual user (respectively each virtual user pair) represents the association of a real user (respectively of a real user pair) with a serving antenna. The preference lists of all virtual user sets are then constructed according to Definition 24, and the matching set of each subband $s_i \in \mathcal{S}_{\text{rem}}$ is initialized. In Phase 2.1, each real user (respectively each real user pair) chooses its proposing virtual user (respectively its proposing virtual user pair). While each real user chooses its proposing virtual user according to Eq. (3.33), the latter decision metric must be modified when the real user set consists of a pair of BE users. In fact, due to NOMA, the co-channel interference, experienced among users scheduled on the same channel, must be accounted for in the decision metric. To that effect, each real user pair chooses its proposing virtual user pair according to the following metric:

$$v_p^{us_n,a} = \underset{a' \in \mathcal{A}}{\text{argmax}} \sum_{k_{BE} \in us_n} R_{k_{BE}, s^*_{v^{us_n,a'}}}, \tag{3.45}$$

where $s^*_{v^{us_n,a'}}$ is the very first element of the preference list of virtual user set $v^{us_n,a'}$, i.e., of the virtual user pair representing the association of real user pair us_n and antenna a'. The proposing virtual user set then proposes to its preferred subband. In Phase 2.2, subbands accept their preferred users sets and reject all other proposing user sets. In Phase 2.3, the preference lists of all proposing virtual user sets are updated.

Phases 2.1, 2.2 and 2.3 are then repeated until the preference lists of all virtual user sets are empty. Upon termination, the assignment of NOMA BE users to subbands and antennas is obtained. Both the number of subbands per antenna, N_a, for every antenna $a \in \mathcal{A}$, and the power per subband on each antenna, P_a, for every antenna $a \in \mathcal{A}$, are updated, as per step 32 of the algorithm. This assignment is kept unchanged during the NOMA pairing stage of the algorithm, i.e., during Phase 3.

During Phase 3 of the algorithm, NOMA pairing on subbands assigned to OMA RT users is conducted as in Section 3.4.3.

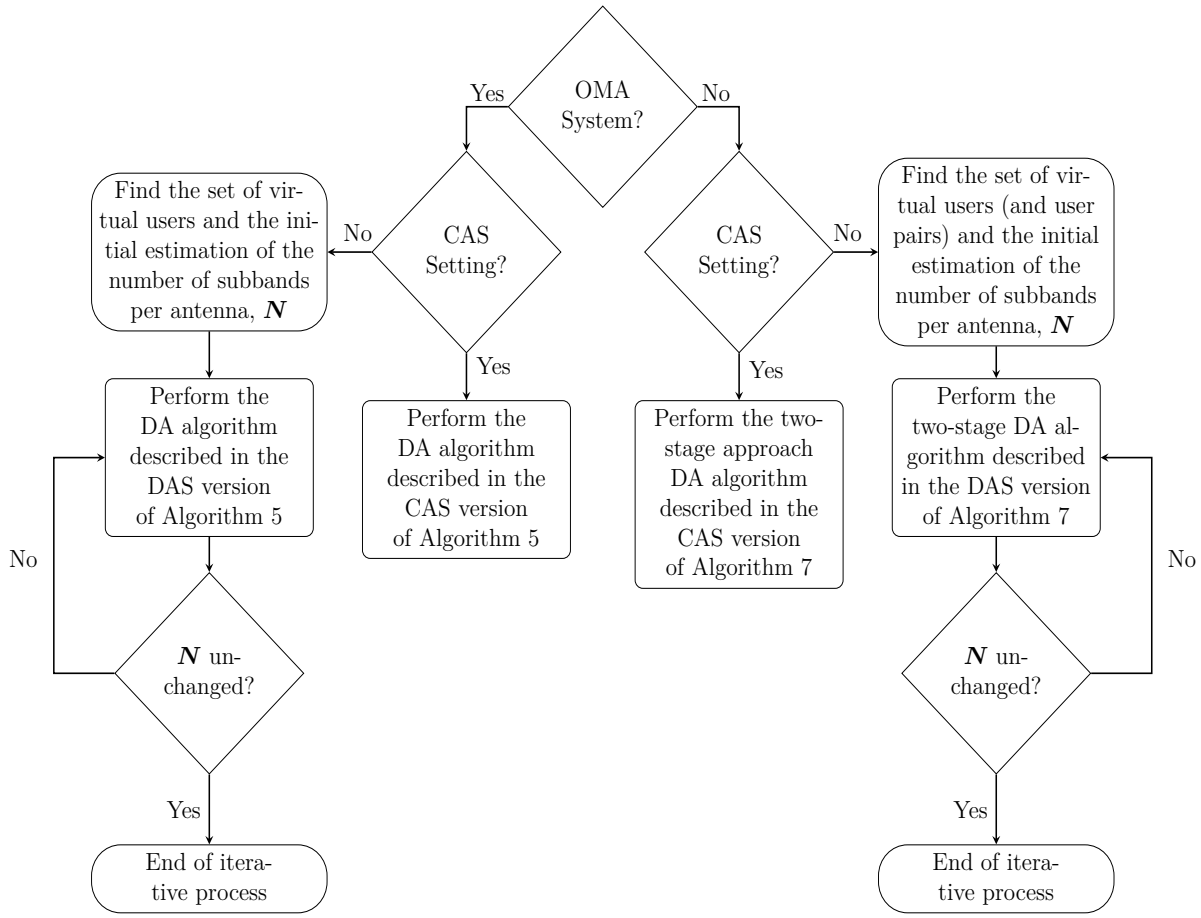


Figure 3.6 – Flowchart of the proposed matching technique in the studied settings

Phases 1, 2 and 3 are then iterated until convergence of the number of subbands per antenna, i.e., until, the number of subbands per antenna, between two consecutive iterations, becomes unchanged.

The proposed method is detailed in the DAS version of Algorithm 7.

To summarize, the flowchart in Fig. 3.6 shows how the matching technique is used in the different considered system settings.

3.4.5 Analysis of Stability, Convergence and Complexity

To analyze the properties of the proposed matching technique in all studied system settings (i.e., in OMA-CAS, OMA-DAS, NOMA-CAS, NOMA-DAS), we separately consider the following two parts:

1. the matching algorithm,

2. the iterative approach proposed to find the number of subbands per antenna in the DAS case.

3.4.5.1 Properties of the Matching Technique

3.4.5.1.1 Stability Analysis

Before discussing the stability property of the matching technique, the definition of *blocking pair* [17] is first recalled.

Definition 26. *Given a matching Ψ and a pair (k_n, s_i), where $k_n \in \mathcal{K}$ in the OMA case whereas $k_n \in \mathcal{US} \cup \mathcal{K}_{RT}$ in the NOMA case, and $s_i \in \mathcal{S}$, let $k_n \notin \Psi(s_i)$ and $s_i \notin \Psi(k_n)$. The pair (k_n, s_i) forms a blocking pair if:*

1. $k_n \succ_{s_i} \Psi(s_i)$,
2. $s_i \succ_{k_n} s_j$, *where* $s_j \in \Psi(k_n)$.

In other words, a pair (k_n, s_i) forms a *blocking pair* if k_n and s_i are not matched together in matching Ψ, while both k_n and s_i prefer each other over their respective matching.

With the definition above, it is now possible to define the concept of stability and prove that the proposed matching technique does indeed converge towards a stable matching.

Definition 27. *If there is no blocking pair $(k_n, s_i) \in \Psi$, matching Ψ is considered* stable.

Theorem 2. *The proposed matching technique is guaranteed to converge to a stable matching Ψ^* in all system settings.*

Proof. Suppose that there exists a pair $(k_n, s_i) \notin \Psi^*$ such that (k_n, s_i) forms a blocking pair. According to both Algorithms 5 and 7, user k_n has already proposed to subband s_i and was rejected at some iteration q, meaning that $\Psi^q(s_i) \succ_{s_i} k_n$, where Ψ^q is the matching state at iteration q. Since the optimal matching satisfies $\Psi^*(s_i) \succ_{s_i} \Psi^q(s_i)$, subband s_i is matched to its final partner $\Psi^*(s_i)$ which it prefers to k_n. Hence, (k_n, s_i) does not form a blocking pair and the matching Ψ^* is stable. It should be noted that a stable matching is guaranteed to exist since the problem is modeled as a one-to-many matching game [17]. □

3.4.5.1.2 Convergence Analysis

Theorem 3. *The matching technique is guaranteed to converge after a limited number of iterations for all studied system settings.*

Proof. See Appendix A.1. □

3.4.5.1.3 Complexity Analysis

Theorem 4. *The maximum complexity of the proposed matching technique, achieved in the NOMA-DAS setting, is $\mathcal{O}\Big((K_{RT} + |\mathcal{US}|)AS^2\Big)$.*

Proof. See Appendix A.2. □

3.4.5.2 Properties of the Iterative Approach

3.4.5.2.1 Convergence Analysis

Theorem 5. *The iterative approach introduced in Algorithm 6 is guaranteed to converge within a limited number of iterations.*

Proof. The number of subbands and antennas of the considered system is limited. Hence, the number of potential assignments of subbands to antennas is finite. Furthermore, system performance is evaluated in terms of RT users satisfaction and BE users sum rate. Since RT users and BE users are competing for the same resources, it can be shown that, after each iteration, the performance of at least one user type is enhanced. In other words, it can be shown that either the performance of RT users or that of BE users is enhanced after each iteration. Moreover, the exchanges between the performance enhancement of the two user types tends to stabilize after a small number of iterations. Since system performance, i.e., the number of satisfied RT users and the sum rate of BE users, has an upper bound because of the limited spectral resources, the iterative approach terminates when this upper bound is reached. Therefore, the number of subbands per antenna can be found in a limited number of iterations. □

3.4.5.2.2 Complexity Analysis

Concerning the complexity of Algorithm 6, the number of iterations cannot be given in a closed form expression because it is not certain at which round the algorithm converges to the final solution. However, an upper bound on the complexity can be derived. Since system performance increases after each iteration, if ΔPG denotes the performance gain yielded by the iterative approach and $\delta_{\min}$ is the minimum increase in performance at each iteration, an upper bound on the complexity of this method is given by $\mathcal{O}(\frac{\Delta PG}{\delta_{min}})$.

3.4.6 Numerical Results

The performance of the proposed subband assignment technique based on matching theory, denoted by "MM", is evaluated in this section through numerical simulations. The performance of MM is tested in the OMA-CAS, OMA-DAS, NOMA-CAS and NOMA-DAS settings. A variation of the MM method, denoted by MM-FA, in the DAS settings is also tested. MM-FA adopts the approach of [23] and determines the number of subbands per antenna, $N_a, \forall\, a \in \mathcal{A}$, at the beginning of the resource allocation algorithm based on the average path-loss experienced by all users. For fair comparison, each antenna a is assigned the N_a subbands having the highest average channel gain for all users. The performance achieved by the low-complexity greedy method presented in Section 3.3 is denoted by "GM".

The parameters used in the simulations are summarized in Table 3.3 [103]. When CAS is considered, one antenna is assumed to be located at the center of the cell. In the case of DAS, $A = 4$ RRHs are assumed to be deployed in the cell. In addition to the centrally located antenna, the other 3 antennas are equally distanced and positioned on a circle of radius $2R_d/3$, R_d being the cell radius, with an angular separation of 120°. The power budget per cell is 40 W; when DAS is considered, this budget is equally partitioned between antennas leading to 10 W per antenna. The number of RT users K_{RT}, is varied between 5 and 30 to depict different system congestion levels. To reflect different RT

application requirements, RT users are partitioned among 3 classes (C1, C2 and C3). Users in all 3 classes request $D^{\text{req}} = 10^5$ bits. However, a user $k_{RT} \in$ C1 has a latency limit $L_{k_{RT}} = 6$ timeslots, whereas if $k_{RT} \in$ C2 (resp. $k_{RT} \in$ C3), $L_{k_{RT}} = 10$ timeslots (resp. $L_{k_{RT}} = 15$ timeslots). For all K_{RT} values, it is assumed that 20% of the users belong to each of classes C1 and C2 while the remaining 60% users belong to class C3. The number of BE users K_{BE} is maintained at 20 throughout the simulations. Moreover, each simulation is run for 20 timeslots, allowing all RT users enough time to reach their latency limits (since the largest latency limit is equal to 15 timeslots). During the last 5 timeslots, BE users compete for system resources among each other.

Table 3.3 – Simulation Parameters

Cell Radius R_d	500 m
Overall Transmission Bandwidth	10 MHz
Number of subbands	16
Cell Power Budget	40 Watts (46 dBm)
Number of RT users in the cell	5, 10, 15, 20, 25, 30
Number of BE users in the cell	20
Distance Dependent Path Loss	$128.1 + 37.6 \log_{10}(d)$(dB), d in Km
Receiver Noise Density	4.10^{-18} mW/Hz

3.4.6.1 Convergence of the Proposed Method

First, the convergence of the iterative method that aims at finding the number of subbands per antenna is observed. Fig. 3.7a plots the cumulative distribution function (CDF) of the number of iterations needed for Algorithm 6 to converge for both the OMA and NOMA settings, when the number of RT users takes the values 10 and 20. The CDF shows that the proposed method converges within a small number of iterations (90% of the cases converge within 8 iterations) for both the OMA and the NOMA settings, as well as for both $K_{RT} = 10$ and $K_{RT} = 20$ users. The convergence of the NOMA settings is slightly slower than its OMA counterpart. Moreover, Fig. 3.7a shows that the convergence of the setting with $K_{RT} = 10$ users is slower than the one with 20 RT users. In fact, when $K_{RT} = 10$, BE users have a higher chance of being scheduled. Hence, all 30 users (10 RT users and 20 BE users) contribute to deciding on the assignment of subbands to the RRHs. However, when $K_{RT} = 20$, the system is more congested and BE users have a harder time getting resources. Therefore, the assignment of subbands to the distributed antennas is mostly decided upon by the $K_{RT} = 20$ RT users in this case, explaining the faster convergence.

In Fig. 3.7b, the convergence of the MM technique is shown for the different settings, when $K_{RT} = 10$ users. As expected, the OMA settings converge faster than the NOMA ones. However, it can be seen that the maximum number of iterations needed for MM to converge is $AS = 64$ in the DAS settings, which is a relatively small number of iterations. Moreover, Fig. 3.7b validates the statement of Theorem 3.

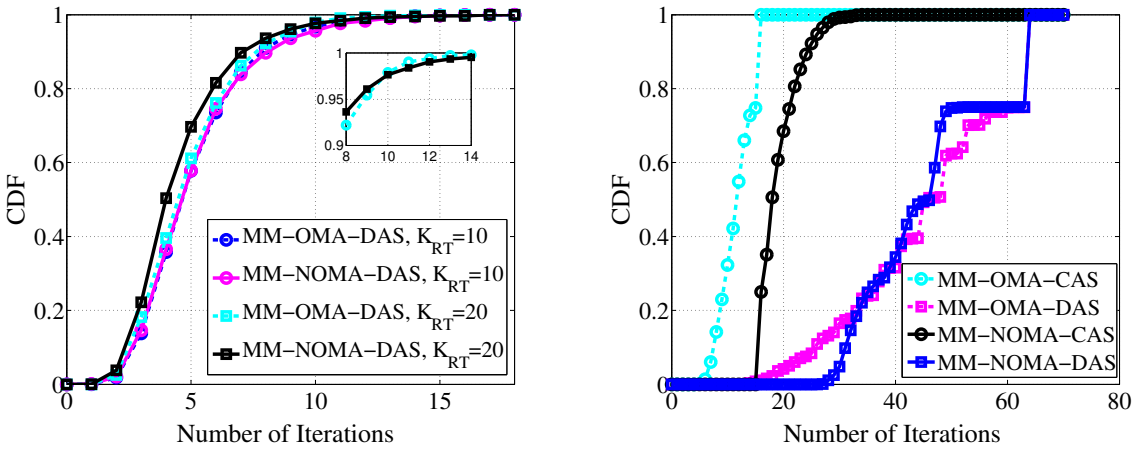


Figure 3.7 – CDF of the number of iterations needed: (a) to find the number of subbands per antenna, (b) for the matching method with $K_{RT} = 10$

3.4.6.2 Performance of the MM Technique

3.4.6.2.1 Evaluation of the Performance of the MM Technique for RT Users

In Fig. 3.8, the performance of the proposed technique in terms of RT users satisfaction is evaluated. It can be noted that, until $K_{RT} = 15$ users, MM and GM perform similarly regardless of the considered scenario. However, as the cell becomes more congested with a larger number of RT users, MM outperforms GM in all its variations. More concretely, when $K_{RT} = 30$, GM achieves almost no satisfaction for RT users. However, MM-OMA-CAS (resp. MM-OMA-DAS) outperforms its GM equivalent by almost 28% (resp. 62 %). Also, in the NOMA cases, MM-NOMA-CAS (resp. MM-NOMA-DAS) outperforms its GM equivalent by almost 30% (resp. 63%). Fig. 3.8 also shows the gains achieved by the iterative method to find the number of subbands per antenna introduced in Algorithm 6. For example, when $K_{RT} = 30$ users, MM-OMA-DAS (resp. MM-NOMA-DAS) outperforms MM-FA-OMA-DAS (resp. MM-FA-NOMA-DAS) by almost 30% (resp. 26%). Additionally, the results show the gain achieved by using a DAS setting, in comparison to CASs, as it can increase the performance by more than 30%.

Having observed the performance enhancement brought by the DAS settings, hereinafter, only the performance of MM-NOMA-CAS will be compared to the other methods as it has the best behavior in the CAS setting. Also, MM-FA-OMA-DAS will be dropped since it is outperformed by its NOMA version.

In Fig. 3.9, the percentage of satisfied RT users per class is shown. As previously stated, GM and MM both satisfy all RT users when $K_{RT} \leq 15$ users. However, when $K_{RT} \geq 20$, the performance of GM degrades and the satisfaction of RT users belonging to the strictest class, C1 (with a latency of 6 timeslots), is mostly affected, as shown in Fig. 3.9a. In fact, for $K_{RT} = 25$ for example, the GM technique achieves close to 10% satisfaction for users belonging to class C1 in the case of a DAS system, while a CAS cannot satisfy any user in C1. On the other hand, MM-OMA-DAS and MM-NOMA-DAS achieve both almost 96% satisfaction while the CAS versions achieve almost 70% satisfaction. For the more relaxed classes, as shown in Fig. 3.9b and Fig. 3.9c,

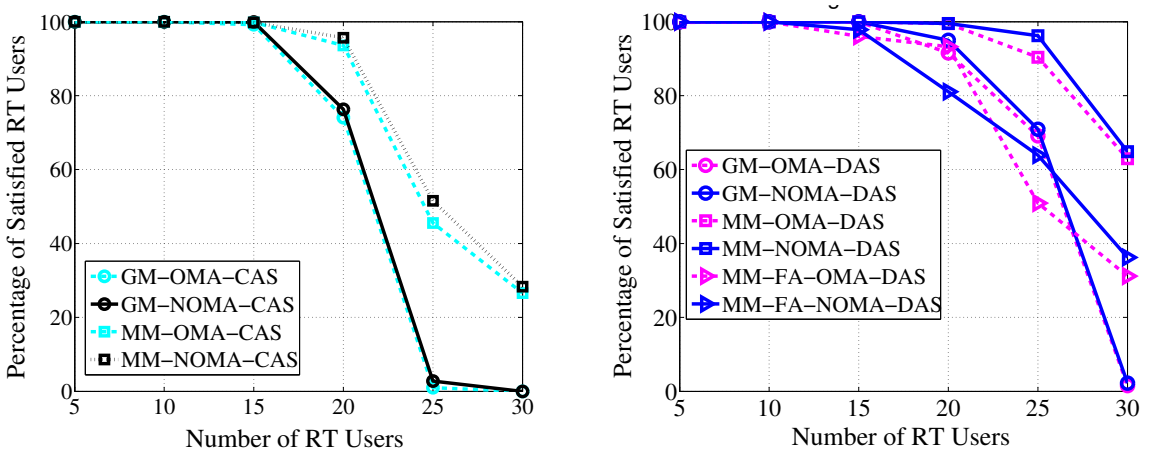


Figure 3.8 – Percentage of satisfied RT users for: (a) the CAS settings, (b) and the DAS settings

the percentage of satisfied RT users with the GM method is higher than for class C1, since their requirements are more relaxed. Hence, even if the GM algorithm achieves an acceptable global percentage of satisfaction for $K_{RT} = 20$ and 25, this performance results from the satisfaction of the users in the most relaxed classes. This is not the case for the MM algorithm which prioritizes users having more strict requirements. Therefore, it achieves an acceptable performance for all classes.

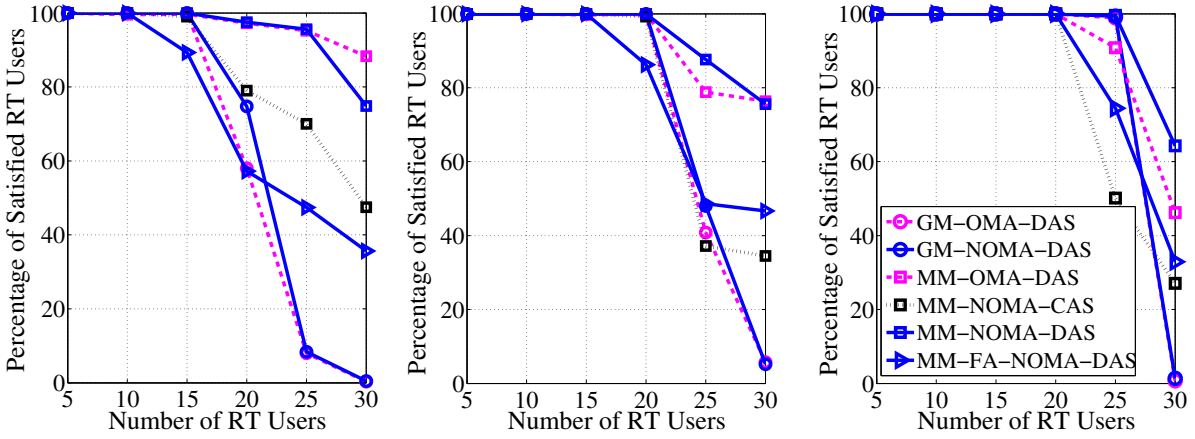
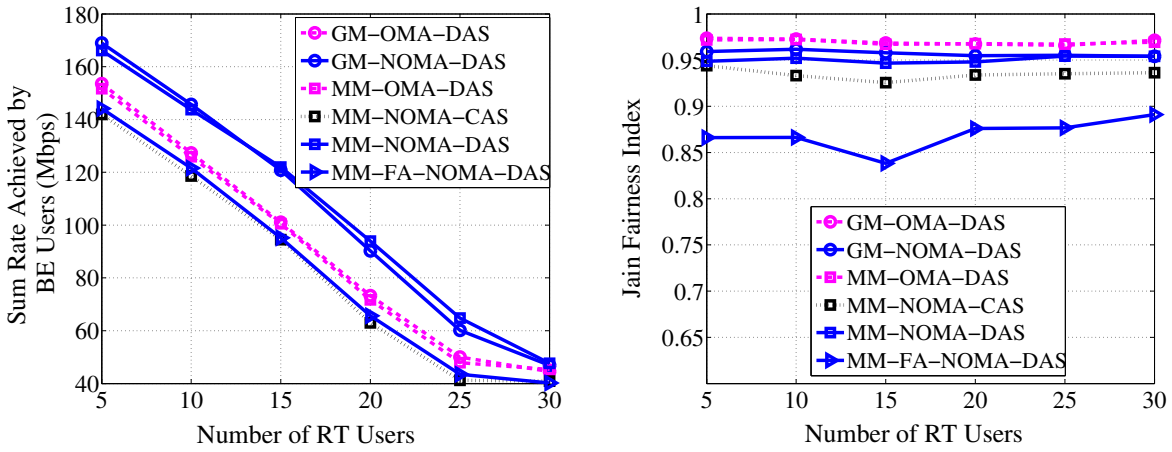


Figure 3.9 – Percentage of satisfied RT users in: (a) class C1, (b) class C2, (c) class C3

Remark 2. *In Fig. 3.9a, for class C1, we notice that MM-OMA-DAS outperforms MM-NOMA-DAS for a system consisting of 30 RT users. This results from the fact that at timeslot t, for NOMA pairing, all additional power allocated to a satisfied RT user k_{RT} is taken away from it to accommodate more users via NOMA. Hence, the rate achieved by k_{RT} at timeslot t is exactly equal to its required rate. At a subsequent timeslot t′*

($t < t' \leq L_{k_{RT}}$), user k_{RT} might not be able to reach its required rate (for example, because of a bad channel state). However, if k_{RT} was allowed to exceed its required rate at timeslot t, the required rate at t' would be reduced and hence user k_{RT} could have been satisfied. A solution for this problem might be to allow RT users to keep a small amount of additional power during the NOMA pairing step.

3.4.6.2.2 Evaluation of the Performance of the MM Technique for BE Users



(a) (b)

Figure 3.10 – (a) Achieved rate for BE users as K_{RT} increases, (b) System Fairness in terms of K_{RT}

Fig. 3.10a shows the achieved rate of BE users as the number of RT users K_{RT} increases. As expected, the sum rate of BE users decreases for all methods as K_{RT} grows, as less resources are available for BE users. Both methods (GM and MM) perform similarly when it comes to the sum rate achieved by BE users. For example, for the NOMA-DAS case, GM-NOMA-DAS achieves almost 1 Mbps gain over MM-NOMA-DAS when $K_{RT} = 5$ or 10 users. However, for $K_{RT} = 20$ or 25 users, MM-NOMA-DAS achieves almost 3 Mbps gain over GM-NOMA-DAS. Therefore, while significantly enhancing the satisfaction of RT users, MM does not degrade the sum rate of BE users. Moreover, MM-NOMA-DAS greatly outperforms MM-FA-NOMA-DAS. Hence, the complexity added by the use of Algorithm 6 is justified by the enhanced performance for both RT and BE users. In Fig. 3.10b, we show the fairness achieved by the different methods in terms of K_{RT}. System fairness is assessed through Jain's fairness index [24], given in Eq. (3.28). It can be seen that MM-NOMA-DAS outperforms its FA counterpart. Putting MM-FA apart, Fig. 3.10b shows that all considered methods have a good performance in terms of fairness with a Jain index higher than 0.9, with an advantage for the DAS settings. Therefore, system fairness is not a deciding factor in the evaluation of the different methods. It should be noted that although the OMA versions slightly outperform the NOMA ones in terms of system fairness, NOMA is able to increase the minimum individual rate of BE users with respect to OMA, as shown in Fig. 3.11. Hence, the slightly decreased fairness is due to some users having slightly more rate that others in the NOMA setting.

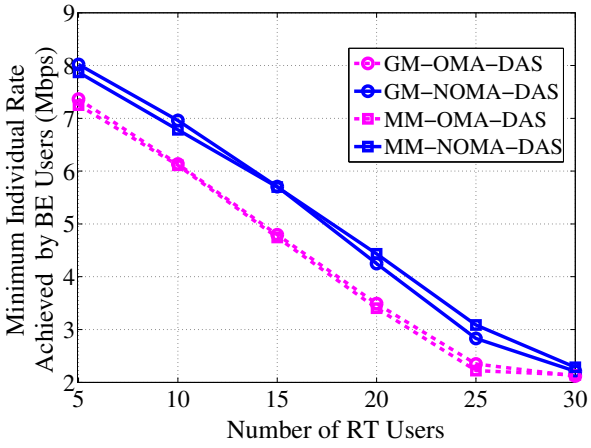


Figure 3.11 – Minimum individual achieved rate by BE users in terms of K_{RT}

3.4.6.2.3 Evaluation of the Proposed Metrics in (3.30) and (3.31)

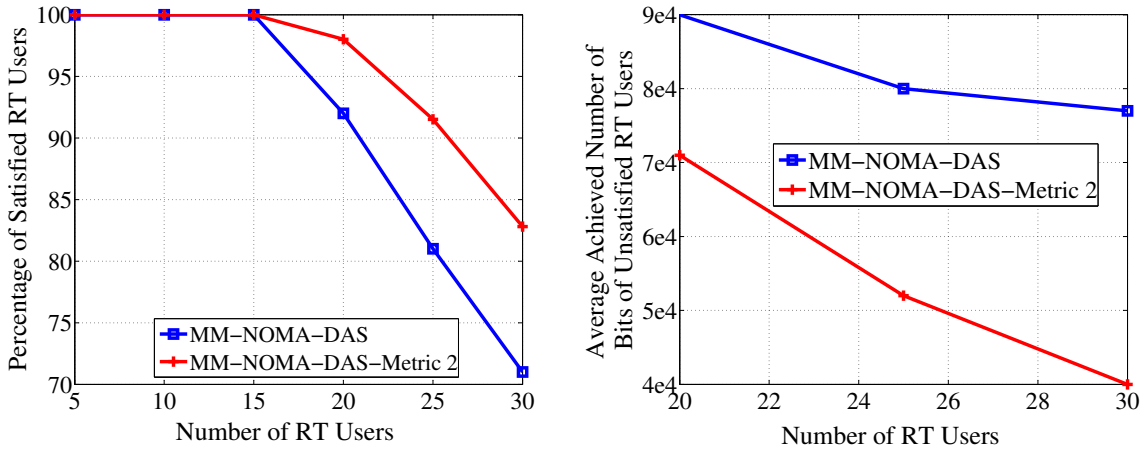
To show the adequacy of the proposed metrics in Eq. (3.30) and (3.31), the performance achieved when using these metrics is compared against the performance obtained when the unified metric proposed in [83] is used. The unified metric is given by:

$$\mathcal{UM}_{k,s}(t) = \gamma_k R^t_{k,s} \exp(\theta_k(t)), \tag{3.46}$$

where $\gamma_k = 1$ for BE users and $\gamma_k = (1 + t)$ for RT users. Variable $\theta_k(t)$ denotes the normalized delay. To compare the performance of the proposed metric and that of the unified metric, a NOMA-DAS setting is considered and Algorithm 7 is employed. However, the preference relation of the subbands is modified depending on the tested metric. The performance of the unified metric is denoted by MM-NOMA-DAS-Metric 2 in the simulation results. As we were interested in comparing the performance of the two metrics for systems with RT and BE users, each simulation is run for 15 timeslots. During these 15 timeslots, both RT and BE users compete for system resources.

Fig. 3.12 compares the performance of the two metrics for RT users. In terms of the percentage of satisfied RT users, Fig. 3.12a shows that both metrics achieve a similar performance for $K_{RT} \leq 15$ users. As K_{RT} increases, the unified metric in (3.46) outperforms our proposed metric in Eq. (3.30) by up to 13% for $K_{RT} = 30$. That is because metric (3.46) privileges users with a high rate. In contrast, our proposed metric (3.30) seeks to achieve a high fairness between RT users by accounting for the received rate before timeslot t. In other words, even if it were unable to satisfy all users, our proposed metric (3.30) aims at approaching most RT users to their requirements, as validated by Fig. 3.12b. Indeed, Fig. 3.12b shows that our proposed metric (3.30) increases the amount of received data bits of unsatisfied RT users by up to 3.6×10^4 bits, when compared to the unified metric (3.46).

Fig. 3.13 compares the performance of both metrics for BE users. As expected, the unified metric (3.46) achieves a higher sum rate for BE users than our proposed metric (3.31). This is due to the fact that in the unified metric (3.46), only the achievable rate of BE users is taken into account, without any fairness consideration. This is not the case of our proposed metric (3.31), which seeks to achieve a tradeoff between rate and fairness maximization. The superior performance of our proposed metric (3.31) in terms of fairness is shown in Fig. 3.13b.



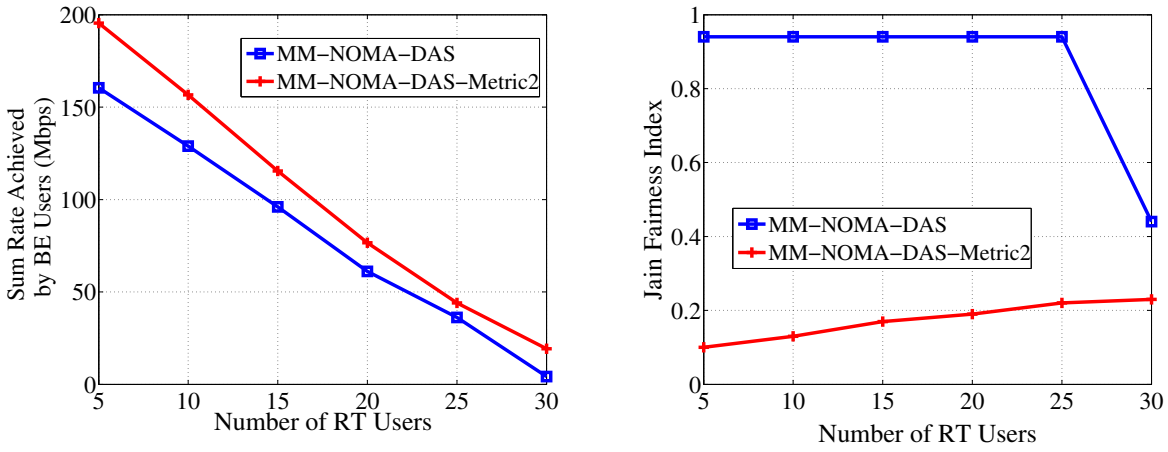
(a) (b)

Figure 3.12 – (a) Percentage of satisfied RT users as K_{RT} increases, (b) Average number of bits received by unsatisfied RT users as K_{RT} increases

Fig. 3.12 and 3.13 show the existing tradeoffs between our proposed metrics and the one in (3.46). In the current work, our goal is to formulate a matching theory based solution for the resource allocation problem in CAS, DAS, OMA and NOMA settings, without necessarily focusing on the optimal scheduling metric. However, to find new metrics reaping the best of these compared ones, a possible future work could consider the formulation, analysis and comparison between multiple metrics. Then, the new metrics can be readily plugged into our proposed algorithms.

3.4.6.2.4 Comparison of the Performance of the MM Technique with the Optimal Solution

On a final note, the performance of the proposed matching-based technique was compared with the optimal solution found by exhaustive search in the NOMA-DAS setting. The exhaustive search method is denoted by ES-NOMA-DAS. For moderate values of the system parameters ($S = 2$, and 4 subbands, $K_{RT} = 2$ users and K_{BE} ranging from 2 to 6 users), the matching-based technique was able to achieve more than 90% of the performance of ES-NOMA-DAS, albeit with a much lower complexity. More precisely, MM-NOMA-DAS requires 2.78% and 5.36×10^{-4}% of the complexity of ES-NOMA-DAS when $K_{BE} = 6$ users, for 2 and 4 subbands, respectively. Moreover, Fig. 3.14 shows a comparison of the Pareto Frontier (the achieved data rate of BE users vs. the minimum percentage of satisfied RT users) between ES-NOMA-DAS and MM-NOMA-DAS, for $K_{RT} = K_{BE} = 4$ users and $S = 2$ subbands. It can be clearly seen that the total throughput achieved by BE users decreases when the minimum percentage of satisfied RT users increases. This is due to RT users requiring more resources for a higher percentage of satisfied RT users, leaving fewer resources for BE users. Fig. 3.14 also shows that the slope of decrease of the exhaustive search method and the matching technique is relatively the same. Moreover, the matching technique achieves 90% of the performance of ES-NOMA-DAS on average, with a much lower complexity. As a conclusion, depending on the system requirements regarding the satisfaction of RT users, a different performance



(a) (b)

Figure 3.13 – (a) Achieved sum rate of BE users in terms of K_{RT}, (b) Achieved Jain fairness index in terms of K_{RT}

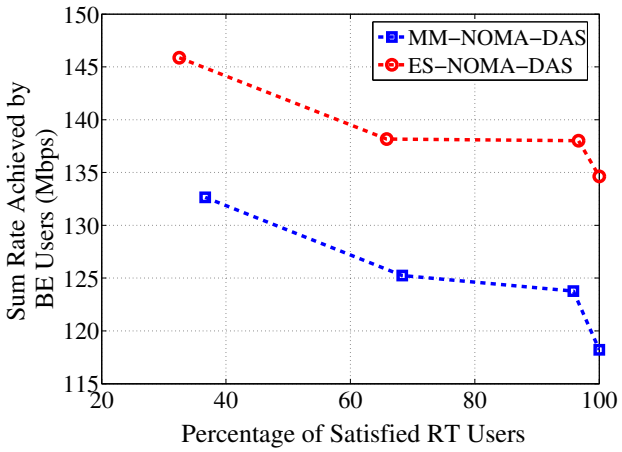


Figure 3.14 – Pareto frontier comparison between MM-NOMA-DAS and ES-NOMA-DAS

of BE users can be achieved.

3.5 Summary

In this chapter, we have proposed a mobile traffic-aware resource allocation framework for the coexistence of latency-constrained and best-effort user types in different system settings. An optimization problem aiming at satisfying the largest number of latency-constrained users and maximizing the achieved rates and fairness level of best-effort users was formulated. To solve this problem, a low-complexity greedy algorithm was first proposed. The greedy algorithm prioritizes RT users by assigning them their preferred resources. A second solution based on matching theory was also proposed for the different considered system settings, i.e., for OMA-CAS, OMA-DAS, NOMA-CAS and NOMA-DAS. The proposed algorithm adapts the deferred-acceptance algorithm to account for different traffic types. Moreover, an iterative approach to optimize the number of sub-

bands per antenna in the distributed setting was also proposed. Simulation results have shown that the proposed methods yield significant performance gains when compared to other conventional solutions that do not account for the particularities of the different traffic types and that do not optimize the subband assignment among the distributed antennas.

3.6 Appendix A

A.1 Proof of Theorem 3

At the beginning of the algorithm, each active user k (respectively each virtual user $v^{k,a}$ in the DAS setting) builds a preference list $\mathcal{PL}_k$ (respectively $\mathcal{PL}_{v^{k,a}}$ in the DAS setting) over the S available subbands. Therefore, $\mathcal{PL}_k,\ \forall\, k \in \mathcal{K}_{\text{active}}$ (respectively $\mathcal{PL}_{v^{k,a}}, \forall\, v^{k,a} \in \mathcal{K}_{\text{active}}$ in the DAS setting), has initially S elements, hence has a finite size. At each iteration of the algorithm, after subbands make decisions regarding accepted users (or user combinations in the NOMA case), each active user removes the subband it proposed to at the current iteration from its preference list. Hence, as the number of iterations increases, the preference lists of active users become smaller. When the maximum number of iterations is reached, the preference lists of active users become empty and the algorithm converges. Next, the maximum number of iterations needed by each setting is evaluated.

1. OMA-CAS: Each user can propose to, at most, S subbands, leading to a maximum number of iterations of S for OMA-CAS.

2. OMA-DAS: The DAS context involves duplicating each user A times. Hence, the system consists of $A \times K$ virtual users, each having preferences over the S subbands. In addition to that, during each iteration, only one of the duplicated users is allowed to propose to its favorite subband. Therefore, at each iteration, K entries of the preference lists are removed, which leads to the maximum number of iterations being upper bounded by $A \times S$.

3. NOMA-CAS: For NOMA, the matching technique is divided into two stages: assignment of subbands followed by user pairing on the subbands assigned to RT users. In the first stage, as the system consists of S subbands, a user (or user pair) can propose to at most S subbands, meaning that the maximum number of iterations before reaching convergence is also S. For the second stage, the maximum number of iterations is given by the number of subbands assigned to RT users, $|\mathcal{S}_{RT}|$, upper bounded by S. Hence, the maximum number of iterations before the CAS version of Algorithm 7 converges is $2 \times S$.

4. NOMA-DAS: The maximum number of iterations in this case is an extension of the OMA-DAS and NOMA-CAS ones. In the first step of the allocation technique, $A \times S$ is the maximum number of iterations, followed by a maximum of S iterations for the second part. Hence, an upper bound for the maximum number of iterations is $S \times (A + 1)$.

A.2 Proof of Theorem 4

The complexity of the proposed matching technique in all system settings is dominated by two steps:

1. sorting the subbands to form the preference lists,
2. the matching step which involves making proposals and decisions.

In Table 3.4, the complexity of each system setting is evaluated. For comparison, the complexity of the optimal method based on exhaustive search is also given.

Table 3.4 – Complexity Analysis

	OMA-CAS	OMA-DAS	NOMA-CAS	NOMA-DAS
Sorting Complexity	$\mathcal{O}\left(KS^2\right)$	$\mathcal{O}\left(KAS^2\right)$	$\mathcal{O}\left(\left(K_{RT}+\lvert\mathcal{US}\rvert\right)S^2\right)$	$\mathcal{O}\left(\left(K_{RT}+\lvert\mathcal{US}\rvert\right)AS^2\right)$
Matching Complexity	$\mathcal{O}\left(KS\right)$	$\mathcal{O}\left(KAS\right)$	$\mathcal{O}\left(\left(K_{RT}+\lvert\mathcal{US}\rvert\right)S\right)$	$\mathcal{O}\left(\left(K_{RT}+\lvert\mathcal{US}\rvert\right)AS\right)$
Overall Complexity	$\mathcal{O}\left(KS^2\right)$	$\mathcal{O}\left(KAS^2\right)$	$\mathcal{O}\left(\left(K_{RT}+\lvert\mathcal{US}\rvert\right)S^2\right)$	$\mathcal{O}\left(\left(K_{RT}+\lvert\mathcal{US}\rvert\right)AS^2\right)$
Exhaustive Search Complexity	$\mathcal{O}\left(K^S\right)$	$\mathcal{O}\left((K\times A)^S\right)$	$\mathcal{O}\left((K+P(K,2))^S\right)$	$\mathcal{O}\left(((K+P(K,2))\times A)^S\right)$

Chapter 4

Uncoordinated Spectrum Access using Multi-Armed Bandits

4.1 Introduction

As mentioned in Chapter 1, future cellular communication networks are expected to support a myriad of new applications and services conceived for both traditional human-type devices and for the growing number of machine-type devices (MTDs) [104]. To meet the exponential growth in connectivity and mobile traffic, new technologies are needed. Among these new technologies, we noted in Chapters 1 and Chapter 2 that the elaboration of novel radio access techniques such as grant-free communications and the densification of cellular networks are promising approaches.

MTDs applications (e.g., smart meters, e-health) generally result in mobile traffic that mostly relies on the uplink transmission [44] of short packet messages. Compared to the small packet sizes of useful information, the signaling overhead resulting from acquiring the channel state information (CSI) at MTDs and sending scheduling requests to a central unit is large. Therefore, optimizing the uplink scheduling of MTDs for efficient spectrum use, by allowing for uncoordinated spectrum access, is of utmost importance.

Moreover, the deployment of different types of access points (APs), e.g., small base stations (SBSs), pico-cells, femto-cells, relays, is of particular importance, since APs can offload mobile traffic from highly congested MBSs [105]. To limit human intervention and reduce planning and maintenance costs, APs can be equipped with self-organizing capabilities [106], allowing them to optimize their resource use in a distributed manner. APs normally have a lower transmit power budget and a smaller coverage range when compared to traditional MBSs. However, thanks to their denser deployment, APs benefit from the ability to consume less transmit power, leading to significant gains in power consumption as shown in [107, 108]. That said, by introducing APs into the network, the problem of inter-cell interference (ICI) is aggravated, necessitating the application of adequate resource allocation algorithms to limit the interference [109]. Various solutions to mitigate the problem of ICI, ranging from centralized solutions, to reinforcement learning (RL)-based solutions, were proposed in the literature.

In Chapter 1 also, we mentioned that to improve system performance, novel multiple access techniques such as non-orthogonal multiple access (NOMA) were recently proposed for future communication systems [28, 82, 110]. From an information-theoretical point of view, it is well-known that non-orthogonal user multiplexing using superposition coding

at the transmitter and proper decoding techniques at the receiver not only outperforms orthogonal multiplexing, but is also optimal in the sense of achieving the capacity region of the downlink broadcast channel [111]. NOMA allows multiple users to be scheduled on the same time-frequency resource by multiplexing them in the power domain. At the receiver side, successive interference cancellation (SIC) is performed to retrieve superimposed signals.

The work presented in this chapter was conducted as part of a research visit at the Coordinated Science Laboratory (CSL) of the University of Illinois at Urbana-Champaign, under the supervision of Professor Venugopal V. Veeravalli. This work has appeared in the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications 2020 [112] and an extended journal version has been submitted to the IEEE Transactions on Communications in July 2020 [113]. Next, we give an overview of the related literature before presenting our problem statement and contributions.

4.1.1 Related Work

The problem of ICI in self-organizing networks (SONs) was extensively studied in the literature. In [114], the weighted sum-rate of the system is optimized through ICI coordination between SBSs. The authors adopt a blanking method where at the level of each SBS, some wireless channels are not used to mitigate the ICI. In [115], an algorithm for ICI coordination between SBSs based on asynchronous inter-cell signaling is proposed. The authors of [116] propose an algorithm based on a semi-static frequency allocation to mitigate ICI and enhance the performance of cell-edge users. The proposed solutions of [114, 115, 116] rely on explicit communication between the distributed SBSs to mitigate ICI, resulting in excessive signaling among SBSs. To limit signaling overhead, decentralized algorithms, based on RL, are preferred.

The use of RL in wireless communications has recently garnered significant attention. In the context of uncoordinated spectrum access, to maximize the accumulated data rate and the number of successful transmissions, [117] and [118] adopt Q-learning, while [119] considers a NOMA system and applies deep RL. The related framework of multi-player multi-armed bandits (MP-MAB) [11] has also been widely used to study multiple problems in wireless communication systems ranging from SONs [40, 120, 121], to uncoordinated spectrum access [12, 13, 15, 122], to fast uplink grant allocation [123], to unmanned-aerial vehicles positioning and path-planning [16]. In the context of SONs, in [40, 120], a solution is proposed based on the stochastic multi-armed bandits (MAB) framework to allow SBSs to partition efficiently the available frequency resources in an effort to mitigate ICI. In [41], a method based on learning automata is proposed where femto-cells adjust their resource use based on the feedback received from users. In [121], the authors resort to the EXP3 algorithm from the adversarial MAB framework to mitigate the ICI while allowing each base station (BS) to access multiple frequency bands. The MAB framework was also widely used to study the opportunistic and the uncoordinated spectrum access problems. For example, in [12], [13] and [14], the MAB model is used to study the opportunistic spectrum access problem in cognitive radio networks where secondary users compete to access the part of the spectrum not occupied by primary users. In addition to studying the opportunistic channel access problem, in [14], the authors also solve the distributed power allocation problem. In contrast to opportunistic channel access, the authors of [15, 122, 124] employ the MAB framework to study the

uncoordinated spectrum access problem without distinguishing between the users. The distributed power control problem is studied in [124], and solutions are proposed based on the Upper-Confidence-Bound (UCB) algorithm, and on the ϵ-greedy algorithm. In [125], the channel and power allocation problem in a device-to-device (D2D) system is modeled using the MAB framework; a game-theoretic solution based on the potential game framework is proposed to minimize regret of users.

With the exception of [121], all previous work on wireless communications solutions based on the MAB framework assumes that each player chooses one channel at each timeslot. However, removing this assumption is expected to improve performance for the players if a suitable algorithm is formulated, especially for the case of a SON. Indeed when an AP can access multiple channels simultaneously, an increase in both the probability of a successful transmission and the achieved reward or rate is expected, allowing the AP to serve more end-users. Moreover, with the exception of [15, 122, 125], all previous work based on the MAB framework considered a zero reward for multiple players accessing the same channel. By alleviating this assumption and adopting NOMA, system performance is expected to further improve.

4.1.2 Problem Statement and Contributions

To limit the ICI in a SON, studying the resource allocation in the fronthaul portion of the network is of utmost importance [40, 120, 121]. When coupled with optimizing the resource allocation in the backhaul link, optimizing the fronthaul portion leads to significant performance gains [126].

In this chapter, we first consider the fronthaul part of a self-organizing wireless network where multiple APs aim at organizing their uplink transmissions with a central unit in a distributed manner. Both the uncoordinated channel access and the distributed power control problems are studied. A solution based on the MAB framework, which does not necessitate any coordination or communication between APs, is proposed. The considered setting is closest to the ones studied in [122] and [127], where a game-theoretic approach is used to solve the uncoordinated channel access problem. Our study extends that of [122] and [127] by allowing each AP to access multiple channels simultaneously and by proposing a model for the distributed power control problem. Second, we propose a simplified model of the uncoordinated spectrum access problem in a cellular network, where users aim to organize their transmissions over multiple channels in an orthogonal multiple access (OMA) manner. The main contributions of this chapter can be summarized as follows:

- A two-phase algorithm based on the MAB framework, extending the work in [122, 127], is proposed for the uncoordinated channel access and distributed power control problems.

- For the first phase, i.e., the uncoordinated channel access phase, in addition to considering varying channel rewards between APs, each AP is allowed to simultaneously access multiple channels. This is in contrast to the work in [122, 127] where each player accesses one channel in a timeslot. Moreover, each channel can accommodate multiple APs at once using NOMA, leading to an MP-MAB problem with varying player rewards, multiple plays and non-zero reward on collision.

- For the power control phase, varying rewards across users for the different power

levels are considered and an algorithm to solve the power control problem on each channel is proposed.

- The proposed technique is shown to achieve a sublinear regret of $\mathcal{O}(\log^2 T)$. In addition, simulation results validating the theoretical results and the performance of the proposed technique are presented.

The rest of this chapter is organized as follows. The system model is presented in Section 4.2. In Sections 4.3, 4.4, 4.5 and 4.6, the proposed algorithm is presented along with an analysis of the system-wide regret. Section 4.7 presents the simplified uncoordinated spectrum access in an OMA setting. The performance of the proposed technique is evaluated in Section 4.8. Finally, a summary of this chapter is given in Section 4.9.

4.2 System Model

Consider the uplink of a cellular system as shown in Fig. 4.1 where K APs aim to organize their communications with the core network, over M available wireless channels, in an uncoordinated manner. The communication occurs over a finite time horizon T that may not be known in advance to the APs. At each timeslot t, every AP k chooses N channels, adjusts its transmission power, and transmits over the chosen channels. We assume that NOMA is employed, enabling multiple APs to choose the same channel for communication and achieve a non-zero rate. That said, if two or more APs choose the

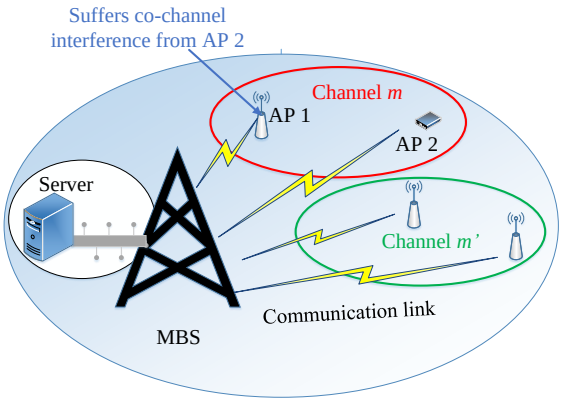


Figure 4.1 – System Model

same channel, the received power levels of these APs must be different at the receiving BS level in the core network, to enable SIC decoding at the receiver side. To ensure the reception of different received power levels for the signals transmitted by the APs, we generalize the uplink NOMA power allocation model introduced in [52] and discussed in Section 2.1.1.2 of Chapter 2, where for a constant signal-to-interference-plus-noise-ratio (SINR) requirement, L received power levels, ensuring the SINR requirement for L users scheduled on the same channel, are calculated. In this work, we extend the study of [52] to allow for L distinct SINR requirements per channel, $\mathbf{\Gamma} = \{\Gamma_1, \ldots, \Gamma_L\}$, sorted by decreasing order. Note that allowing for distinct SINR levels inherently encompasses the special case of constant SINR levels. An AP k choosing SINR requirement Γ_l over channel

m achieves the following uplink data rate:

$$R_{k,m,l} = \log_2 (1 + \Gamma_l), \tag{4.1}$$

where Γ_l is given by:

$$\Gamma_l = \frac{v_l}{V_l + N_0 B_c}. \tag{4.2}$$

In Eq. (4.2), v_l is the received power level of AP k, the expression of which will be given in Section 4.2.2, N_0 is the noise power spectral density and B_c the channel bandwidth. At the receiver side, when the AP transmissions are received with different power levels, SIC is employed to decode the received messages in a descending order. In other words, the AP choosing a higher SINR requirement Γ_1, and consequently a higher received power level v_l, suffers interference from all APs choosing a lower SINR requirement. Once decoded, the signal of the AP choosing Γ_l is removed using SIC before decoding the remaining messages. Hence, variable V_l of Eq. (4.2) is the power level of the interfering transmissions, not canceled with SIC, expressed as: $V_l = \sum_{l'=l+1}^{L} v_{l'}$. To limit the decoding complexity at the receiving BS in the core network, as well as the error propagation in SIC, the number of APs allowed to access a channel and achieve a non-zero rate is limited to β, such that $\beta M \geq KN$. Moreover, it is assumed that when an AP k accesses a channel m, k knows the total number of APs currently accessing channel m. No *a priori* knowledge of the channel gain experienced over each channel is assumed. Moreover, these channel gains are distinct for each AP. To solve the channel and power allocation problems in an uncoordinated manner, we proceed in two steps, the first, of length T_C, dedicated to channel allocation and the second, of length T_P, dedicated to power allocation. Note that both T_C and T_P may not be known to the APs.

4.2.1 Uncoordinated Channel Allocation

To allow each AP to access N channels simultaneously in a NOMA manner, the problem of uncoordinated multiple access is modeled as a stochastic MP-MAB problem with multiple plays and non-zero reward on collision. The set of players is the set of APs $\mathcal{K}$ and the set of arms is the set of channels $\mathcal{M}$. The action of each AP k at each timeslot t is $\boldsymbol{a}_k^t \in \{0,1\}^{M\times 1}$ such that $a_k^t(m) = 1$ if AP k pulls channel m at timeslot t. Moreover, $\sum_{m=1}^{M} a_k^t(m) = N,\ \forall k \in \mathcal{K}$. The action space of each AP k, $\mathcal{A}_k$, consists of all possible combinations of N channels, hence $|\mathcal{A}_k| = \binom{M}{N}$. Let $\boldsymbol{a}^t = \{\boldsymbol{a}_1^t, \ldots, \boldsymbol{a}_K^t\}$ denote the strategy profile of all APs in timeslot t. Upon choosing an action $\boldsymbol{a}_k^t \in \boldsymbol{a}^t$, AP k receives the following reward:

$$g_k^t(\boldsymbol{a^t}) = \sum_{m=1}^{M} a_k^t(m)\mu_M(k, m, k_m), \tag{4.3}$$

where k_m is the number of users choosing channel m at timeslot t. Variable $\mu_M(k, m, k_m)$ is the mean reward of AP k over channel m when k_m APs access it. We assume that the mean reward of AP k when accessing channel m alone is equal to the normalized channel gain of AP k over channel m, i.e.,

$$\mu_M(k, m, 1) = h_{k,m}/\mu_M^{max}, \tag{4.4}$$

where $h_{k,m}$ is the channel gain of AP k over channel m and $\mu_M^{max} = \max_{k\in\mathcal{K}, m\in\mathcal{M}} h_{k,m}$. Note that it is assumed that the BS at the core network performs channel estimation on the

received signals from all APs. Hence, the channel gains $h_{k,m}, \forall k \in \mathcal{K}, \forall m \in \mathcal{M}$ are assumed to be perfectly known by the receiving BS. For $1 < k_m \leq \beta$, the mean reward of an AP must account for the added interference brought by the $k_m - 1$ other APs scheduled on the same channel m. Ideally, the mean reward should take into account the interference brought by each particular AP. However, that would result in a prohibitive complexity since any channel, for each $1 < k_m \leq \beta$, would have $\binom{K-1}{k_m}$ distinct reward values. To simplify the analysis, in this work, we assume that the mean reward for $1 < k_m \leq \beta$, is a decreasing function of the number of interfering APs on the same channel. In other words,

$$\mu_M(k, m, k_m) = \mu_M(k, m, 1)/k_m. \tag{4.5}$$

When $k_m > \beta$, $\mu_M(k, m, k_m) = 0$. The normalization in Eq. (4.4) leads to: $\mu_M(k, m, k_m) \in [0, 1]$ for every AP $k \in \mathcal{K}$, on every channel $m \in \mathcal{M}$ and for every number of APs $k_m \in [\beta]$. Hence, $g_k^t(\boldsymbol{a^t}) \in [0, N]$.

In addition to receiving the achieved rewards, we assume that each AP k receives the total number of APs simultaneously accessing its chosen channels. In other words, for all channels m such that $a_k^t(m) = 1$, AP k receives the total number of APs accessing channel m, i.e., receives $k_m = \sum_{k \in \mathcal{K}} a_k^t(m)$. Note that this assumption is necessary for the correct estimation of the mean rewards, allowing APs to learn and settle on the optimal allocation. Moreover, β is normally kept small. Hence, the assumption that each AP k receives the total number of APs simultaneously accessing its chosen channels results in an overhead of a few bits only.

APs make their decisions in a distributed manner observing neither the channels chosen by other APs nor the rewards received by other APs. Each AP k can only observe the reward it gets on each of its chosen channels. Our aim is to propose a distributed algorithm allowing APs to organize their transmissions on the available channels, without communicating together, in such a way as to maximize the sum reward of the system. By definition, the action profile yielding the highest sum reward $\boldsymbol{a}^*$ is given by:

$$\boldsymbol{a}^* = \underset{\boldsymbol{a} \in \mathcal{A}}{\operatorname{argmax}} \sum_{k=1}^{K} \sum_{m=1}^{M} a_k(m) \mu_M(k, m, k_m). \tag{4.6}$$

The expected regret incurred during T_C is the difference between the achieved reward when playing $\boldsymbol{a}^*$ at all timeslots, and the actually achieved reward by the learning users during the T_C timeslots [11]. In our case, it is given by:

$$\bar{R} = T_C \sum_{k,m} a_k^*(m) \mu_M(k, m, k_m^*) - \mathbb{E}\left(\sum_{t,k,m} a_k(m) \mu_M(k, m, k_m) \right), \tag{4.7}$$

where k_m^* is the optimal number of APs scheduled over channel m under $\boldsymbol{a}^*$.

After T_C timeslots, the APs receive a signal from the core network to terminate the channel allocation phase. At the end of the channel allocation phase, at most β APs are scheduled over each channel $m \in \mathcal{M}$. Moreover, as an outcome of this first phase, each AP k computes an estimate of its channel gain over each channel m, denoted by $\hat{h}_{k,m}$.

4.2.2 Distributed Power Allocation

Once settled over their chosen channels, the APs receive a signal from the core network to move to the power allocation stage. Since different frequency bands are allocated to

different channels, power allocation over each channel m can be done independently of other channels $m' \in \mathcal{M} \setminus \{m\}$. In the following, we will focus on the power allocation over channel $m \in \mathcal{M}$, where the set of scheduled users is $\mathcal{K}_m$.

To simplify the distributed power allocation, we assume that each AP chooses, for each of its allocated channels, one SINR level among a fixed set of $L \geq \beta$ available SINR levels, with $\mathbf{\Gamma}$ being the set of pre-determined available SINR levels. The AP then calculates the necessary power level v_l for the chosen SINR level Γ_l. For successful SIC decoding, each power level can support one AP only. In other words, if multiple APs choose the same power level, SIC fails and the signals of all K_m APs are not decodable. Inspired by [52], it can be shown that, to satisfy Eq. (4.2), the power level v_l must be set as:

$$v_l = \Gamma_l N_0 B_c \prod_{l'=l+1}^{L} (\Gamma_{l'} + 1)\,. \tag{4.8}$$

The expression of v_l ensures the SINR requirement Γ_l when considering that all subsequent SINR requirements have been chosen by other APs, which corresponds to the worst case scenario. Note that our setting allows for similar SINR levels. However, for similar or distinct SINR levels, the power levels chosen by users need to be distinct to allow for SIC decoding.

To ensure SIC stability, i.e., successful decoding of the received signals in descending order [49], the distributed power control scheme must ensure that the power of each signal scheduled for decoding at the BS is larger than the received power of the interference generated by the combination of the remaining signals. In other words, the distributed power control scheme must ensure $v_l > V_l$. From Eq. (4.8), the power level v_l depends on the associated SINR level Γ_l as well as on the interfering SINR levels $\Gamma_{l'}, l' = l+1, \ldots, L$.

Proposition 2. *To ensure SIC stability, the available SINR levels must satisfy:*

$$\Gamma_l > \frac{2^{(L-l-1)} \times \Gamma_L}{\prod_{l'=l+1}^{L} (\Gamma_{l'} + 1)}. \tag{4.9}$$

Proof. By proceeding backwards, to get $v_{L-1} > v_L$, the following must hold:

$$\Gamma_{L-1} > \frac{\Gamma_L}{\Gamma_L + 1} = \frac{2^{(L-(L-1)-1)}\Gamma_L}{\Gamma_L + 1}. \tag{4.10}$$

Similarly, to get $v_{L-2} > v_{L-1} + v_L$, the following must hold:

$$\Gamma_{L-2} > \frac{2\Gamma_L}{(\Gamma_{L-1} + 1)(\Gamma_L + 1)} = \frac{2^{(L-(L-2)-1)}\Gamma_L}{\prod_{l'=L-1}^{L} (\Gamma_{l'} + 1)}. \tag{4.11}$$

To get $v_l > \sum_{l'=l+1}^{L} v_{l'}$, assume that Eq. (4.9) holds. By induction, to get $v_{l-1} > \sum_{l'=l}^{L} v_{l'}$, we must have:

$$\Gamma_{l-1} > \frac{2^{(L-(l-1)-1)}\Gamma_L}{\prod_{l'=l}^{L} (\Gamma_{l'} + 1)}. \tag{4.12}$$

□

Knowing the available SINR levels, each AP $k \in \mathcal{K}_m$ calculates the associated received power levels using Eq. (4.8). Then, using the estimated channel gain over m, $\hat{h}_{k,m}$, AP $k \in \mathcal{K}_m$ calculates the necessary transmit power for each power level v_l according to:

$$p_{k,m,l} = v_l / \hat{h}^2_{k,m}. \tag{4.13}$$

Each AP is assumed to have a power budget per channel P^m_k. Hence, AP k can transmit over channel m using power level v_l if $p_{k,m,l} \leq P^m_k$. AP $k \in \mathcal{K}_m$ builds the set of possible power levels, $\mathcal{P}^a_{k,m}$, where $\mathcal{P}^a_{k,m} = \{v_l |\ p_{k,m,l} \leq P^m_k, l \in [L]\}$. Note that the set of possible power levels are AP-dependent because of their dependency on the estimated channel gain of each AP, $\hat{h}_{k,m}$, and on the AP power budget.

The power allocation among APs on the same channel consists of APs choosing SINR levels, and hence received power levels, in a distributed manner, and without any inter-AP coordination. Since APs choosing the same SINR level result in an unsuccessful SIC decoding, the APs must aim at organizing their transmissions using different SINR levels. For this purpose, the power allocation on each channel is modeled using the MAB framework with single play and zero-reward on collision. Over channel m, the set of players is $\mathcal{K}_m$ and the set of arms is the set of power levels $\mathcal{VL}$. Since $L = |\mathcal{VL}| \geq \beta \geq K_m = |\mathcal{K}_m|$, a solution where each AP accesses one power level, without collision, is achievable. At each timeslot, each AP $k \in \mathcal{K}_m$ chooses an action $a^t_{k,m}$, i.e., a power level $v_l \in \mathcal{P}^a_{k,m}$, and transmits using $p_{k,m,l}$. The action space of AP k is $\mathcal{P}^a_{k,m}$. Let $\boldsymbol{a}^t_m$ denote the strategy chosen by all APs in $\mathcal{K}_m$ over channel m at timeslot t. Upon choosing action $a^t_{k,m} \in \boldsymbol{a}^t_m$, AP k receives the following reward on channel m:

$$g^t_{k,m}(\boldsymbol{a}^t) = \mu_P(k, m, a^t_{k,m})\eta(\boldsymbol{a}^t_m),, \tag{4.14}$$

where $\mu_P(k, m, a^t_{k,m})$ is the mean reward of AP k when choosing $a^t_{k,m}$. The mean reward $\mu_P(k, m, a^t_{k,m})$ is chosen in a way to strike a trade-off between SINR maximization and transmit power minimization. Therefore, it is set as:

$$\mu_P(k, m, a^t_{k,m} = v_l) = w^1_k \frac{\Gamma_l}{\Gamma_1} + w^2_k \frac{1}{p_{k,m,l}\ \max\limits_{k,m,l}(\frac{1}{p_{k,m,l}})}, \tag{4.15}$$

where w^1_k and w^2_k are weight parameters relative to user $k \in \mathcal{K}_m$ satisfying $w^1_k + w^2_k = 1$. Note that $\mu_P(k, m, a^t_{k,m}) \in [0, 1]$ and is not known by the AP in advance. Let $\mathcal{N}^m_{v_l}(\boldsymbol{a}^t_m)$ be the set of APs choosing power level v_l at timeslot t, i.e., $\mathcal{N}^m_{v_l}(\boldsymbol{a}^t_m) = \{k \in \mathcal{K}_m \mid a^t_{k,m} = v_l\}$. Variable $\eta(\boldsymbol{a}^t_m)$ is the collision indicator of the strategy profile of all APs, $\boldsymbol{a}^t_m$, i.e., $\eta(\boldsymbol{a}^t_m) = 1$ if $|\mathcal{N}^m_{a^t_{k,m}=v_l}(\boldsymbol{a}^t_m)| \leq 1, \forall\ v_l \in \mathcal{VL}$, and 0 otherwise.

APs choose power levels in a distributed manner without any coordination, with each AP only observing the reward received on the chosen power level. The proposed power allocation scheme aims at maximizing the sum reward of the system. Let $\boldsymbol{a}^{*P}_m$ be the action profile yielding the highest sum reward over channel m:

$$\boldsymbol{a}^{*P}_m = \underset{\boldsymbol{a}_m \in \mathcal{P}^a_m}{\operatorname{argmax}} \sum_{k \in \mathcal{K}} \mu_P(k, m, a^t_{k,m})\, \eta(\boldsymbol{a}^t_m). \tag{4.16}$$

The expected regret incurred during the time horizon T_P over all M channels is given by:

$$\bar{R}_p = \sum_{m \in \mathcal{M}} \left\{ T_P \sum_{k \in \mathcal{K}} \mu_P(k, m, a^{*P}_{k,m}) - \mathbb{E}\left(\sum_{t,k} \mu_P(k, m, a^t_{k,m})\, \eta(\boldsymbol{a}^t_m) \right) \right\}. \tag{4.17}$$

4.3 Proposed Solution

4.3.1 Proposed Algorithm for the Channel Allocation Problem

Since the time horizon T_C is not necessarily known in advance, the proposed solution, presented in Algorithm 8, proceeds in epochs, each epoch consisting of three phases, namely, *exploration, matching and exploitation*. The exploration phase aims at estimating the previously unknown means of each channel, as well as the total number of APs competing for system resources. During this phase, each AP uniformly accesses one channel at a time to estimate its mean reward. AP k accessing channel m gets as feedback the achieved reward on m as well as the total number of APs simultaneously accessing channel m. This phase runs for a constant number of timeslots given by T_C^0. Upon termination, all APs have an estimate $\hat{\boldsymbol{\mu}}_M$ of the means of the channels and of the channel gain experienced over each channel. Each AP also calculates an estimate of the number of APs $\hat{K}$, as done in [15]. These estimated means and number of APs are used in the second phase of the algorithm where APs play a non-cooperative game with the aim of maximizing the achieved sum rewards. The estimated reward means are taken to be the actual utilities achieved in the matching phase. In other words, after choosing a channel m, if the received reward is non-zero, AP k assumes that this reward is equal to:

$$u_k(m) = \hat{\mu}_M(k, m, k_m), \tag{4.18}$$

where k_m is the received total number of users simultaneously accessing channel m.

The dynamics of this matching phase, adopted from [128], are described in Section 4.3.2. The matching phase runs for $c_1 l^{1+\delta}$ frames, l being the epoch number. The third and final phase is an exploitation phase in which APs settle on the channels that resulted in the best performance in the previous matching phase. The exploitation phase runs for $c_2 2^l$ timeslots.

4.3.2 Matching Dynamics

Each AP k is associated with a state $[\bar{\boldsymbol{a}}_k, \bar{\boldsymbol{u}}_k, S]$. The baseline action of AP k is $\bar{\boldsymbol{a}}_k \in \{0,1\}^{M\times 1}$, such that $\sum_{m=1}^{M} \bar{a}_k(m) = N$. The baseline utility of AP k is $\bar{\boldsymbol{u}}_k$, such that $|\bar{\boldsymbol{u}}_k| = N$. Variable $S \in \{C, D\}$ is the mood of AP k and reflects whether k is content or discontent with the current action and utility. At each frame of the matching phase, each AP chooses an action according to the game dynamics and receives a reward that depends on the collective choices of all the APs. Define $u_{k,\max} = \underset{\boldsymbol{a}}{\operatorname{argmax}} \sum_{m=1}^{M} a_k(m)\mu_M(k, m, k_m)$, where $u_{k,\max}$ is the highest reward achievable by AP k, with a number of estimated APs given by $\hat{K}$.

At each frame t during the matching phase, AP k adheres by the following dynamics to decide on the action to choose:

- A content AP plays its baseline action with high probability:

$$p_k^{\boldsymbol{a}_k} = \begin{cases} \frac{\epsilon^c}{|\mathcal{A}_k|-1}, & \text{if } \boldsymbol{a}_k \neq \bar{\boldsymbol{a}}_k, \\ 1-\epsilon^c, & \text{if } \boldsymbol{a}_k = \bar{\boldsymbol{a}}_k. \end{cases} \tag{4.20}$$

where $\epsilon > 0$ is a small perturbation and c is a constant satisfying $c \geq KN$.

Algorithm 8

Initialization: Set $\hat{\mu}_M(k,m,k_m)=0,\ \forall k\in\mathcal{K}, m\in\mathcal{M}, k_m\in[\beta]$. Set $b_k^t=0,\ \forall k\in\mathcal{K}$, $co_k^t(k_m)=0,\ \forall k\in\mathcal{K}, k_m\in[\beta]$. Let $\epsilon>0$ and $c\geq KN$.

for $l=1,\ldots,L_C$ **do**

1- Exploration Phase:

  for $t=1:T_C^0$ **do**

    Choose one channel $m\in\mathcal{M}$ uniformly.

    Receive the achieved reward $x_k^t(m)$, and the total number of users, k_m, accessing channel m simultaneously.

    $W_k^t(m,k_m)=W_k^{t-1}(m,k_m)+x_k^t(m)$,
    $co_k^t(k_m)=co_k^t(k_m)+1$.

    if $k_m>1$ **then**

      $b_k^t=b_k^{t-1}+1$

    end if

  end for

  Estimate means: $\hat{\mu}_M(k,m,k_m)=\frac{W_k^t(m,k_m)}{co_k^t(k_m)},\ \forall\ k_m\in[\beta]$.

  Estimate the number of APs according to: $\hat{K}=\min\left\{\text{round}\left(\frac{\log\left(\frac{T_C^0-b_k^t}{T_C^0}\right)}{\log\left(1-\frac{1}{M}\right)}+1\right),\beta M\right\}$.

2- Matching Phase: for the next $c_1 l^{1+\delta}$ frames, play according to the dynamics described in section 4.3.2.

  If $S_k=C$, choose the action to play according to Eq. (4.20). If $S_k=D$, choose the action according to Eq. (4.21).

  If the achieved reward for some chosen channel $u_k(m)$, found from Eq. (4.18), is 0, the AP becomes discontent as per Eq. (4.23).

  If $\boldsymbol{a}_k\neq\bar{\boldsymbol{a}}_k$ or $\boldsymbol{u}_k\neq\bar{\boldsymbol{u}}_k$ or player k is discontent, the state transition happen according to Eq. (4.24).

  Each AP keeps a counter of the number of times each action $\boldsymbol{a'_k}$ was played and resulted in it being content:

$$F_k^l(\boldsymbol{a'_k})=\sum_{t=1}^{c_2 l^{1+\delta}}\mathbb{I}\left(\boldsymbol{a}_k^t=\boldsymbol{a'_k},S_k^t=C\right),\tag{4.19}$$

  with $\mathbb{I}$ being the indicator function.

3- Exploitation phase: for $c_2 2^l$ timeslots:

  Play the action $\boldsymbol{a}_{\boldsymbol{k}}^{\boldsymbol{l*}}=\underset{\boldsymbol{a}_k\in\mathcal{A}_k}{\text{argmax}}\, F_k^l(\boldsymbol{a}_k)$.

end for

- A discontent AP chooses its action uniformly at random:

$$p_k^{\boldsymbol{a}_k} = \frac{1}{|\mathcal{A}_k|}, \quad \forall\ \boldsymbol{a}_k \in \mathcal{A}_k. \tag{4.21}$$

After deciding on the action and observing the reward $u_k(m)$ for chosen channels, the state transition of each AP k occurs according to:

- If $\boldsymbol{a}_k = \bar{\boldsymbol{a}}_k$ and $\boldsymbol{u}_k = \bar{\boldsymbol{u}}_k$, a content AP remains content with probability one:

$$[\bar{\boldsymbol{a}}_k, \bar{\boldsymbol{u}}_k, C] \rightarrow [\bar{\boldsymbol{a}}_k, \bar{\boldsymbol{u}}_k, C]. \tag{4.22}$$

- If $u_k(m) = 0$ for some $m = 1, \ldots, N$, AP k becomes discontent with probability one:

$$[\bar{\boldsymbol{a}}_k, \bar{\boldsymbol{u}}_k, C/D] \rightarrow [\boldsymbol{a}_k, \boldsymbol{u}_k, D]. \tag{4.23}$$

- If $\boldsymbol{a}_k \neq \bar{\boldsymbol{a}}_k$ or $\boldsymbol{u}_k \neq \bar{\boldsymbol{u}}_k$ or player k is discontent, the state transition occurs according to:

$$[\bar{\boldsymbol{a}}_k, \bar{\boldsymbol{u}}_k, C/D] \rightarrow \begin{cases} [\boldsymbol{a}_k, \boldsymbol{u}_k, C] & \text{with probability } \epsilon^{u_{k,\max} - \sum_{n=1}^{N} u_{k,n}}, \\ [\boldsymbol{a}_k, \boldsymbol{u}_k, D] & \text{with probability } 1 - \epsilon^{u_{k,\max} - \sum_{n=1}^{N} u_{k,n}}. \end{cases} \tag{4.24}$$

4.3.3 Proposed Solution for the Distributed Power Allocation

A simplified version of Algorithm 8 can be used to solve the power allocation problem over each channel m. The solution is divided into three phases:

1. Exploration phase: This phase runs for T_P^0 timeslots and aims at estimating the reward of each power value. During this phase, each AP chooses each of its possible power levels, i.e., power levels in $\mathcal{P}_k^a$, uniformly at random. Upon termination, APs have estimates of the reward associated to each power value, denoted by $\hat{\boldsymbol{\mu}}_P$.

2. Matching phase: In this phase, APs play a non-cooperative game according to the dynamics presented in Section 4.3.2, after replacing $\mathcal{A}_k$ in Eq. (4.20) and (4.21) by $\mathcal{P}_{k,m}^a$. Each AP keeps a counter of the number of times each action was played and resulted in content behavior.

3. Exploitation phase: During this phase, each AP k exploits the action, i.e., the power level, that resulted in the most content behavior during the matching phase.

4.4 Regret Analysis

The time horizon of the channel allocation phase can be lower-bounded by [127]:

$$T_C \geq \sum_{l=1}^{L_C - 1} (T_C^0 + c_1 l^{1+\delta} + c_2 2^l) \geq c_2(2^{L_C} - 2), \tag{4.25}$$

where L_C is the total number of epochs occurring within T_C and upper bounded by:

$$L_C \leq \log\left(T_C/c_2 + 2\right). \tag{4.26}$$

Similarly, the number of epochs, L_P, occurring within the time horizon T_P dedicated to the power allocation stage is upper bounded by $L_P \leq \log(T_P/c_2 + 2)$.

4.4.1 Regret in the Exploration Phase

In the exploration phase of the channel allocation, each AP samples channels uniformly to get estimates of their means. Even though the purpose of this work is to assign to each AP N channels at each timeslot, the number of channels sampled by each AP at a timeslot is set to one in the exploration phase. The expected regret incurred by all APs in the exploration phase of the channel allocation, R_C^1, can be upper bounded by:

$$R_C^1 \leq \sum_{l=1}^{L_C} KNT_C^0 \leq KNT_C^0 \log\left(T_C/c_2 + 2\right). \tag{4.27}$$

Similarly, the expected regret incurred by all APs in the exploration phase of the power allocation, R_P^1, can be upper bounded by:

$$R_P^1 \leq \sum_{m=1}^{M} \sum_{l=1}^{L_P} K_m T_P^0 \leq KT_P^0 \log(T_P/c_2 + 2). \tag{4.28}$$

4.4.2 Regret in the Matching Phase

The expected regret in the matching phase of the channel allocation, R_C^2, can be upper bounded by:

$$R_C^2 \leq \sum_{l=1}^{L_C} KNc_1 l^{1+\delta} \leq KNc_1 \log^{2+\delta}\left(T_C/c_2 + 2\right). \tag{4.29}$$

Similarly, the expected regret in the matching phase of the power allocation, R_P^2, can be upper bounded by:

$$R_P^2 \leq \sum_{m=1}^{M} \sum_{l=1}^{L_C} K_m c_1 l^{1+\delta} \leq Kc_1 \log^{2+\delta}\left(T_P/c_2 + 2\right). \tag{4.30}$$

4.4.3 Regret in the Exploitation Phase

In the exploitation phase of epoch l of the channel allocation, each AP k plays the action that it played the most and resulted in content behavior in the matching phase of epoch l. The exploitation phase fails in two cases:

1. If the exploration phase of epoch l fails: This happens with a probability $\leq 4(M\beta)^2 e^{-l}$ as shown in Lemma 2.

2. If the most played action of the matching epoch differs from the optimal action: This happens with a probability $\leq A_1 e^{-l^{1+\delta}}$ as shown in Lemma 5.

The expected regret incurred by all APs in the exploitation phase can be upper bounded by:

$$R_C^3 \leq \sum_{l=1}^{L_C} KNc_2 2^l \left(4(M\beta)^2 e^{-l} + A_1 e^{-l^{1+\delta}}\right) \leq A_3. \tag{4.31}$$

Similarly, the regret incurred by the APs in the exploitation phase of the power allocation is $R_P^3 \leq A_3$.

4.4.4 Regret of the Proposed Technique

Theorem 6. *The expected regret of the proposed allocation solution can be upper bounded as:*

$$R \leq R_C^1 + R_C^2 + R_C^3 + R_P^1 + R_P^2 + R_P^3 = \mathcal{O}\left(\log^{2+\delta}(T)\right). \tag{4.32}$$

4.5 Exploration Phase

The exploration phase is performed so that APs learn estimates of the channel mean reward in the channel allocation phase, and of the power level mean reward in the power allocation phase. Moreover, by keeping track of the number of times each channel was accessed with one or more other users in the channel allocation phase, the APs can estimate the total number of APs in the system. In this section, we find the minimum length of the exploration phase ensuring an accurate estimation of both the reward means and the number of APs.

4.5.1 Estimation of the Reward Means

Since the estimation may not always be perfect, the result of the assignment with the estimated means ($\hat{\boldsymbol{\mu}}_M$ and $\hat{\boldsymbol{\mu}}_P$) might differ from the result of the assignment calculated with the true means ($\boldsymbol{\mu}_M$ and $\boldsymbol{\mu}_P$). However, if the estimation inaccuracy is kept small as in [122] and [127], the result of the assignment would not be affected.

Lemma 1. *Let J_M^1 and J_M^2 be the sum reward achieved by the best channel assignment and the second best channel assignment and let $\Delta_M = \frac{J_M^1 - J_M^2}{2KN}$. Moreover, let J_P^1 and J_P^2 be the sum reward achieved by the best power allocation on each channel m and the second best power assignment and let $\Delta_P = \frac{J_P^1 - J_P^2}{2K_m}$. If the difference between the estimated and the correct reward means satisfies:*

$$|\mu_M(k,m,k_m) - \hat{\mu}_M(k,m,k_m)| < \Delta_M, \forall k \in \mathcal{K}, m \in \mathcal{M}, k_m \in [\beta], \tag{4.33}$$

$$|\mu_P(k,m,v_l) - \hat{\mu}_P(k,m,v_l)| < \Delta_P, \forall k \in \mathcal{K}_m, m \in \mathcal{M}, v_l \in \mathcal{VL}, \tag{4.34}$$

then, the best assignment result does not change due to the estimation inaccuracy.

Proof. See Appendix A.1. □

Next, we upper bound the probability of error, i.e., the probability of having channel reward estimates (resp. power level reward estimates) that do not satisfy the condition in (4.33) (resp. (4.34)) in the exploration epoch l. We also provide a lower bound of the length of the exploration epoch $T_{\hat{\boldsymbol{\mu}}_M}$ in the channel allocation phase, and T_P^0 in the power allocation phase.

Lemma 2. *If $T_{\hat{\boldsymbol{\mu}}_M} = \left\lceil \frac{2Me\left(\frac{K-1}{M-1}\right)}{\Delta_M^2 (M-1)^{1-\beta}} \right\rceil$, all players have an estimate of the channel means satisfying the condition in (4.33), with probability $\geq 1 - \gamma_{e,l}^M$. Moreover, the probability of error in the l^{th} exploration phase of the channel allocation solution, $\gamma_{e,l}^M$ is upper bounded by $4(M\beta)^2 e^{-l}$. For the power allocation exploration phase, if $T_P^0 = \left\lceil \frac{2Le\left(\frac{\beta-1}{L-1}\right)}{\Delta_p^2} \right\rceil$, all players have an estimate of the power level means satisfying the condition in (4.34), with probability $\geq 1 - \gamma_{e,l}^P$, where $\gamma_{e,l}^P$ is upper bounded by $4\beta L e^{-l}$.*

Proof. See Appendix A.2. □

We now turn our attention to finding the minimum length of the exploration phase in the channel allocation stage ensuring an accurate estimate of the number of APs $\hat{K}$.

4.5.2 Estimating the number of APs

For AP k, b_k^t found in step 7 of Algorithm 8 denotes the number of timeslots player k was not the sole occupier of some channel m until t.

Lemma 3. *If the length of the exploration epoch in the channel allocation step satisfies:*

$$T_{\hat{K}} = \left\lceil 2.08 \log\left(\frac{2}{\eta}\right) M^2 e^{2\left(\frac{M\beta-1}{M-1}\right)} \right\rceil, \tag{4.35}$$

then all APs have an estimate of the number of APs $\hat{K}$ satisfying $\hat{K} = K$ with probability higher than $1-\eta$.

Proof. See Appendix A.3. □

4.5.3 Length of the Channel Allocation Exploration Phase

To ensure an accurate estimate of the channel reward means and of the number of APs, the minimum length of the exploration phase in the channel allocation solution, T_C^0, must satisfy the conditions in Lemma 2 and Lemma 3. Hence, the following must hold:

$$T_C^0 = \max\left\{ \left\lceil \frac{2Me^{\left(\frac{K-1}{M-1}\right)}}{\Delta_M^2 (M-1)^{1-\beta}} \right\rceil, \left\lceil 2.08 \log\left(\frac{2}{\eta}\right) M^2 e^{2\left(\frac{M\beta-1}{M-1}\right)} \right\rceil \right\}. \tag{4.36}$$

4.6 Matching Phase

The matching phase of the channel allocation solution aims at reaching a final assignment in which every AP accesses N channels, such that the achieved sum reward is maximized.

The dynamics presented in Section 4.3.2 and adopted in the matching phase induce a Markov chain over the state space $\mathcal{Z} = \prod_{k=1}^{K} \{\mathcal{A}_K \times [0,1]^{N\times 1} \times \{C, D\}\}$. Let P^ϵ denote the transition matrix of the regular perturbed Markov chain $\mathcal{Z}$. The following theorem was proved in [128]:

Theorem 7. *For a K-player interdependent game over a finite joint action space $\mathcal{A}$, the dynamics presented in Section 4.3.2 ensure that a state $z = \prod_{k=1}^{K} [\bar{\boldsymbol{a}}_k, \bar{\boldsymbol{u}}_k, S_k] \in \mathcal{Z}$ is stochastically stable if and only if:*

1. *the action profile $\bar{\boldsymbol{a}} = \{\bar{\boldsymbol{a}}_1, \ldots, \bar{\boldsymbol{a}}_K\}$ maximizes the achieved utility $\sum_{k=1}^{K} \sum_{n=1}^{N} \bar{u}_k(n)$,*
2. *the baseline actions and rewards are aligned, i.e., $\bar{u}_k(n) = \mu(k, \bar{a}_k(n), k_n),\ \forall k \in \mathcal{K}, n \in \{1, \ldots, N\}$,*
3. *the mood of each user is content, i.e., $S_k = C,\ \forall k \in \mathcal{K}$.*

Theorem 7 guarantees that, when playing according to the dynamics of Section 4.3.2, the optimal state, i.e., the one maximizing the sum rewards, is played most often. The proof relies on the theory of resistance trees for regular perturbed Markov chains [68]. The dynamics used in this chapter differ from those in [128] in two aspects:

1. If AP k receives a reward equal to 0 on some channel m, AP k is discontent with probability one. In [128], the game is assumed to be interdependent which means that it is not possible to partition APs into two groups that do not interact with each other. However, this property does not hold in the considered setting as shown in [127]. Therefore, as in [127], to characterize the stable states of the unperturbed chain when $\epsilon = 0$, a player with 0 reward on some channels is discontent with probability one.

2. For the transition probabilities between content and discontent in (4.24), instead of using $\epsilon^{N-\sum_{n=1}^{N} u_{k,n}}$, we use $\epsilon^{u_{k,\max}-\sum_{n=1}^{N} u_{k,n}}$, since the maximum utility achievable by each user k is $u_{k,\max}$.

Next, the recurrence classes of $\mathcal{Z}$ are characterized.

Lemma 4. *Let D^0 denote the set of states where all users are discontent. Moreover, let C^0 denote all singleton states where all users are content and their baseline actions and utilities are aligned. The only recurrence classes of $\mathcal{Z}$ are D^0 and all singletons in C^0 [128].*

Proof. In the unperturbed Markov process, all discontent APs remain so with probability one. Therefore, D^0 is a recurrence class. Moreover, in each state $z \in C^0$, all APs play their baseline action and receive their baseline utility with probability one, hence they remain content with probability one. Therefore, all singletons $z, z' \in C^0$ are recurrence classes.

To show that these are the only recurrence classes, consider a state $z \notin D^0 \cup C^0$. At least one AP k is discontent in such a state z. In the next timeslot, k chooses its channels uniformly at random. Hence, there is a positive probability that k will choose some channel m occupied by some other content AP k'. The utility of AP k' will become misaligned with its baseline utility leading AP k' to become discontent, while AP k remains discontent. This continues until all APs become discontent. The state $z \notin D^0 \cup C^0$ cannot therefore be a recurrence class. Now consider a state z where all APs are content but the baseline action and utility of player k are not aligned. At the next timeslot, k becomes discontent and it follows from the above argument that all players become discontent.

Hence, the only recurrence classes are D^0 and all singletons in C^0. □

Similarly to [128], the resistance of moving from D^0 to any state $z \in C^0$ is:

$$r(D^0 \to z) = \sum_{k=1}^{K} \left(u_{k,\max} - \sum_{n=1}^{N} u_{k,n} \right). \tag{4.37}$$

The transition $z \in C^0 \to D$ has a resistance of:

$$r(z \to D^0) = c, \tag{4.38}$$

and the resistance of moving from any state $z \in C^0$ to $z' \in C^0$ is:

$$c \leq r(z \to z') \leq 2c. \tag{4.39}$$

The stochastic potential of any state $z \in C^0$ is of the form:

$$\zeta(z) = c[|C^0| - 1] + \sum_{k=1}^{K} u_{k,\max} - \sum_{m=1}^{M} a_k(m)\hat{\mu}(k, m, k_m). \tag{4.40}$$

Since the stochastic potential of any state $z \in C^0$ given in Eq. (4.40) is the same as that in [128], the proof of Theorem 7 given [128] still holds. This proof shows that the stable state is the one minimizing the stochastic potential, hence maximizing the achieved sum reward. This stable state is guaranteed to be played the majority of times for a small enough perturbation ϵ [127], [128]. In the exploitation phase, as the state that was most played and that resulted most in the players being content is played, the stable state is hence expected to be played with high probability. Next, the probability of error in the matching epoch l is found.

Let π denote the stationary distribution of the Markov chain $\mathcal{Z}$ and let $\boldsymbol{z}^* = [\bar{\boldsymbol{a}}^*, \bar{\boldsymbol{u}}^*, C^K]$ denote the optimal state. According to [127], $\pi(\boldsymbol{z}^*) > 1/2$ for a small enough perturbation ϵ. The following lemma finds the probability of error in the matching phase of the l^{th} epoch, $\delta_{m,l}$.

Lemma 5. *Let $\boldsymbol{a}^{(l)}$ denote the action that was most played in some epoch l. The probability of error in the matching phase in epoch l, $\delta_{m,l}$, is upper bounded by:*

$$\delta_{m,l} = Pr(\boldsymbol{a}^* \neq \boldsymbol{a}^{(l)}) \leq A_0 \left\|\phi\right\|_{\pi} \exp\left(\frac{-\theta^2 \pi(\boldsymbol{z}^*) c_2 l^{1+\delta}}{72 T_m(1/8)}\right), \tag{4.41}$$

where A_0 is a constant, ϕ_π is the probability distribution of the initial state played in epoch l and $T_m(1/8)$ is the mixing time of the Markov chain $\mathcal{Z}$ with an accuracy of 1/8 [69].

Proof. See Appendix A.4. □

The analysis of the matching phase of the power allocation solution is similar to the one given above and is omitted for concision.

4.7 Uncoordinated Spectrum Access for Grant-Free Communications

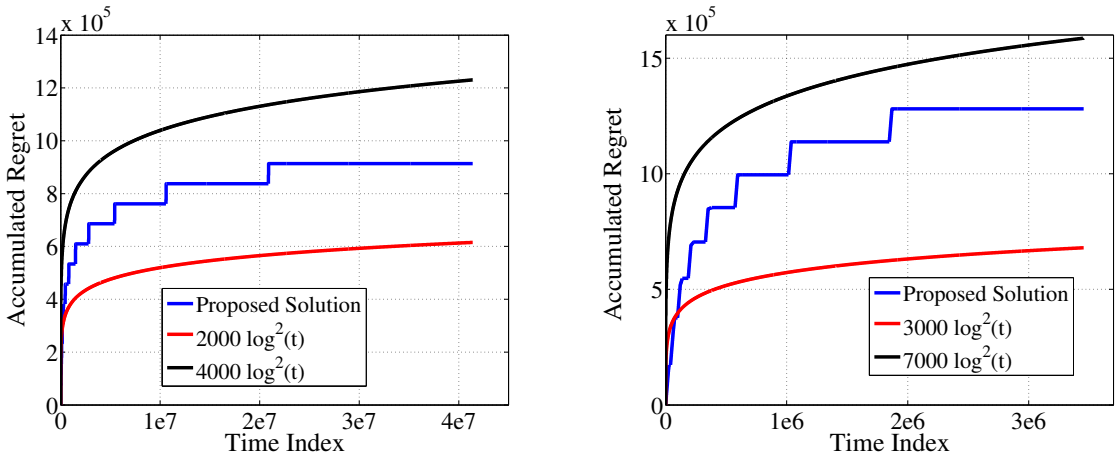
A simplified version of the solution to the uncoordinated spectrum access based on the MAB framework proposed in Section 4.3 is also tested for grant-free communications. This simplified version assumes OMA scheduling, i.e., assumes that each channel can accommodate one user only. A user is allowed to access N channels in each timeslot. When two or multiple users access the same channel, these users receive a zero reward, indicating a collision has occurred. Hence, this problem can be modeled as an MP-MAB with multiple plays, and varying rewards across users. Algorithm 8, proposed in Section 4.3, can be used to solve the simplified uncoordinated spectrum access problem, after setting $\beta = 1$. The sub-linear regret of this algorithm, found in Section 4.4.4 holds. Moreover, the minimum length of the exploration phase can be found from Eq. (4.36), after setting $\beta = 1$.

4.8 Simulation Results

Extensive simulations of the proposed algorithm were conducted to validate its performance. We first evaluate the performance of the simplified version of the uncoordinated spectrum access for grant-free communications in Section 4.8.1. Then, we turn our attention to the evaluation of the performance of the solution to the uncoordinated spectrum access and power control in a SON in Section 4.8.2.

4.8.1 Performance of the OMA Version of the Uncoordinated Spectrum Access

The following simulation parameters were chosen: $c_1 = 3000, c_2 = 5000, \epsilon = 10^{-4}, c = KN, \delta = 0$. Note that the definitions of these variables are given in Section 4.3.



(a) (b)

Figure 4.2 – Accumulated regret as time progresses for (a) $K = 2, M = 6, N = 3$, (b) $K = 4, M = 8, N = 2$

Fig. 4.2 shows the average accumulated regret as a function of time, averaged over 50 realizations of the algorithm, for different system settings. The results show that the average accumulated regret increases with time as $\mathcal{O}(\log(t)^2)$. More specifically, the regret for a system consisting of $K = 2$ users, $M = 6$ channels, and where each user pulls $N = 3$ channels in each timeslot, is bounded between $2000\log(t)^2$ and $4000\log(t)^2$, as shown in Fig. 4.2a. The regret of a system consisting of $K = 4$ users, $M = 8$ channels and with $N = 2$ channels pulled at each timeslot, is bounded between $3000\log(t)^2$ and $7000\log(t)^2$ as shown in Fig. 4.2b. The higher regret observed for the case of $K = 4$ users is due to the system taking a longer time to converge as shown next.

In Fig. 4.3, the mean achieved reward normalized by the reward of the optimal allocation is shown for different system settings. The performance of the proposed method is compared against that of the UCB algorithm [11]. The simulated UCB algorithm consists of each player choosing N arms according to the UCB technique without consideration for the other players in the system. The performance of the random selection technique, where each user chooses an action uniformly at random (i.e., with a probability given by Eq. (4.21)), is also shown for comparison. Fig. 4.3 shows the superior performance of

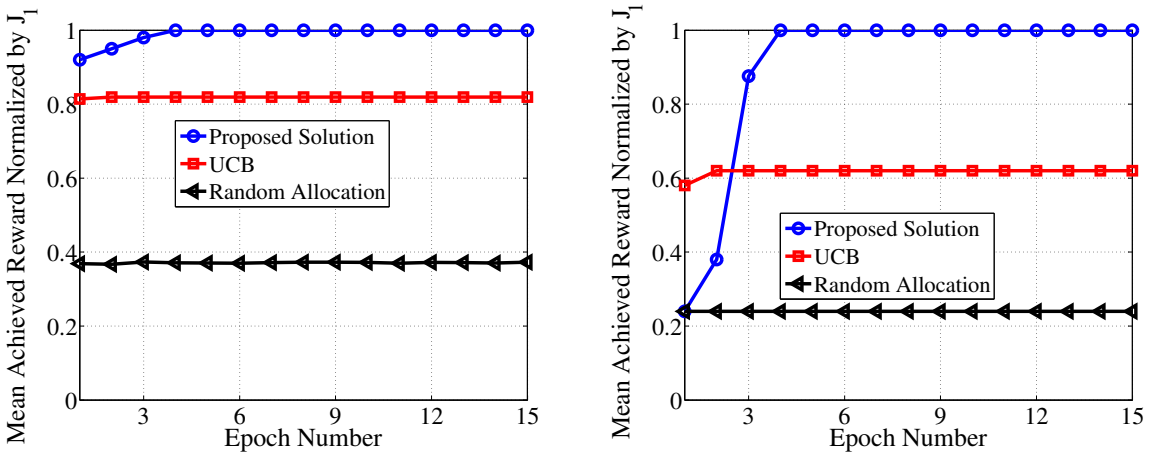


Figure 4.3 – Mean of the achieved reward normalized by the reward of the optimal allocation J_1 for (a) $K = 2, M = 6, N = 3$, (b) $K = 4, M = 8, N = 2$

the proposed method. Moreover, Fig. 4.3 shows that the algorithm convergence is much faster for a system consisting of a smaller number of users.

Method	**Average Hitting Time of Optimal Allocation** $(K = 2, M = 6, N = 3)$, $(K = 4, M = 8, N = 2)$	**Percentage of Optimal Allocation Hits** $(K = 2, M = 6, N = 3)$, $(K = 4, M = 8, N = 2)$
Proposed Solution	$(1.8 \times 10^4), (2.3 \times 10^4)$	$(97), (96)$
UCB	$(9 \times 10^5), (10^7)$	$(17.9), (0.1)$
Random Selection	$(1.8 \times 10^4), (2.4 \times 10^6)$	$(0.25), (2 \times 10^{-4})$

Table 4.1 – Comparison of the average hitting timeslot number and the percentage of hitting the optimal allocation

Table 4.1 compares the average hitting time of the optimal allocation and the percentage of hitting the optimal allocation for different system settings. The hitting time of the optimal allocation is defined as the first time this optimal allocation is played. Table 4.1 also shows the superiority of the proposed method over the UCB algorithm and the random selection technique. In fact, the proposed algorithm has a smaller average hitting time for both simulated system settings. Moreover, the percentage of optimal allocation hits of the proposed method greatly exceeds the percentage of the UCB and the random technique.

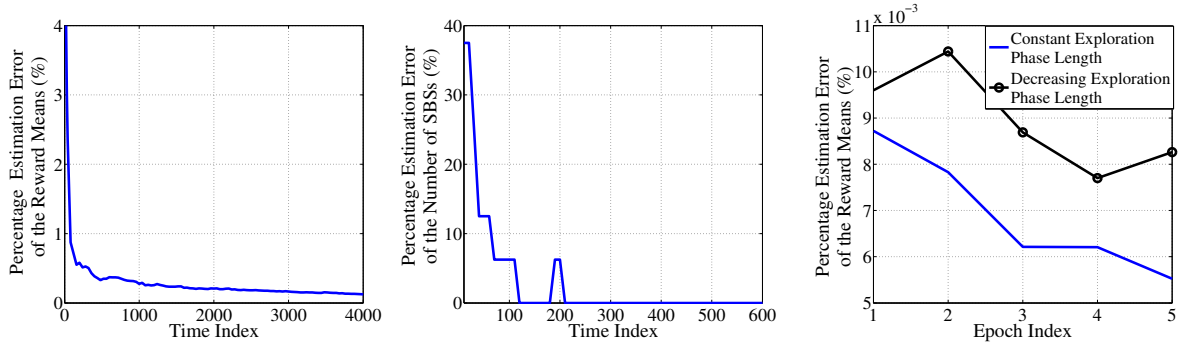
4.8.2 Performance of the Uncoordinated Spectrum Access and Power Control in a SON

To evaluate the performance of Algorithm 8 in the SON setting, the following simulation parameters were chosen: $K = 4, M = 4, N = \beta = L = 2, B_c = 2.5$ MHz, $c_1 = 3000, c_2 = 5000, \epsilon = 5 \times 10^{-5}, \delta = 0$. The available SINR values are $\mathbf{\Gamma} = \{24, 4.77\}$ (dB) leading to achieved rates of 20 and 5 Mbps respectively. For the channel allocation stage, the parameter c used in the matching phase (Cf. Section 4.3.2) is set as: $c = KN$, whereas for the power allocation stage $c = K_m$ for each channel $m \in \mathcal{M}$. Two of the APs are assumed to have a power budget of 1W per channel, while the remaining two have a power budget of 2W per channel. Additional simulation parameters are given in Table 4.2 [103].

Cell Radius R_d	150 m
Overall Transmission Bandwidth	10 MHz
Number of channels	4
Number of APs	4
Power Budget per AP per channel $P^m_{(.)}$	$\{1, 1, 2, 2\}$ (W)
Available SINR Requirements	$\mathbf{\Gamma} = \{24, 4.77\}$(dB)
Distance Dependent Path Loss	$128.1 + 37.6 \log_{10}(d)$(dB), d in Km
Receiver Noise Density	4.10^{-18} mW/Hz

Table 4.2 – Simulation parameters

4.8.2.1 Estimation Accuracy of the Exploration Phase



(a) (b) (c)

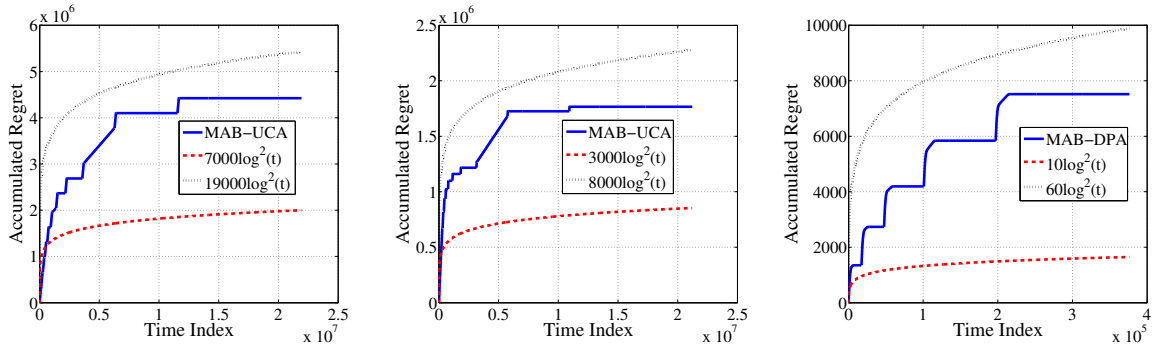
Figure 4.4 – Estimation error as time progresses in the channel allocation stage for (a) the estimation of the rewards, (b) the estimation of the number of APs. (c) Comparison of the estimation error as a function of the epoch index in the channel allocation stage for the estimation of the rewards

First, we evaluate the estimation accuracy of the exploration phase in the channel allocation stage. As shown in Fig. 4.4a and Fig. 4.4b, the estimation of both the reward

means and the total number of APs converges rather quickly to the correct values. Having observed that the estimation of the exploration phase converges quickly, a version of the proposed algorithm where the exploration phase length is divided by the epoch index was tested. The estimation error of this version with a decreasing exploration phase length was compared against the version with a constant exploration phase length. Fig. 4.4c plots the channel rewards estimation error for both versions. Although the constant length version outperforms the version with a decreasing exploration phase length, the estimation error achieved by both methods is lower than $1.1 \times 10^{-2}\%$, hence negligible. When it comes to the estimation of the number of APs, both versions accurately estimate $\hat{K}$, without error, when convergence is reached.

For the power allocation stage, the power level rewards estimation, not shown here for brevity, also converges quickly to a negligible error value.

4.8.2.2 Performance Analysis



(a) (b) (c)

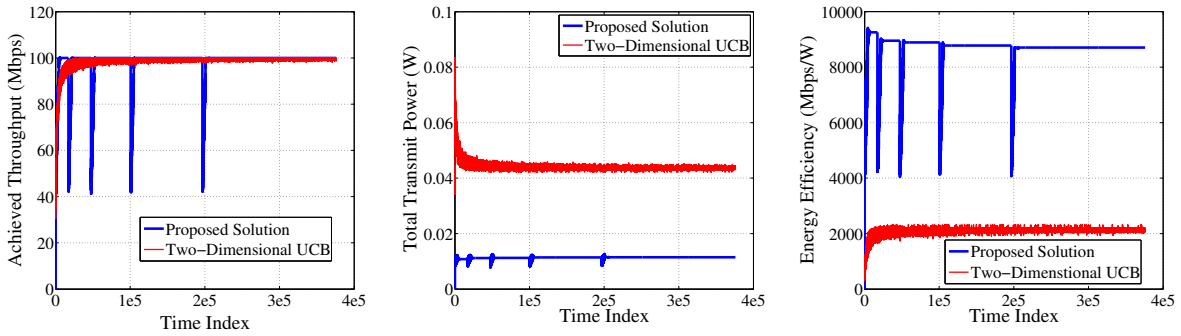
Figure 4.5 – Accumulated regret as time progresses (a) for the channel allocation phase with a constant exploration phase length, (b) for the channel allocation phase with a decreasing exploration phase length, (c) for the power allocation stage

Fig. 4.5 shows the average accumulated regret as a function of time in the channel allocation stage for both the constant and the decreasing length exploration phase versions. The results show that the average accumulated regret for both versions increases with time as $\mathcal{O}(\log(t)^2)$. More specifically, the regret incurred for the constant length exploration phase version is bounded between $7000\log(t)^2$ and $19000\log(t)^2$, as shown in Fig. 4.5a. The regret incurred for the decreasing length exploration phase version is bounded between $3000\log(t)^2$ and $8000\log(t)^2$. In fact, most of the regret is accumulated during the exploration phase where APs choose a channel uniformly at random. Hence, decreasing the length of the exploration phase lowers the value of the accumulated regret as shown in Fig. 4.5b, without jeopardizing the estimation accuracy as was shown in Section 4.8.2.1.

The regret incurred during the power allocation stage is bounded between $10\log(t)^2$ and $60\log(t)^2$, as shown in Fig. 4.5c. The lower regret observed during the power allocation stage, when compared to the channel allocation stage, results from the smaller number of APs competing for a smaller number of arms. In fact, during the power allocation stage, the number of APs competing for each channel $m \in \mathcal{M}$ is $K_m \leq \beta = 2$

while the number of arms is $L = 2$. In contrast, during the channel allocation stage, the number of players is $K = 4$ with $\binom{M}{N} = 6$ available arms.

In Fig. 4.6, we compare the performance of the proposed method with a technique based on the UCB algorithm proposed in [124] and similar to the one proposed in [14], denoted by Two-Dimensional UCB. In the Two-Dimensional UCB method, channel and power allocation are conducted at the same time, using the UCB algorithm, by considering all possible combinations of the channels and the power levels. For the considered setting, the number of arms in the Two-Dimensional UCB method is hence $\binom{M}{N} \times L^N = 24$ arms. In Fig. 4.6a, the achieved rate is plotted as a function of time. Both methods



(a) (b) (c)

Figure 4.6 – Performance comparison as a function of time of (a) the achieved rate, (b) the total transmit power, (c) the energy efficiency

converge relatively quickly to the highest achievable rate, with small variations for the Two-Dimensional UCB technique. The sharp falls in the achieved rate of the proposed method are due to the exploration phase during each epoch of the power allocation stage where APs choose the power levels uniformly at random, causing collisions and leading to zero rates.

The total transmit power used by the APs as a function of time is shown in Fig. 4.6b. While both methods converge to the same highest achievable rate, the power used by our proposed method is significantly lower than the one needed by the Two-Dimensional UCB method. Hence, our proposed method achieves a better allocation for the channel and power when compared to the UCB-based method. Moreover, our proposed method has performance guarantees in terms of regret and optimality, while the Two-Dimensional UCB method [124] does not.

To check the combined effect of rate and power on the performance of the compared methods, the achieved energy efficiency (EE), which is the ratio of the achieved rate to the used power, is plotted in Fig. 4.6c. Once again, the sharp falls in the performance of our proposed method are due to the exploration phase in each epoch of the power allocation stage. Fig. 4.6c shows that our proposed method greatly outperforms the UCB-based method in terms of EE.

4.9 Summary

In this chapter, the uncoordinated channel and power allocation problems in a SON were studied. The considered framework allows each AP to choose N channels at each

timeslot, and allows each channel to simultaneously accommodate multiple APs in a NOMA manner. The considered problem was modeled using the MP-MAB framework, with varying user rewards, multiple plays, and non-zero reward on collision. A game-theoretic approach was used to develop an algorithm with a sub-linear regret of $\mathcal{O}(\log^2 T)$. Simulation results validated the sub-linear regret of the proposed method and showed its superior performance, when compared with one of the most used algorithms in the MAB literature.

4.10 Appendix A

A.1 Proof of Lemma 1

In the channel allocation phase, denote by $\boldsymbol{a}^{(1)}$ the optimal assignment, and by J_M^1 the sum rewards achieved when $\boldsymbol{a}^{(1)}$ is played, which is then given by:

$$J_M^1 = \sum_{k=1}^{K}\sum_{m=1}^{M} a_k^{(1)}(m)\mu_M(k,m,k_m^*). \tag{4.42}$$

Furthermore, denote the second best assignment and the sum reward achieved under it by $\boldsymbol{a}^{(2)}$ and J_M^2 respectively. Let the estimated mean of AP k over channel m with k_m APs on channel m be written as:

$$\hat{\mu}_M(k,m,k_m) = \mu_M(k,m,k_m) + z(k,m,k_m), \tag{4.43}$$

where $z(k,m,k_m)$ is the estimation inaccuracy during the channel allocation phase satisfying $|z(k,m,k_m)| \leq \Delta_M$. The sum reward achieved when $\boldsymbol{a}^{(1)}$ is played with the estimated channel means satisfies:

$$\begin{aligned}\sum_{k=1}^{K}\sum_{m=1}^{M} a_k^{(1)}(m)\hat{\mu}_M(k,m,k_m) = &\sum_{k=1}^{K}\sum_{m=1}^{M} a_k^{(1)}(m)(\mu_M(k,m,k_m) + z(k,m,k_m)) > \\ &\sum_{k=1}^{K}\sum_{m=1}^{M} a_k^{(1)}(m)\mu_M(k,m,k_m) - KN\Delta_M.\end{aligned} \tag{4.44}$$

Any other assignment $\boldsymbol{a} \neq \boldsymbol{a}^{(1)} \neq \boldsymbol{a}^{(2)}$ must perform at most as well as $\boldsymbol{a}^{(2)}$:

$$\begin{aligned}\sum_{k=1}^{K}\sum_{m=1}^{M} a_k(m)\hat{\mu}_M(k,m,k_m) = &\sum_{k=1}^{K}\sum_{m=1}^{M} a_k(m)(\mu_M(k,m,k_m) + z(k,m,k_m)) < \\ &\sum_{k=1}^{K}\sum_{m=1}^{M} a_k^{(2)}(m)\mu_M(k,m,k_m) + KN\Delta_M.\end{aligned} \tag{4.45}$$

To avoid changing the optimal assignment because of the estimation inaccuracy, the following must hold $\forall \boldsymbol{a} \neq \boldsymbol{a}^{(1)}$:

$$\sum_{k=1}^{K}\sum_{m=1}^{M} a_k^{(1)}(m)\hat{\mu}_M(k,m,k_m) > \sum_{k=1}^{K}\sum_{m=1}^{M} a_k(m)\hat{\mu}_M(k,m,k_m). \tag{4.46}$$

To ensure Eq. (4.46), we need to have: $J_M^1 - KN\Delta_M > J_M^2 + KN\Delta_M$, which holds if:

$$\Delta_M < \frac{J_M^1 - J_M^2}{2KN}. \tag{4.47}$$

In the power allocation phase, following a similar approach over each channel m, we get:

$$\Delta_P < \frac{J_P^1 - J_P^2}{2K_m}. \tag{4.48}$$

A.2 Proof of Lemma 2

A.2.1 Lower Bound of the Length of the Exploration Phase in the Channel Allocation Step

To find a lower bound of the length of the exploration phase in the channel allocation step, we first find the required number of observations of each channel by each AP to guarantee condition (4.33) [15, 129]. To do so, the probability of each AP not having a correct estimation of the channel means should be bounded. Let $\gamma = \gamma_{e,l}^M/2$ and let us define the following events:

- A: all players have an estimate satisfying condition (4.33),
- B: all players have $\geq Q$ observations of each channel m for every s in $[\beta]$,
- A_k: player k has an estimate satisfying condition (4.33),
- B_k: player k has $\geq Q$ observations of each channel m for every s in $[\beta]$.

The following must hold:

$$\Pr(\bar{A}_k \mid B_k) \leq \frac{\gamma}{K}. \tag{4.49}$$

In fact,

$$\begin{aligned}
&\Pr(\bar{A}_k|B_k) \leq \Pr\left(\exists\, m, s, \text{s.t. } |\mu_M(k,m,s) - \hat{\mu}_M(k,m,s)| > \Delta_M \mid B_k\right) \overset{(a)}{\leq} \\
&\sum_{m=1}^{M}\sum_{s=1}^{\beta} \Pr\left(|\mu_M(k,m,s) - \hat{\mu}_M(k,m,s)| > \Delta_M \mid B_k\right) = \\
&\sum_{m=1}^{M}\sum_{s=1}^{\beta}\sum_{q=Q}^{\infty} \Pr\left(|\mu_M(k,m,s) - \hat{\mu}_M(k,m,s)| > \Delta_M \,|k \text{ has } q \text{ observations of } (m,s)\right) \times p_2 \overset{(b)}{\leq} \\
&\sum_{m=1}^{M}\sum_{s=1}^{\beta}\sum_{q=Q}^{\infty} 2p_2 e^{(-2q\Delta_M^2)} \leq \sum_{m=1}^{M}\sum_{s=1}^{\beta} 2e^{(-2Q\Delta_M^2)} = 2M\beta e^{(-2Q\Delta_M^2)},
\end{aligned} \tag{4.50}$$

where (m,s) refers to channel m with s players on it, (a) results from applying the union bound and (b) from using Hoeffding's inequality [130], and $p_2 = \Pr\left(q \text{ observations of } (m,s) \mid q \geq Q\right)$.

To ensure $\Pr(\bar{A}_k|B_k)$ is lower than $\frac{\gamma}{K}$, Q must satisfy:

$$Q \geq \frac{1}{2\Delta_M^2}\log\left(\frac{2KM\beta}{\gamma}\right). \tag{4.51}$$

Then,

$$\Pr\left(A|B\right) = 1 - \Pr(\bar{A}|B) \geq 1 - \sum_{k=1}^{K} \Pr\left(\bar{A}_k|B_k\right) = 1 - \gamma, \tag{4.52}$$

leading to all APs having an estimate of every channel satisfying condition (4.33) with probability higher than $1-\gamma$.

Next, we need to find a time horizon T_h for the exploration phase of the channel allocation step large enough such that all players have $\geq Q$ observations of each arm with probability higher than $1-\gamma$. Note that the length of each exploration phase $T_{\hat{\boldsymbol{\mu}}}$ does not necessarily satisfy $T_{\hat{\boldsymbol{\mu}}} \geq T_h$. In other words, all players can get $\geq Q$ observations of each arm with probability higher than $1-\gamma$ after multiple exploration phases.

Let $A_{k,m,s}(t) = 1$ if player k observed channel m with s APs on it at timeslot t, and 0 otherwise. For $0 < \tau < 1$, we have:

$$\Pr\left(\ k \text{ has} \leq (1-\tau)T_h\mathbb{E}[A_{k,m,s}] \text{ observations}\right) = \Pr\left(\sum_{t=1}^{T_h} A_{k,m,s}(t) \leq (1-\tau)T_h\mathbb{E}[A_{k,m,s}]\right) =$$
$$\Pr\left(e^{\left(-d\sum_{t=1}^{T_h} A_{k,m,s}(t)\right)} \geq e^{\left(-d(1-\tau)T_h\mathbb{E}[A_{k,m,s}]\right)}\right) \overset{(a)}{\leq} \frac{\mathbb{E}\left[e^{\left(-d\sum_{t=1}^{T_h} A_{k,m,s}(t)\right)}\right]}{e^{\left(-d(1-\tau)T_h\mathbb{E}[A_{k,m,s}]\right)}}, \tag{4.53}$$

where $d > 0$ and (a) results from applying the Chernoff bound. By noting that all players are randomly and uniformly sampling every channel during the exploration phase, for any $k \in \mathcal{K}, m \in \mathcal{S}, s \in [\beta]$, variables $A_{k,m,s}$ are independent and identically distributed (i.i.d.) across time. Hence:

$$\mathbb{E}\left[e^{\left(-d\sum_{t=1}^{T_h} A_{k,m,s}(t)\right)}\right] = \prod_{t=1}^{T_h} \mathbb{E}\left[e^{\left(-dA_{k,m,s}(t)\right)}\right]. \tag{4.54}$$

Moreover, $A_{k,m,s}(t)$ is a Bernoulli random variable that takes the value 1 with probability p_A. Therefore, we have:

$$\mathbb{E}\left[e^{\left(-dA_{k,m,s}(t)\right)}\right] = 1 + p_A(e^{-d} - 1) \overset{(a)}{\leq} e^{\left(p_A\left(e^{-d}-1\right)\right)}, \tag{4.55}$$

where (a) follows since $1 + y \leq e^y$. Eq. (4.54) can hence be expressed as:

$$\mathbb{E}\left[e^{\left(-d\sum_{t=1}^{T_h} A_{k,m,s}(t)\right)}\right] \leq e^{\sum_{t=1}^{T_h}\left(p_A\left(e^{-d}-1\right)\right)} = e^{\left(T_h\mathbb{E}[A_{k,m,s}]\left(e^{-d}-1\right)\right)}. \tag{4.56}$$

By inserting Eq. (4.56) into Eq. (4.53), we get:

$$\Pr\left(\text{player } k \text{ has} \leq (1-\tau)T_h\mathbb{E}[A_{k,m,s}]\ \right) \leq e^{\left(T_h\mathbb{E}[A_{k,m,s}]\left(e^{-d}-1\right)\right)+\left(d(1-\tau)T_h\mathbb{E}[A_{k,m,s}]\right)}. \tag{4.57}$$

To make the bound as tight as possible, d is chosen such that the right hand side of Eq. (4.57) is minimized, leading to $d = -\log(1-\tau)$. By substituting d by its value in Eq. (4.57), we get:

$$\Pr\left(\text{player } k \text{ has} \leq (1-\tau)T_h\mathbb{E}[A_{k,m,s}]\ \right) \leq e^{\left(-T_h\mathbb{E}[A_{k,m,s}](\tau-(1-\tau)\log(1-\tau))\right)} =$$
$$\left(\frac{e^{-\tau}}{(1-\tau)^{(1-\tau)}}\right)^{\left(T_h\mathbb{E}[A_{k,m,s}]\right)} \overset{(a)}{\leq} e^{-\frac{\tau^2}{2}T_h\mathbb{E}[A_{k,m,s}]}, \tag{4.58}$$

where (a) results from having $(1-\tau)\log(1-\tau) > -\tau + \frac{\tau^2}{2}$, obtained by using a Taylor expansion.

Taking $\tau = 1/2$ and using a union bound on (4.58), we get:

$$\Pr\left(\exists\, k, m, s \text{ s.t. } k \text{ has} \leq \frac{T_h}{2}\mathbb{E}[A_{k,m,s}(t)] \text{ observations}\right) \leq KM\beta e^{\left(\frac{-\frac{1}{4}T_h\mathbb{E}[A_{k,m,s}]}{2}\right)}, \tag{4.59}$$

which is upper bounded by γ if T_h satisfies:

$$T_h \geq \frac{8}{\mathbb{E}[A_{k,m,s}]}\log\left(\frac{KM\beta}{\gamma}\right). \tag{4.60}$$

Moreover, the number of observations of each arm during T_h timeslots, $\sum_{t=1}^{T_h} A_{k,m,s}(t)$, must be at least equal to Q. Hence we need:

$$\sum_{t=1}^{T_h} A_{k,m,s}(t) > \frac{T_h}{2}\mathbb{E}[A_{k,m,s}] \geq Q > \frac{1}{2\Delta_M^2}\log\left(\frac{2KM\beta}{\gamma}\right), \tag{4.61}$$

which holds if:

$$T_h \geq \left\lceil \max\left\{\frac{8}{\mathbb{E}[A_{k,m,s}]}\log\left(\frac{KM\beta}{\gamma}\right),\ \frac{1}{\Delta_M^2\mathbb{E}[A_{k,m,s}]}\log\left(\frac{2KM\beta}{\gamma}\right)\right\}\right\rceil. \tag{4.62}$$

Note that:

$$\begin{aligned}\mathbb{E}[A_{k,m,s}] &= \binom{K-1}{s-1}\left(\frac{1}{M}\right)^s\left(1-\frac{1}{M}\right)^{K-s} \overset{\text{(a)}}{\geq} \left(\frac{1}{M}\right)^s\left(1-\frac{1}{M}\right)^{K-s} \geq \\ &\left(\frac{1}{M}\right)\left(\frac{1}{M}\right)^{s-1}\left(1-\frac{1}{M}\right)^{K-1}\left(1-\frac{1}{M}\right)^{1-s} \overset{\text{(b)}}{\geq} \frac{1}{Me^{\left(\frac{K-1}{M-1}\right)}}(M-1)^{1-s} \overset{\text{(c)}}{\geq} \frac{(M-1)^{1-\beta}}{Me^{\left(\frac{K-1}{M-1}\right)}},\end{aligned} \tag{4.63}$$

where (a) follows from having $\binom{K-1}{s-1} \geq 1$, (b) from $(1-\frac{1}{x})^{x-1} \geq \frac{1}{e}$, and (c) from $s \leq \beta$.

Hence, T_h can be re-written as:

$$T_h \geq \left\lceil \max\left\{\frac{8Me^{\left(\frac{K-1}{M-1}\right)}}{(M-1)^{1-\beta}}\log\left(\frac{KM\beta}{\gamma}\right),\ \frac{Me^{\left(\frac{K-1}{M-1}\right)}}{\Delta_M^2(M-1)^{1-\beta}}\log\left(\frac{2KM\beta}{\gamma}\right)\right\}\right\rceil. \tag{4.64}$$

Having T_h, the probability of all APs having an estimate of the channel means satisfying Eq. (4.33) is lower bounded by:

$$\begin{aligned}\Pr(A) &= 1 - \Pr(\bar{A}) = 1 - \Big(\Pr(\bar{A}|B)\Pr(B) + \Pr(\bar{A}|\bar{B})\Pr(\bar{B})\Big) \\ &\geq 1 - \Big(\Pr(\bar{A}|B)) + \Pr(\bar{B})\Big) \geq 1 - (\gamma+\gamma) = 1 - \gamma_{e,l}^M.\end{aligned} \tag{4.65}$$

Since $\Delta_M = \frac{J_M^1 - J_M^2}{2KN} \leq \frac{KN-0}{2KN} \leq \frac{1}{2}$, Eq. (4.64) is satisfied if:

$$T_h = \frac{2Me^{\left(\frac{K-1}{M-1}\right)}}{\Delta_M^2(M-1)^{1-\beta}}\log\left(\frac{4KM\beta}{\gamma_{e,l}^M}\right). \tag{4.66}$$

Having found the minimum needed length of the exploration epoch in the channel allocation phase, next, we upper bound the error probability in the l^{th} exploration epoch. To do so, we first note that:

$$T_{\hat{\boldsymbol{\mu}}_{\boldsymbol{M}}} \times l = T_h = \frac{2Me^{\left(\frac{K-1}{M-1}\right)}}{\Delta_M^2(M-1)^{1-\beta}}\log\left(\frac{4KM\beta}{\gamma_{e,l}^M}\right). \tag{4.67}$$

To have $\gamma_{e,l}^{M} \leq 4KM\beta e^{-l} \leq 4(M\beta)^2 e^{-l}$, the length of each exploration epoch must satisfy:

$$T_{\hat{\boldsymbol{\mu}}_{\boldsymbol{M}}} \geq \frac{2Me^{\left(\frac{K-1}{M-1}\right)}}{\Delta_M^2 \left(M-1\right)^{1-\beta}}. \tag{4.68}$$

A.2.2 Lower Bound of the Length of the Exploration Phase in the Power Allocation Step

By following a similar analysis of the one in Appendix A.2.1, the minimum length of the length of the exploration phase on each channel m in the power allocation step can be given by:

$$T_P^0 = \left\lceil \frac{2Le^{\left(\frac{\beta-1}{L-1}\right)}}{\Delta_p^2} \right\rceil. \tag{4.69}$$

If the length of the exploration phase in the power allocation step on each channel m satisfies Eq. (4.69), then all players have an estimate of the power level means satisfying the condition in (4.34), with probability $\geq 1-\gamma_{e,l}^{P}$, where $\gamma_{e,l}^{P}$ is upper bounded by $4\beta Le^{-l}$.

A.3 Proof of Lemma 3

Let p be the true probability of player k not being the sole occupier of some channel m when k accesses the M channels uniformly at random:

$$p = 1 - \sum_{m=1}^{M} \frac{1}{M}\left(1-\frac{1}{M}\right)^{K-1} = 1 - \left(1-\frac{1}{M}\right)^{K-1}. \tag{4.70}$$

From Eq. (4.70), the number of users K is given by:

$$K = \text{round}\left(\frac{\log(1-p)}{\log(1-\frac{1}{M})} + 1\right). \tag{4.71}$$

The estimated probability of player k not accessing channel m alone at time t is: $\hat{p}_t = b_k^t/t$. For a correct estimation of the number of APs, we need to find a time t sufficiently large to guarantee with high probability that:

$$\hat{K} = \text{round}\left(\frac{\log(1-\hat{p}_t)}{\log(1-\frac{1}{M})} + 1\right) = \text{round}\left(\frac{\log(1-p)}{\log(1-\frac{1}{M})} + 1\right) = K. \tag{4.72}$$

To ensure Eq. (4.72), if $\kappa < 1/2$, the following must hold:

$$\left|\frac{\log(\frac{t-b_k^t}{t})}{\log(1-\frac{1}{M})} - \frac{\log(1-p)}{\log(1-\frac{1}{M})}\right| = \left|\frac{\log\left(\frac{1-\hat{p}_t}{1-p}\right)}{\log\left(1-\frac{1}{M}\right)}\right| \leq \kappa. \tag{4.73}$$

Let $\hat{p}_t - p = \xi$. After some calculations, Eq. (4.73) can be expressed as:

$$(1-p)\left(1-\left(1-\frac{1}{M}\right)^{-\kappa}\right) \leq \xi \leq (1-p)\left(1-\left(1-\frac{1}{M}\right)^{\kappa}\right). \tag{4.74}$$

With high probability, $K = \hat{K}$ when $\kappa < \frac{1}{2}$, if $|\hat{p}_t - p| \leq \xi_1$, where:

$$\xi_1 = \min\left\{\left|(1-p)\left(1-\left(1-\frac{1}{M}\right)^{-\kappa}\right)\right|,\ \left|(1-p)\left(1-\left(1-\frac{1}{M}\right)^{\kappa}\right)\right|\right\}. \tag{4.75}$$

Let $T_{\hat{K}}$ be a large enough time horizon for which the estimated probability $\hat{p}_{T_{\hat{K}}}$ is an average of i.i.d. random variables with expectation p. Using Hoeffding's inequality [130], we get:

$$\Pr\left(|\hat{p}_{T_{\hat{K}}} - p| \geq \xi_1\right) \leq 2e^{-2T_{\hat{K}}\xi_1^2}. \tag{4.76}$$

To bound the probability of an incorrect estimation of $\hat{K}$ by some small value η, $T_{\hat{K}}$ must be lower bounded by:

$$T_{\hat{K}} \geq \frac{\log(2\eta)}{2\xi_1^2}. \tag{4.77}$$

To get a simpler expression of ξ_1 and hence of $T_{\hat{K}}$, suppose that $\kappa = 0.49$. With the expression of p given by Eq. (4.70), the first term in Eq. (4.75) can be lower bounded as:

$$\begin{aligned}\left|\left(1-\frac{1}{M}\right)^{K-1}\left(1-\left(1-\frac{1}{M}\right)^{-0.49}\right)\right| &= -\left(1-\frac{1}{M}\right)^{K-1}\left(1-\left(1-\frac{1}{M}\right)^{-0.49}\right) \overset{(a)}{\geq} \\ \left(1-\frac{1}{M}\right)^{M\beta-1}\left(1-\left(-1+\frac{1}{M}\right)^{-0.49}\right) &\overset{(b)}{\geq} \frac{1}{e^{\left(\frac{M\beta-1}{M-1}\right)}}\left(1-\left(-1+\frac{1}{M}\right)^{-0.49}\right) \overset{(c)}{\geq} \frac{0.49}{Me^{\left(\frac{M\beta-1}{M-1}\right)}},\end{aligned} \tag{4.78}$$

where (a) results from having $M\beta \geq K$, (b) from $(1-\frac{1}{x})^{x-1} \geq \frac{1}{e}$, and (c) from using a Taylor Expansion. Similarly, the second term in Eq. (4.75) can be lower bounded as:

$$\left|\left(1-\frac{1}{M}\right)^{K-1}\left(1-\left(1-\frac{1}{M}\right)^{0.49}\right)\right| \geq \frac{0.49}{Me^{\left(\frac{M\beta-1}{M-1}\right)}}. \tag{4.79}$$

Variable ξ_1 is therefore lower bounded by:

$$\xi_1 \geq \frac{0.49}{Me^{\left(\frac{M\beta-1}{M-1}\right)}}. \tag{4.80}$$

Hence, $\hat{K} = K$ with probability higher than $1-\eta$ if:

$$T_{\hat{K}} = \left\lceil 2.08\log\left(\frac{2}{\eta}\right)M^2e^{2\left(\frac{M\beta-1}{M-1}\right)}\right\rceil. \tag{4.81}$$

A.4 Proof of Lemma 5

In the matching phase, each AP k keeps a counter of the number of times each action was played and resulted in k being content. At the end of the matching phase, AP k chooses the action that was most played and resulted in k being content in order to play it in the exploitation phase. If the optimal strategy profile $\boldsymbol{a}^*$ was played more than $c_1 l^{1+\delta}/2$ times during matching phase l, then each AP has played the optimal action more than half of the timeslots during the matching phase and will play the optimal action during the exploitation phase. Therefore, we can upper bound the probability of error in the matching phase by the probability of the optimal action being played less than half of the

timeslots in the matching phase of epoch l. Let $f(\boldsymbol{z}) = \mathbb{I}(\boldsymbol{z} = \boldsymbol{z}^*)$, where $\mathbb{I}$ is the indicator function:

$$\delta_{m,l} = \Pr(\boldsymbol{a}^* \neq \boldsymbol{a}^{(l)}) \leq \Pr(\sum_{\tau=1}^{c_1 l^{1+\delta}} f(\boldsymbol{z}(\boldsymbol{\tau})) \leq \frac{c_1 l^{1+\delta}}{2}). \tag{4.82}$$

Let $\theta = 1 - \frac{1}{2\pi(\boldsymbol{z}^*)}$, where $0 \leq \theta \leq 1$ since $\pi(\boldsymbol{z}^*) \geq 1/2$, we have:

$$\delta_{m,l} \leq \Pr(\sum_{\tau=1}^{c_1 l^{1+\delta}} f(\boldsymbol{z}(\boldsymbol{\tau})) \leq \frac{c_1 l^{1+\delta}}{2}) = \Pr\left(\sum_{\tau=1}^{c_1 l^{1+\delta}} f(\boldsymbol{z}(\boldsymbol{\tau})) \leq (1-\theta)\pi(\boldsymbol{z}^*) c_1 l^{1+\delta}\right). \tag{4.83}$$

Using the concentration bound of Theorem 3 in [69], the above probability can be bounded by:

$$\delta_{m,l} \leq \Pr\left(\sum_{\tau=1}^{c_1 l^{1+\delta}} f(\boldsymbol{z}(\boldsymbol{\tau})) \leq (1-\theta)\pi(\boldsymbol{z}^*) c_1 l^{1+\delta}\right) \leq A_0 \left\|\phi\right\|_\pi \exp\left(\frac{-\theta^2 \pi(\boldsymbol{z}^*) c_1 l^{1+\delta}}{72 T_m(1/8)}\right), \tag{4.84}$$

where A_0 is a constant, $\left\|\phi\right\|_\pi$ is the distribution of the initial state of the matching phase at the l^{th} epoch, and $T_m(1/8)$ is the mixing time of the Markov chain with an accuracy of $1/8$.

$\delta_{m,l}$ can be further upper bounded as:

$$\delta_{m,l} \leq A_0 \left\|\phi\right\|_\pi \exp\left(\frac{-\theta^2 \pi(\boldsymbol{z}^*) c_1 l^{1+\delta}}{72 T_m(1/8)}\right) \leq A_1 \exp(-l^{1+\delta}). \tag{4.85}$$

Chapter 5

UAV-Enabled Communication Networks

5.1 Introduction

When used as flying base stations (BSs), unmanned aerial vehicles (UAVs) can help increase the throughput and the coverage of traditional communication systems thanks to their mobility, flexibility and low cost [131]. Moreover, they can help alleviate traffic congestion in hotspot areas and establish communication links in remote and disaster areas, where the communication infrastructure is either non-existent or damaged [132]. Therefore, their use in wireless communication systems has received a lot of attention in recent literature [42].

Contrary to most terrestrial BSs, UAVs are connected to the core network through a wireless backhaul link [132]. That said, to reap the benefits promised by the use of UAV-BSs, this backhaul link needs to be dynamically managed according to the traffic state of the network [42]. In fact, if not configured properly, the backhaul link introduces interference, thus limiting the throughput provided by the UAV in the access link (i.e., the link between the UAV and the served users). In-band full-duplex (IBFD) communications were recently investigated to increase the spectral efficiency and reduce the latency of systems relying on a wireless backhaul link [133]. When adopting an in-band wireless backhaul link, the same frequency band is used in the access and the backhaul links, thus optimizing the system spectral efficiency. Additionally, the deployment of in-band wireless backhauling is of great interest to cellular operators, since networks can be upgraded in a short time and a cost-effective way [134]. IBFD allows the simultaneous transmission and reception of backhaul and access information in the same frequency band, at the expense of a self-interference (SI), induced by the transmitter on its own receiver. Nonetheless, SI cancellation schemes have progressed significantly [135], allowing an efficient application of IBFD for wireless backhauling. For example, [135] reported a 150 dB cancellation of SI by a closed-loop echo cancellation technique.

Different frequency bands are used to achieve the wireless backhaul connectivity. In fact, a wireless backhaul link might use, among others, sub-6 GHz spectrum, millimeter wave (mmWave) or satellite spectrum to achieve the required backhaul connectivity. Although mmWave spectrum offers high capacity from an abundance of available spectrum, numerous challenges must be accounted for when considering it for the backhaul link. In a nutshell, mmWaves suffer from a high atmospheric attenuation [134], are highly sensi-

tive to shadowing and have poor penetration capabilities through obstacles, constraining them to operate over line-of-sight (LOS) links. Moreover, a high symbol rate with an order of magnitude increase in carrier frequency makes such systems particularly prone to poor propagation characteristics, high frequency selectivity, hardware impairments, and Doppler effects, especially in a mobile scenario. To improve the link quality of mmWave transmissions, highly directive antennas, with frequent repointing, would have to be used [134, 136]. In contrast, sub-6 GHz spectrum is simpler to deploy and offers several advantages. From the operator point of view, no additional spectrum and no new hardware are required. Sub-6 GHz spectrum ensures a wider coverage due to its low attenuation and does not require antenna alignment. That said, sub-6 GHz spectrum is scarce and its use in the backhaul link results in additional interference, evoking the need for appropriate resource allocation techniques to increase the system spectral efficiency and avoid spectrum miss-use.

In addition to UAV-enabled networks, non-orthogonal multiple access (NOMA) has been recently proposed as a promising solution to enhance the performance of future communication networks [75, 76, 137]. In Power-Domain NOMA, multiple users are scheduled on the same time-frequency resource by multiplexing their signals in the power domain. At the receiver side, successive interference cancellation (SIC) is performed to retrieve the superimposed signals. By allowing multiple users to access the same resource, NOMA enhances spectral efficiency and user fairness.

The objective of the current chapter is to propose a complete solution for the resource allocation and positioning problems in a UAV-enabled network benefiting from both NOMA and full-duplex (FD) communications. The proposed solution aims at serving users with their rate requirements, all the while ensuring the UAV receives the necessary capacity in the backhaul link to be able to satisfy users. The work presented in this chapter has been presented at the IEEE Vehicular Technology Conference Fall 2019 [126] and an extended journal version has been published in the IEEE Transactions on Vehicular Technology in June 2020 [110]. Next, we give an overview of the related work on UAV-aided communication networks.

5.1.1 Background

The optimization of the UAV deployment has been given significant attention in recent literature. In [138], the authors investigated the optimal location of the UAV to minimize its transmit power. The optimal UAV altitude that maximizes its coverage region was evaluated in [63]. In [139], the authors built on the results of [63] and found the 3D location of the UAV that maximizes the coverage for users having different quality of service (QoS) requirements. However, none of these works accounted for the wireless backhaul connectivity of the UAV.

The backhaul connectivity of UAV-enabled networks was discussed in a few recent works. The 3D placement problem of the UAV was investigated in [132] for the sake of maximizing the number of served users while considering a backhaul link with constant transmission rate. The authors in [140] introduced a heuristic algorithm that finds the number of needed UAVs as well as their 3D positions and accounts for the backhaul constraint. However, the authors did not elaborate on the way bandwidth assignment in the backhaul link is conducted. In [25], the authors proposed an algorithm to find the 3D position of the UAVs as well as the user and bandwidth allocation to maximize the loga-

rithmic rates of users. Nevertheless, [25] assumed that access and backhaul transmissions take place on different and sufficiently spaced frequency bands to avoid SI.

Resource allocation for NOMA systems has been extensively studied with different performance measures. For example, the weighted sum rate of a FD-NOMA system was maximized in [89]; however the proposed method has exponential complexity. Maximizing system fairness was the target of [90], while minimizing the transmit power subject to rate requirements was targeted in [91] and [92].

The use of NOMA in UAV-enabled networks was given some attention in a few recent papers. In [141], the macro base station (MBS) performed backhaul transmission to dispatched UAVs using NOMA, with the UAVs cooperating to increase the data rates and reliability of served users. However, the study considered a half-duplex (HD) system, as the backhaul and access transmissions occurred in different timeslots. Moreover, NOMA was only considered for the transmission in the backhaul link. In [142], system performance of a NOMA UAV-enabled system was evaluated using stochastic geometry, and a reinforcement learning approach for the placement and movement of the UAV was proposed. The authors in [143] considered a NOMA-UAV system and maximized the minimum rate achieved with respect to the UAV altitude, beamwidth, power and bandwidth allocation. In [144], the sum rate of the users served by the UAV and the BS was maximized through the optimization of the user scheduling, the UAV trajectory and the precoding at the BS level. However, neither [142], [143] nor [144] considered the backhaul limitation of the UAV.

5.1.2 Problem Statement and Contributions

In this chapter, we consider a scenario where the traditional wireless infrastructure is missing, e.g., in remote areas or due to a disaster or BS failure. An FD-UAV is dispatched to serve users having QoS requirements in that area, and an in-band wireless backhaul link is established between the UAV and an MBS to provide the needed backhaul capacity. To minimize the UAV transmit power, an optimization problem that finds the assignment of subbands in the backhaul link, the 3D position of the UAV as well as the power levels in the access and backhaul links is solved. Moreover, when the UAV power budget is not sufficient to guarantee user rate requirements, a NOMA pairing step is conducted to maximize the achievable rates.

The main contributions of this chapter can be summarized as follows:

- We formulate and solve a feasibility test that checks if the user rate requirements can be simultaneously met with the UAV power budget.
- An optimization problem that minimizes the UAV transmit power is formulated. This problem takes into account the rate requirement per user, the backhaul constraint, and the transmit power budget constraints for the UAV and the MBS.
- A novel framework is introduced to find the best bandwidth assignment in the backhaul link, as well as the region in space in which the UAV can be deployed. Moreover, the number and the assignment of subbands in the backhaul link are optimized.
- After backhaul subband assignment, we determine the UAV position within the identified region and the power variables that lead to the minimum needed UAV

power.

- When the UAV power budget cannot satisfy all users simultaneously, a NOMA pairing algorithm is proposed in order to maximize the achieved sum rate and number of satisfied users.

The rest of this chapter is organized as follows. In Section 5.2, the system model and the problem formulation are presented. Section 5.3 introduces the concept of coverage region of each user, a concept that will help formulating the proposed solution presented in Section 5.4. Section 5.5 presents a simplified solution to the UAV positioning and resource allocation problems that does not benefit from NOMA scheduling. The performance of the proposed technique is evaluated in Section 5.6 through simulation results. Finally, a summary of this chapter is given in Section 5.7.

5.2 System Model

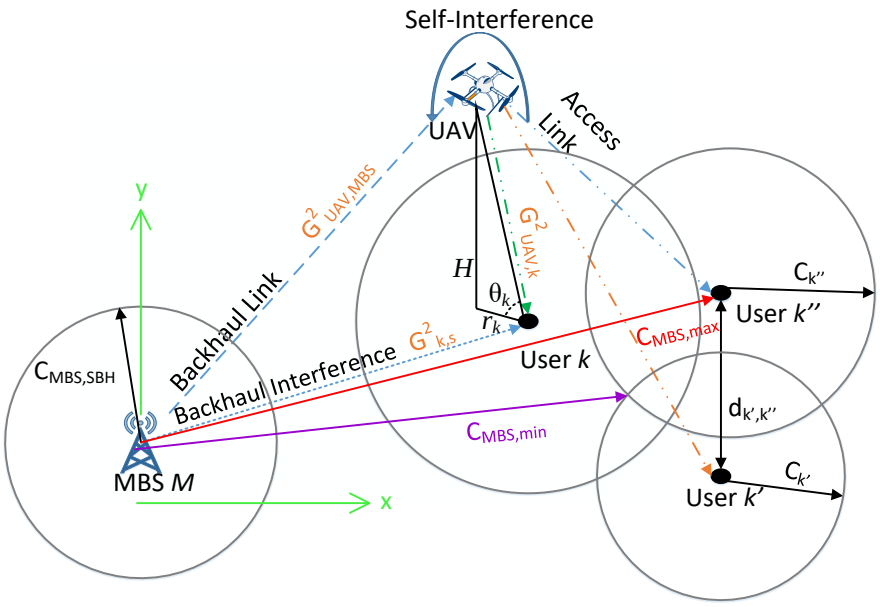


Figure 5.1 – System Model

The downlink UAV-enabled communication network consists of one IBFD-enabled UAV and K randomly deployed users, as shown in Fig. 5.1. Note that the elements of Fig. 5.1 will be detailed as the chapter progresses. The UAV is connected to the core network through an MBS M, located at the origin, via wireless in-band backhauling on the sub-6 GHz spectrum. The set of users is denoted by $\mathcal{K}$, with the horizontal position of each user $k \in \mathcal{K}$ denoted by (x_k, y_k). Each user $k \in \mathcal{K}$ is associated with a rate requirement R_k^{req}. Let the UAV position be denoted by $\boldsymbol{z_{UAV}} = (x_{UAV}, y_{UAV}, H)$, where (x_{UAV}, y_{UAV}) refers to the horizontal position of the UAV, while H refers to its altitude. Clearly, the UAV should be positioned to serve as many users as possible, while being able to receive the necessary rate from the backhaul link.

The IBFD wireless backhaul assumption leads the backhaul and the access links of the UAV to share the same frequency. Hence, a user associated with the UAV suffers from the interference of the backhaul link occurring in the same frequency band. Moreover, due to its FD nature, the UAV suffers from the SI term, caused by its own transmitter on its receiver.

5.2.1 Path Loss Model

The widely adopted air-to-ground (A2G) communication model [62] is considered between the UAV and the users, with two LOS and non-line-of-sight (NLOS) path-loss models. The probability of having a LOS communication link between the UAV and user k is given by:

$$P_{LOS} = \frac{1}{1 + \alpha \exp\left(-\beta\left(\frac{180\theta_k}{\pi} - \alpha\right)\right)}, \tag{5.1}$$

where $\theta_k = \arctan\left(\frac{H}{r_k}\right)$ is the elevation angle, $r_k = \sqrt{(x_k - x_{UAV})^2 + (y_k - y_{UAV})^2}$ is the horizontal distance between user k and the UAV, whereas α and β are constants determined by the environment (e.g., rural, urban, suburban). The NLOS probability is hence: $P_{NLOS} = 1 - P_{LOS}$.

Since classifying a link as LOS or NLOS is not straightforward and requires terrain knowledge, the probabilistic mean path loss is adopted [139]:

$$\begin{aligned} L(H, r_k) &= \eta_L \times P_{LOS} + \eta_{NL} \times P_{NLOS} + 20\log\left(\frac{4\pi f_c d_{UAV,k}}{c}\right) \\ &= \frac{A}{1 + \alpha \exp\left(-\beta\left(\frac{180\theta_k}{\pi} - \alpha\right)\right)} + 20\log(\frac{r_k}{\cos(\theta_k)}) + B, \end{aligned} \tag{5.2}$$

where $20\log\left(\frac{4\pi f_c d_{UAV,k}}{c}\right)$ is the free space path loss, f_c and c being the carrier bandwidth and the speed of light respectively. Variables η_L and η_{NL} are the mean additional losses for LOS and NLOS links, respectively whereas $d_{UAV,k} = \sqrt{(r_k^2 + H^2)}$ is the 3D distance between the UAV and user k. Moreover, $A = \eta_L - \eta_{NL}$ and $B = 20\log\left(\frac{4\pi f_c}{c}\right) + \eta_{NL}$.

5.2.2 Communication Model

We consider the sub-6 GHz spectrum where the system bandwidth BW is equally partitioned into a set $\mathcal{S}$ of S subbands, leading to a subband bandwidth of $B_c = BW/S$. In this study, it is assumed that the number of users is equal to the number of subbands, i.e., $K = S$. Let subband s be the subband allocated to user k and let $G_{k,s}$ be the channel gain between the MBS and each user k over subband s. The channel gain $G_{k,s}$ consists of both small-scale and large-scale fading. The rate achieved by user k over subband s in the access link is given by [140]:

$$R_{k,s} = B_c \log_2(1 + \frac{P_{UAV,k,s}\, G_{UAV,k}^2}{N_0 B_c + b_s\, P_{MBS,s}\, G_{k,s}^2}), \tag{5.3}$$

where $P_{UAV,k,s}$ is the transmit power allocated by the UAV to user k on subband s, while $P_{MBS,s}$ is the transmit power of the MBS on subband s. The noise power spectral density

is N_0. The binary decision variable $b_s \in \{0,1\}$ represents the subband assignment in the backhaul link. In other words, $b_s = 1$ if subband s is used in the backhaul link, and 0 otherwise. Moreover, $P_{MBS,s}G^2_{k,s}$ denotes the backhaul interference (BI) suffered by user k when subband s is used in the backhaul link, i.e., when $b_s = 1$. The channel gain between the UAV and user k is denoted by $G^2_{UAV,k}$ and given by:

$$G^2_{UAV,k} = 10^{-L(H,r_k)/10}. \tag{5.4}$$

The channel between the UAV and the MBS also follows the A2G path-loss model. When subband s is used to provide backhaul information to the UAV, i.e., when $b_s = 1$, the rate achieved by the UAV over s is given by:

$$R_{UAV,s} = B_c \log_2(1 + \frac{P_{MBS,s}\, G^2_{UAV,MBS}}{N_0 B_c + C_{SI} \sum\limits_{k\in\mathcal{K}} P_{UAV,k,s}}). \tag{5.5}$$

In (5.5), the residual SI experienced at the UAV on subband s is given by $C_{SI}\sum_k P_{UAV,k,s}$, with $1/C_{SI}$ being the SI cancellation factor. Moreover, $G^2_{UAV,MBS}$ is the channel gain between the UAV and the MBS given by:

$$G^2_{UAV,MBS} = 10^{-L(H,r_{MBS})/10}, \tag{5.6}$$

where $L(H, r_{MBS})$ is the mean path-loss between the UAV and the MBS, r_{MBS} being the horizontal distance between the UAV and the MBS.

5.2.3 Problem Formulation

UAV communications being energy-limited, the main purpose of this study is to minimize the transmit power of the UAV while meeting the rate requirements of all users. To achieve this purpose, the positioning of the UAV along with the subband and power allocation in the access and the backhaul links are studied. Therefore, the following optimization problem is formulated:

$$\min_{\substack{\boldsymbol{P_{UAV}},\boldsymbol{P_{MBS}},\\ \boldsymbol{z_{UAV}},\boldsymbol{a},\boldsymbol{b}}} \sum_{k\in\mathcal{K}}\sum_{s\in\mathcal{S}} a_{k,s}\, P_{UAV,k,s} \tag{5.7}$$

$$\text{such that } \sum_{s\in\mathcal{S}} a_{k,s}\, R_{k,s} \geq R^{\text{req}}_k,\ \forall\, k \in \mathcal{K}, \tag{5.7a}$$

$$\sum_{k\in\mathcal{K}}\sum_{s\in\mathcal{S}} a_{k,s}\, R_{k,s} \leq \sum_{s\in\mathcal{S}} b_s\, R_{UAV,s}, \tag{5.7b}$$

$$\sum_{k\in\mathcal{K}}\sum_{s\in\mathcal{S}} P_{UAV,k,s} \leq P^{\max}_{UAV}, \tag{5.7c}$$

$$\sum_{s\in\mathcal{S}} b_s\, P_{MBS,s} \leq P^{\max}_{MBS}, \tag{5.7d}$$

$$H_{\min} \leq H \leq H_{\max}, \tag{5.7e}$$

$$a_{k,s},\, b_s \in \{0,1\}. \tag{5.7f}$$

In (5.7), $a_{k,s}$ is a binary variable that is equal to 1 if user k is scheduled on subband s in the access link and 0 otherwise. Constraint (5.7a) is the rate requirement constraint for each user $k \in \mathcal{K}$. Moreover, constraint (5.7b) is the backhaul data rate constraint which

ensures that the total rate delivered by the UAV to users does not exceed its backhaul capacity. In (5.7c) and (5.7d), $P_{UAV}^{\max}$ and $P_{MBS}^{\max}$ are the transmit power budgets of the UAV and the MBS respectively. Finally, (5.7e) is the altitude constraint for the UAV.

The optimization problem in (5.7) consists of three subproblems: the UAV placement problem, as well as the problems of bandwidth and power allocation in the access and backhaul links. One can see that, at the optimum, constraint (5.7a) is satisfied with equality for all users. Hence, the left hand side of constraint (5.7b) can be substituted with $\sum_k R_k^{\text{req}}$. Nonetheless, the resulting formulated optimization problem is mixed-integer and multivariate. To solve it, we propose a multi-step algorithm that targets the three subproblems.

Once the subband assignment in the access link, $\boldsymbol{a}$, is known, solving the optimization problem (5.7) resorts to finding the backhaul subband assignment $\boldsymbol{b}$, the UAV position $\boldsymbol{z_{UAV}}$ and the power variables $\boldsymbol{P}_{UAV}$ and $\boldsymbol{P}_{MBS}$ minimizing the UAV transmit power. Therefore, the proposed solution proceeds as follows:

1. Perform the subband assignment in the access link.
2. Solve a feasibility test to check whether $P_{UAV}^{\max}$ can accommodate the rate requirements without considering the BI.
3. Decide on the assignment of backhaul subbands. To do so:
   (a) Determine the minimum number of required backhaul subbands. This initial value does not take into account the impact of the resulting BI at the user side.
   (b) Retain the backhaul subband assignment minimizing the needed UAV power and accounting for the resulting BI at the user side.
4. Find $\boldsymbol{z_{UAV}}, \boldsymbol{P_{UAV}}$, and $\boldsymbol{P_{MBS}}$.
5. If needed, perform a NOMA pairing step to further improve achieved user data rates.

5.3 Coverage Region of Each User and of the MBS

To be able to serve user k with its rate requirement when the latter is scheduled on subband s, the UAV position as well as the power allocated to subband s in the access and the backhaul links must be optimized. For each value of the power in the access link $P_{UAV,k,s}$ and the power in the backhaul link $P_{MBS,s}$, the rate requirement of user k in constraint (5.7a) can be translated into a requirement on the maximum tolerable path-loss. In fact, to guarantee the rate requirement of user k, from Eq. (5.3), the channel gain between the UAV and user k, $G_{UAV,k}^2$, must satisfy:

$$G_{UAV,k}^2 \geq \left(2^{\frac{R_k^{\text{req}}}{B_c}} - 1\right) \frac{N_0 B_c + b_s P_{MBS,s} G_{k,s}^2}{P_{UAV,k,s}}. \tag{5.8}$$

For Eq. (5.8) to hold, from Eq. (5.4), the path-loss experienced by user k in the access link, $L(H, r_k)$, should satisfy:

$$L(H, r_k) \leq L_k^{\text{th}} = -10 \log_{10} \left\{ \left(2^{\frac{R_k^{\text{req}}}{B_c}} - 1\right) \frac{N_0\, B_c + b_s\, P_{MBS,s}\, G_{k,s}^2}{P_{UAV,k,s}} \right\}, \tag{5.9}$$

where L_k^{th} is the maximum tolerable path-loss by k to reach R_k^{req}. In other words, user k meets its rate requirement if the experienced path-loss between user k and the UAV, $L(H, r_k)$, is less than or equal to the maximum tolerable path-loss L_k^{th}.

According to Eq. (5.2), the path-loss experienced by user k depends on the relative 3D distance between user k and the UAV through its dependence on the elevation angle between k and the UAV, θ_k, and on the 2D distance between user k and the UAV, r_k. Moreover, the expression of the path-loss (5.2) is implicit, as neither r_k nor H can be written as an explicit function of one another [63]. Using Eq. (5.2) and Eq. (5.9), we need the following to hold:

$$\begin{aligned} L(H, r_k) &= \frac{A}{1 + \alpha \exp\left(-\beta\left(\frac{180\theta_k}{\pi} - \alpha\right)\right)} + 20\log(\frac{r_k}{\cos(\theta_k)}) + B \\ &\leq L_k^{\text{th}} = -10 \times \log_{10}\left(\left(2^{R_k^{\text{req}}/Bc} - 1\right)\frac{N_0 B_c + b_s P_{MBS,s} G_{k,s}^2}{P_{UAV,k,s}}\right). \end{aligned} \tag{5.10}$$

From Eq. (5.10), for a given value of the altitude H, if the UAV is located at a horizontal distance r_k from user k, the latter experiences a constant path loss $L(H, r_k)$. Moreover, for the same altitude H, if the UAV is located at a distance $r \leq r_k$ from user k, the latter experiences a path loss $L(H, r) \leq L(H, r_k)$. Recall that for user k to reach its rate requirement, the maximum tolerable path-loss between k and the UAV is L_k^{th}. For a UAV altitude H, let $C_k(H)$ be the maximum 2D distance between the UAV and user k guaranteeing this maximum tolerable path-loss, i.e.,:

$$C_k(H) = r \,|_{L_{(H,r)} = L_k^{\text{th}}}. \tag{5.11}$$

From Eq. (5.10), as long as the horizontal distance between user k and the UAV r_k is lower than $C_k(H)$, the rate requirement of user k is guaranteed. This translates into a coverage region for user k that is a disk centered at user k and having $C_k(H)$ as radius. Put differently, if the UAV is positioned at an altitude H and in a coverage disk having user k as center and $C_k(H)$ as radius, user k experiences a path loss lower than or equal to L_k^{th}, hence guaranteeing its rate requirement. In Fig. 5.2, the path-loss experienced by a randomly positioned user k as a function of the 2D distance separating it from the UAV, for a fixed UAV altitude, is plotted. Fig. 5.2 shows that while the 2D distance between user k and the UAV satisfies $r_k \leq C_k(H)$, the path-loss experienced by user k is lower than the path-loss threshold L_k^{th}, guaranteeing its rate requirement.

The study in [63] focused on finding the optimal UAV altitude guaranteeing the maximum value of the 2D distance $C_k(H)$ between user k and the UAV, for a certain maximum tolerable path-loss L_k^{th}. The difference between our study and that of [63] is that in [63], the UAV is positioned at the center of the disk having $C_k(H)$ as radius, with all users positioned at a distance lower than $C_k(H)$ from the UAV, and having the maximum tolerable path-loss L_k^{th}, assumed to be covered by the UAV. In this chapter, we propose to make use of the inverse relation which equally holds, since the channel between the UAV and the users is assumed symmetrical. To find the optimal UAV altitude leading to the maximum accepted 2D distance $C_k(H)$, [63] solved the following equation:

$$\frac{\partial r_k}{\partial H} = 0. \tag{5.12}$$

First, it was noted that, for a fixed maximum tolerable path-loss L_k^{th}, the coverage radius of the UAV as a function of its altitude H has only one optimum point corresponding

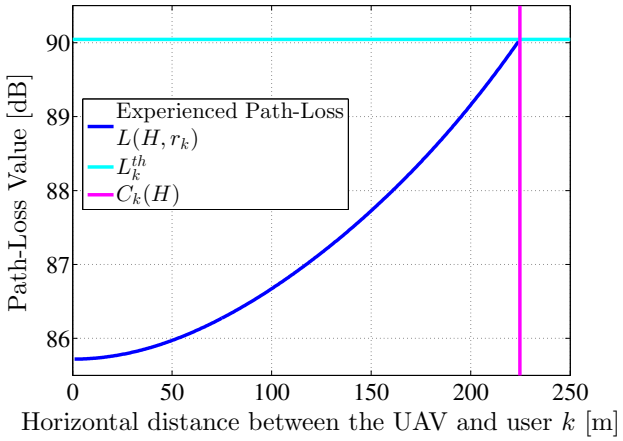


Figure 5.2 – Path-loss experienced by user k in terms of the 2D distance with the UAV r_k.

to the maximum value of the coverage radius. Drawing parallels, in our work, from the perspective of user k, the coverage radius $C_k(H)$ has one maximum value denoted as C_k. As shown in Fig. 5.3, where the coverage radius $C_k(H)$ in terms of H is sketched, the maximum value of $C_k(H)$ depends on the environment type. Finding the optimum point

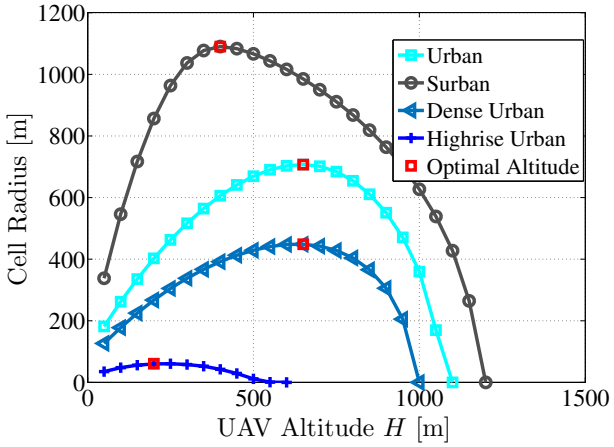


Figure 5.3 – Cell Radius for different propagation environments as a function of the UAV altitude H for $L_k^{th} = 100$ dB.

can be done numerically since $C_k(H)$ is an implicit function of the altitude H. To solve Eq. (5.12), both [63] and [145] solved:

$$\frac{\partial r_k}{\partial H} = \frac{\partial r_k}{\partial \theta_k}\frac{\partial \theta_k}{\partial H} = 0. \tag{5.13}$$

Moreover, the following holds:

$$\frac{\partial \theta_k}{\partial H} = \frac{\partial \tan^{-1}(\frac{H}{r_k})}{\partial H} = \frac{r_k}{H^2 + r_k^2} > 0. \tag{5.14}$$

Hence, solving for the value of H that maximizes the coverage radius can be achieved by searching for the value of the elevation angle θ_k satisfying $\frac{\partial r_k}{\partial \theta_k} = 0$. Denote this value by θ_{opt}. Finding θ_{opt} comes down to solving the following equation numerically [63]:

$$\frac{\pi}{9\ln(10)}\tan(\theta_{\text{opt}}) + \frac{\alpha\beta A \exp\left(-\beta\left(\frac{180}{\pi}\theta_{\text{opt}} - \alpha\right)\right)}{\left(\alpha\exp\left(-\beta\left(\frac{180}{\pi}\theta_{\text{opt}} - \alpha\right)\right) + 1\right)^2} = 0. \tag{5.15}$$

It can be clearly seen from (5.15) - and this was also noted in [63] and [145] - that the optimal value of the elevation angle, θ_{opt}, depends only on the propagation environment. Solving (5.15) yields: $\theta_{\text{opt}} = 20.34°, 42.44°, 54.62°$ and $75.52°$ for the suburban, urban, dense urban and high-rise urban environments, respectively.

Having found the value of the optimal elevation angle θ_{opt}, the maximum coverage radius C_k can be found using Eq. (5.2) as follows:

$$C_k = \cos(\theta_{\text{opt}}) \times 10^{L_k^{\text{th}}/20} \times 10^{-D/20} = E \times 10^{L_k^{\text{th}}/20}, \tag{5.16}$$

where

$$\begin{cases} D = \frac{A}{1+\alpha\exp(-\beta(180\theta_{\text{opt}}/\pi - \alpha))} + B, \\ E = \cos(\theta_{\text{opt}}) \times 10^{-D/20}. \end{cases} \tag{5.17}$$

Hence, we propose to associate every user k with a 2D coverage region of radius C_k (cf. Fig. 5.1) that guarantees its rate requirement, given the power values in the access and backhaul links. Let D_k be the disk centered at user k and of radius C_k. It should be noted that C_k will only be used to find the acceptable 2D deployment region for the UAV where all users can be served simultaneously.

Similarly, when considering the backhaul link, the UAV should be positioned in a region where it is able to receive the required backhaul rate from the MBS. Let S_{BH} be the value of the total number of backhaul subbands. To meet the required backhaul rate, from Eq. (5.5), the path-loss between the UAV and the MBS should satisfy:

$$\begin{aligned} L(H, r_{MBS}) \leq L_{S_{BH}}^{\text{th}} = -10\log_{10}\Big\{\left(2^{R_{UAV,s}/B_c} - 1\right) \times \\ \frac{N_0 B_c + C_{SI}\sum_{k\in\mathcal{K}} P_{UAV,k,s}}{P_{MBS,s}}\Big\}. \end{aligned} \tag{5.18}$$

Note that Eq. (5.18) depends on the value of the total number of backhaul subbands, S_{BH}, through both the rate and the power allocated by the MBS to each subband used in the backhaul link, $R_{UAV,s}$ and $P_{MBS,s}$, respectively.

As in the users case, the maximum coverage region of the MBS satisfying the backhaul path-loss requirement is found using Eq. (5.16) after replacing L_k^{th} with $L_{S_{BH}}^{\text{th}}$. Being a function of the number of backhaul subbands, the maximum coverage radius of the MBS is denoted by $C_{MBS,S_{BH}}$. Let $D_{MBS,S_{BH}}$ be the disk centered at the MBS with $C_{MBS,S_{BH}}$ as radius.

5.4 Holistic Solution to the Positioning and Resource Allocation Problems in UAV-Enabled Networks

5.4.1 Preliminaries

First, the subband assignment in the access link is discussed before detailing the initialization steps.

5.4.1.1 Subband Assignment in the Access Link

As per Eq. (5.3), to reach the users required rates while achieving the objective of problem (5.7), the subband assignment in the access link should take into account:

- maximizing the average channel gain between the UAV and its users, i.e., maximizing $G_{UAV,k}, \forall\, k$ in $\mathcal{K}$,
- minimizing the average BI experienced by the users, i.e., minimizing the average channel gain between the MBS and each user, $G_{k,s}, \forall k$ in $\mathcal{K}$.

Since the adopted A2G channel model [62] does not consider frequency selectivity, the channel gain between the UAV and user k, $G_{UAV,k}$ is constant for every subband $s \in \mathcal{S}$. Hence, the value of the channel gain between the UAV and each user k, $G_{UAV,k}$, does not affect the optimization of problem (5.7) from the perspective of subband assignment in the access link. Consequently, achieving the objective of problem (5.7) comes down to minimizing the average experienced BI by minimizing the average channel gain between the MBS and each user on its allocated subband. Therefore, we propose to solve the following optimization problem:

$$\min_{\boldsymbol{a}} \sum_{k\in\mathcal{K}} \sum_{s\in\mathcal{S}} a_{k,s} G_{k,s} \tag{5.19}$$

$$\text{such that } \sum_{s\in\mathcal{S}} a_{k,s} = 1,\ \forall\, k \in \mathcal{K}. \tag{5.19a}$$

Problem (5.19) aims at finding the assignment $\boldsymbol{a}$ that minimizes the average channel gain between users and the MBS, hence the average BI. Constraint (5.19a) restricts each user to be assigned one subband only. Being an assignment problem, problem (5.19) can be solved efficiently using the Hungarian method [146].

5.4.1.2 Feasibility Test

Recall that the 2D coverage regions guaranteeing (5.7a) for each user and (5.7b) in the backhaul link are found in Section 5.3. Next, a feasibility test that checks if the available UAV power is sufficient for guaranteeing the intersection of the different coverage regions is performed. If positive, the number of backhaul subbands minimizing the needed UAV power when accounting for the BI is then found. In the opposite case, problem (5.7) is still solved while initially relaxing constraint (5.7c). After finding the UAV position, a subsequent NOMA pairing step is then performed to maximize the achieved rates while enforcing back (5.7c).

The feasibility test that checks if the UAV power budget $P_{UAV}^{\max}$ is sufficient to guarantee the rate requirements of all user is done when considering that none of the subbands is used in the backhaul link, i.e., $b_s = 0$, for every subband $s \in \mathcal{S}$. If the test is negative without considering any BI, the system is infeasible under any combination of used backhaul subbands because of the added interference. A negative feasibility test means that the current rate requirements cannot be met simultaneously with the available UAV power.

Let s_k be the subband assigned to user k in the access link. When replacing L_k^{th} in Eq. (5.16) by its expression from Eq. (5.9), and by noting that:

$$10^{L_k^{\text{th}}/20} = \frac{\sqrt{P_{UAV,k,s_k}}}{\sqrt{\left(2^{R_k^{\text{req}}/B_c} - 1\right) N_0 B_c}} \tag{5.20}$$

in the absence of BI, the maximum value of the coverage region, C_k, can be formulated as:

$$C_k = \frac{E \times \sqrt{P_{UAV,k,s_k}}}{\sqrt{\left(2^{R_k^{\text{req}}/B_c} - 1\right) N_0 B_c}}, \tag{5.21}$$

with the expression of E given by Eq. (5.17).

For the problem to be feasible, a set of access power values $\boldsymbol{P_{UAV}}$, that guarantees the intersection of the different user coverage regions and satisfies the UAV power budget, i.e., constraint (5.7c), must be found. To this end, the following optimization problem is formulated:

$$\min_{\boldsymbol{P_{UAV}}} \sum_{k \in \mathcal{K}} P_{UAV,k,s_k} \tag{5.22}$$

$$\text{such that } \bigcap_{k \in \mathcal{K}} D_k \neq \emptyset. \tag{5.22a}$$

Constraint (5.22a) ensures that the coverage disk of each user intersects with the coverage disks of every other user. In other words, constraint (5.22a) ensures the existence of a space region where, if positioned, the UAV can simultaneously guarantee the rate requirements of all users. Note that the MBS coverage region is not taken into account in problem (5.22) since the feasibility test is performed while considering that no subband is used in the backhaul link, i.e., $b_s = 0$ for every subband $s \in \mathcal{S}$.

To account for constraint (5.22a) in the solution of problem (5.22), Helly's theorem [147] can be used. Helly's theorem states that, given M convex sets of dimension n, if the intersection of every $(n+1)$ combination of the M sets is non empty, so is the intersection of all M sets. Although the solution using Helly's theorem is precise, the execution time is in the order of $\mathcal{O}(M^3)$. Therefore, a less computationally demanding solution is presented next.

We note that the joint intersection of all coverage disks is non empty if we can find at least one point that belongs to all disks. Therefore, constraint (5.22) is reformulated as follows:

$$\min_{\boldsymbol{P_{UAV}}, \boldsymbol{o}} \sum_{k \in \mathcal{K}} P_{UAV,k,s_k} \tag{5.23}$$

$$\text{such that } d_{k,\boldsymbol{o}} \leq C_k, \ \forall k \in \mathcal{K}, \tag{5.23a}$$

where $\boldsymbol{o}$ is the intersection point in 2D space of all coverage disks. The coverage radius of user k, C_k, is expressed in terms of P_{UAV,k,s_k} using Eq. (5.21). Constraint (5.23a) ensures that point $\boldsymbol{o}$ belongs to all K disks by enforcing the distance between each user k and point $\boldsymbol{o}$, $d_{k,\boldsymbol{o}}$, to be less than the radius of the coverage disk of user k. Hence, constraint (5.23a) guarantees the intersection of all K coverage disks. Note that constraint (5.23a) can be reformulated as:

$$d_{k,\boldsymbol{o}}^2 \leq C_k^2, \ \forall \ k \in \mathcal{K}. \tag{5.24}$$

A solution for problem (5.23) can always be found, since the UAV power constraint is not enforced at this stage. Moreover, (5.23) is a convex optimization problem since it involves a linear objective function and convex inequality constraints. The Lagrangian of problem (5.23), with constraint (5.23a) reformulated as in (5.24), is given by:

$$\begin{aligned}\mathcal{L} &= \sum_{k\in\mathcal{K}} P_{UAV,k,s_k} - \tau_k \left(d_{k,\boldsymbol{o}}^2 - C_k^2\right) \\ &= \sum_{k\in\mathcal{K}} P_{UAV,k,s_k} - \tau_k \left((x_k - x_{\boldsymbol{o}})^2 + (y_k - y_{\boldsymbol{o}})^2 - \frac{E^2 P_{UAV,k,s_k}}{\left(2^{\frac{R_k^{\text{req}}}{B_c}} - 1\right) N_0 B_c} \right).\end{aligned} \tag{5.25}$$

The solution of problem (5.23) must satisfy the following Karush-Kuhn-Tucker (KKT) conditions:

$$\frac{\partial \mathcal{L}}{\partial P_{UAV,k,s_k}} = 0, \tag{5.26}$$

$$\frac{\partial \mathcal{L}}{\partial x_{\boldsymbol{o}}} = 0, \tag{5.27}$$

$$\frac{\partial \mathcal{L}}{\partial y_{\boldsymbol{o}}} = 0, \tag{5.28}$$

$$\tau_k \left((x_k - x_{\boldsymbol{o}})^2 + (y_k - y_{\boldsymbol{o}})^2 - \frac{E^2 P_{UAV,k,s_k}}{\left(2^{\frac{R_k^{\text{req}}}{B_c}} - 1\right) N_0 B_c} \right) = 0. \tag{5.29}$$

Solving the above equations yields the closed-form expressions of the power values in the access link guaranteeing the intersection of all coverage disks, $\boldsymbol{P}_{\boldsymbol{UAV}}^*$, and the 2D coordinates of the intersection point $\boldsymbol{o}$, $(x_{\boldsymbol{o}}, y_{\boldsymbol{o}})$:

$$x_{\boldsymbol{o}} = \frac{\sum_k \tau_k x_k}{\sum_k \tau_k}, \tag{5.30}$$

$$y_{\boldsymbol{o}} = \frac{\sum_k \tau_k y_k}{\sum_k \tau_k}, \tag{5.31}$$

$$P_{UAV,k,s_k}^* = \tau_k \left[(x_k - x_{\boldsymbol{o}})^2 + (y_k - y_{\boldsymbol{o}})^2 \right], \tag{5.32}$$

where

$$\tau_k = \frac{\left(2^{R_k^{\text{req}}/B_c} - 1\right) N_0 B_c}{E^2}. \tag{5.33}$$

The values of $\boldsymbol{P}_{\boldsymbol{UAV}}^*$ satisfy the UAV power budget if:

$$\sum_{k\in\mathcal{K}} P_{UAV,k,s_k}^* \leq P_{UAV}^{\max}. \tag{5.34}$$

If the condition in Eq. (5.34) is satisfied, problem (5.7) is feasible. In the opposite case, problem (5.7) is still solved according to the technique detailed in Section 5.4.2. However, constraint (5.7c) is relaxed. Additional steps are then performed to re-enforce the UAV power budget constraint while enhancing system performance. These additional steps are detailed in Section 5.4.3.

5.4.1.3 Initialization of the Access Power Values

As a result of solving the feasibility problem in (5.23), when the optimization problem in (5.7) is feasible, the power values $\boldsymbol{P}^{*}_{\boldsymbol{UAV}}$ result in an intersection region of all coverage disks that is just a point in the 2D space. Indeed, in order to achieve the objective of problem (5.23), the smallest power values satisfying constraint (5.23a) are found. Therefore, for at least two coverage disks, the intersection is just a point.

To maximize the size of the intersection region and be able to account for the BI, the initial power values must be increased as much as possible so that the smallest pairwise intersection is maximized. Indeed, it is clear from Eq. (5.9) that an increase of the power value $P_{UAV,k,s}$ results in an increase of the value of L_k^{th} and therefore in the value of C_k. Put differently, for the two disks with the smallest intersection, their radii, hence the associated power values, are recomputed to increase the intersection. Therefore, the following optimization problem is formulated and solved:

$$\max_{\boldsymbol{P}^{i}_{UAV}} \ \min_{k,k' \in \mathcal{K}} \{(C_k + C_{k'}) - d_{k,k'}\} \tag{5.35}$$

$$\text{such that } P^{i}_{UAV,k,s_k} \geq P^{*}_{UAV,k,s_k}, \ \forall\, k \in \mathcal{K}, \tag{5.35a}$$

$$\sum_{k \in \mathcal{K}} P^{i}_{UAV,k,s_k} = P^{\max}_{UAV}. \tag{5.35b}$$

In problem (5.35), $d_{k,k'}$ is the 2D distance between users k and k' (cf. Fig. 5.1). Constraint (5.35a) ensures that the updated power variables are at least equal to the ones found by the solution of problem (5.23), therefore preserving the intersection point, the coordinates of which are found in Eq. (5.30) and Eq. (5.31).

Problem (5.35) is convex, since it involves the minimization of the maximum of multiple convex functions. Hence, it can be solved efficiently to find the initial power values $\boldsymbol{P}^{i}_{UAV}$. When the feasibility test (5.23) is negative, the minimum required power in the access link to ensure the intersection of all coverage disks, found by Eq. (5.32), already exceeds the UAV power budget. In this case, the initial values of the power in the access link, $\boldsymbol{P}^{i}_{UAV}$, are set to the same values of $\boldsymbol{P}^{*}_{UAV}$.

5.4.1.4 MBS Coverage Radius

In this section, the maximum MBS coverage radius for each potential number of subbands S_{BH} is found.

To avoid the backhaul subband assignment step at this stage of the solution, for each potential number of subbands S_{BH}, equal power and rate repartition in the backhaul link are assumed. Note that this assumption is done in order to initialize the MBS coverage radius for each value of S_{BH}, and will be relaxed at later stages of the solution. With this assumption, the following observation is made.

Proposition 3. *If the values of $P_{MBS,s}$ and $R_{UAV,s}$ are constant for every subband $s \in \mathcal{S}$, the path-loss threshold $L^{th}_{S_{BH}}$ given by Eq. (5.18) decreases (i.e., becomes more strict) when the power in the access link increases.*

Following Proposition 3, the minimum (strictest) value of the required path-loss $L^{\text{th}}_{S_{BH}}$ is achieved when the access power value is the largest. To account for the worst case scenario in the computation of $L^{\text{th}}_{S_{BH}}$, the highest value of $P^{i}_{UAV,k,s}$, found from the solution of problem (5.35), is used. The minimum value of $L^{\text{th}}_{S_{BH}}$ is hence given by:

$$L^{\text{th}}_{S_{BH}} = -10\log_{10}\left\{\left(2^{\frac{R^{\text{req}}}{S_{BH}\times B_c}}-1\right)\times \frac{N_0 B_c + C_{SI}\times \max_s(\sum_k P^{i}_{UAV,k,s})}{P^{\max}_{MBS}/S_{BH}}\right\}, \tag{5.36}$$

where $R^{\text{req}} = \sum_k R^{\text{req}}_k$.

By considering this worst case situation, the strictest required path-loss from (5.18) is accounted for. The MBS coverage radius for S_{BH} backhaul subbands, $C_{MBS,S_{BH}}$, is then found using Eq. (5.16) after replacing L^{th}_k with $L^{\text{th}}_{S_{BH}}$. With $L^{\text{th}}_{S_{BH}}$ being the minimum value of the required path-loss, the computed radius $C_{MBS,S_{BH}}$ is also the smallest achievable value of the MBS coverage radius. Hence, $C_{MBS,S_{BH}}$ is guaranteed under any assignment of backhaul subbands.

Proposition 4. *As shown in Appendix A.1, $C_{MBS,S_{BH}}$ is an increasing function of S_{BH}.*

With the increase of S_{BH} and $L^{\text{th}}_{S_{BH}}$, the value of $C_{MBS,S_{BH}}$ increases to the point where the MBS coverage region encompasses the coverage regions of all users, when using the whole MBS power budget. However, an optimization problem that aims at minimizing the UAV transmit power is expected to position the UAV at a position that is close to the MBS and users at once. In other words, the UAV location should not be very far from the MBS since it would logically require more backhaul power to reach its backhaul rate requirement. This translates into more BI at the user side, which in its turn increases the needed UAV power to guarantee the access rate requirements. Hence, the following observation is made:

Proposition 5. *The maximum needed backhaul coverage radius, $C_{MBS,max}$ (cf. Fig. 5.1), is given by the distance between the MBS and the user farthest from it.*

Consequently, when the value of $C_{MBS,S_{BH}}$ is larger than the maximum needed value $C_{MBS,\max}$, it is replaced by $C_{MBS,\max}$.

When finding the MBS coverage radius in this way, the needed MBS power is ensured to respect the MBS power budget. In fact, for small values of S_{BH} resulting in $C_{MBS,S_{BH}} < C_{MBS,\max}$, the needed power is equal to the whole MBS power budget divided equally between the S_{BH} subbands. However, when S_{BH} increases and results in substituting $C_{MBS,S_{BH}}$ with $C_{MBS,\max}$, the needed backhaul power is reduced with respect to the budget $P^{\max}_{MBS}$. Therefore, the needed MBS power, P^{needed}_{MBS}, always satisfies: $P^{\text{needed}}_{MBS} \leq P^{\max}_{MBS}$.

5.4.2 Proposed Solution

5.4.2.1 Finding the minimum number of required backhaul subbands

To ensure that the UAV can provide the access rate to its users while meeting its backhaul rate requirement, it must be positioned in the intersection of the coverage region of all

users and the coverage region of the MBS. The latter depends on S_{BH}. When S_{BH} is small, the required data rate per subband, $R_{UAV,s}$, is naturally large. This causes the maximum tolerable path-loss $L^{\text{th}}_{S_{BH}}$, given by Eq. (5.18), to be small, resulting in a narrow MBS coverage radius. On the other hand, if all subbands are used in the backhaul link, the MBS coverage radius becomes large. However, for some users with high rate requirements, the added BI may prove too much and shrink their coverage regions. This could result in an impossibility to serve all users simultaneously. Therefore, both the number and the choice of backhaul subbands should be optimized.

To find the number of backhaul subbands that results in the minimum needed UAV power, for each potential value of S_{BH}, i.e., for $S_{BH} = \{1, \ldots, S\}$, the corresponding backhaul subbands must be chosen, and the needed UAV power must be found. However, to reduce the number of backhaul subbands values to be tested, the following observation is made.

Proposition 6. *The largest simultaneous coverage region for the users, D_{int}^{max}, is achieved when none of the subbands is used in the backhaul link, i.e., when $b_s = 0$ for every subband $s \in \mathcal{S}$.*

Proof. Let D_{int} be the simultaneous coverage region when a subband s is used in the backhaul link and let k be the user scheduled on s. According to Eq. (5.9), the value of L^{th}_k decreases when the BI increases. Hence, the coverage region of k becomes smaller when its allocated subband is used in the backhaul link, in comparison to the opposite case. Therefore, the simultaneous coverage region D_{int} is smaller than $D^{\max}_{\text{int}}$, when one or more subbands are used in the backhaul link. □

Following Proposition 6, any MBS coverage region that does not intersect with $D^{\max}_{\text{int}}$ does not intersect with the simultaneous user coverage region that accounts for BI. Let $(x_0, y_0) \in D^{\max}_{\text{int}}$ be the closest point in 2D space to the MBS, i.e.,:

$$(x_0, y_0) = \underset{(x,y) \in D^{\max}_{\text{int}}}{\operatorname{argmin}} \sqrt{x^2 + y^2}. \tag{5.37}$$

In order to have an intersection with $D^{\max}_{\text{int}}$, the MBS coverage radius must satisfy:

$$C_{MBS,S_{BH}} \geq C_{MBS,\min} = \sqrt{{x_0}^2 + {y_0}^2}. \tag{5.38}$$

The minimum value of the number of backhaul subbands, $S_{BH,\min}$, is the smallest value of S_{BH} to satisfy Eq. (5.38). The value of $S_{BH,\min}$ is found using bisection search.

Having found $S_{BH,\min}$ and the total backhaul power needed for every potential value of $S_{BH} \in \mathcal{S}^{\text{poss}}_{BH} = \{S_{BH,\min}, \ldots, S\}$ (in Section 5.4.1.4), we now turn our attention to the assignment of backhaul subbands. Clearly, this assignment should be done carefully since it has an undeniable impact on the interference levels experienced by the users.

5.4.2.2 Deciding on the number and choice of backhaul subbands

When finding the required MBS path-loss and maximum MBS coverage radius in Section 5.4.1.4, the maximum power value in the access link was taken without considering the experienced BI. Hence, the value of $S_{BH,\min}$ was found while considering the maximum of SI but in the absence of BI. When accounting for the BI, the coverage regions of users decrease in size. Therefore, it is not guaranteed that all values in the set $\mathcal{S}^{\text{poss}}_{BH}$ ensure

that the coverage regions of users intersect with the one of the MBS on one hand, and with each other on the other, while respecting the UAV power budget. In this section, we introduce an algorithm to find the value of S_{BH} that results in the best performance.

Starting from the case where all subbands are used in the backhaul link, i.e., for $S_{BH} = S$, problem (5.39) is solved to find the power per subband in the backhaul link needed to achieve the backhaul capacity:

$$\min_{\boldsymbol{P_{MBS}}} \sum_{s \in \mathcal{S}_{S_{BH}}} P_{MBS,s}^{S_{BH}} \tag{5.39}$$

$$\text{such that} \sum_{s \in \mathcal{S}_{S_{BH}}} R_{UAV,s} = R^{\text{req}}. \tag{5.39a}$$

In constraint (5.39a), $\mathcal{S}_{S_{BH}}$ is the set of subbands used in the backhaul link when the total number of backhaul subbands is S_{BH} and $R_{UAV,s}$ is given by Eq. (5.5) while considering the initial access power values found in Section 5.4.1.3. At this point, the UAV is assumed to be positioned at the maximum tolerable distance from the MBS. In other words, the distance between the UAV and the MBS is equal to the MBS coverage radius $C_{MBS,S_{BH}}$.

Finding the Lagrangian of problem (5.39) and setting it to zero yields:

$$P_{MBS,s}^{S_{BH}} = \frac{\lambda}{\log(2)} - \frac{N_0 B_c + C_{SI} \sum_k P_{UAV,k,s}^{i}}{10^{-L_{S_{BH}}^{\text{th}}/10}}, \tag{5.40}$$

where λ is the Lagrange multiplier given by:

$$\lambda = 2^{\frac{1}{S_{BH}} \left(\frac{R^{\text{req}}}{B_c} - \sum_{s \in \mathcal{S}_{S_{BH}}} \log_2 \left(\frac{10^{-L_{S_{BH}}^{\text{th}}/10}}{\log(2) \left(N_0 B_c + C_{SI} \sum_k P_{UAV,k,s}^{i} \right)} \right) \right)}. \tag{5.41}$$

The MBS power values resulting from Eq. (5.40) are found while using the initial access power values $\boldsymbol{P_{UAV}^{i}}$ which do not account for the BI. Therefore, at this stage, the access power values accounting for the BI are re-calculated in a similar manner to that in Section 5.4.1.2: The maximum coverage radius for user k, C_k, is given by (5.16). However, the path-loss thresholds must be modified to account for the BI. For this purpose, in this section, Eq. (5.21) is substituted with the following expression:

$$C_{k,S_{BH}} = \frac{E \times \sqrt{P_{UAV,k,s_k}}}{\sqrt{\left(2^{R_k^{\text{req}}/B_c} - 1\right) \left(N_0 B_c + G_{k,s_k}^2 P_{MBS,s_k}^{S_{BH}}\right)}}. \tag{5.42}$$

To find the access power needed for S_{BH} backhaul subbands, the following convex optimization problem is solved:

$$\min_{\boldsymbol{P_{UAV}^{S_{BH}}}, \boldsymbol{o}} \sum_{k \in \mathcal{K}} P_{UAV,k,s_k}^{S_{BH}} \tag{5.43}$$

$$\text{such that } d_{k,\boldsymbol{o}} \leq C_{k,S_{BH}}, \ \forall\, k \in \mathcal{K}, \tag{5.43a}$$

$$d_{MBS,\boldsymbol{o}} \leq C_{MBS,S_{BH}}. \tag{5.43b}$$

Problem (5.43) is a generalization of problem (5.23) that accounts for the BI term. Constraint (5.43b) ensures the joint intersection of the MBS coverage region and the coverage region of each user.

Being mutually dependent, to find the power values used in the backhaul and the access links for S_{BH} used subbands, problems (5.39) and (5.43) are repeatedly solved until convergence. The latter is reached when the change in the values of $\boldsymbol{P}_{\boldsymbol{MBS}}^{\boldsymbol{S_{BH}}}$ and $\boldsymbol{P}_{\boldsymbol{UAV}}^{\boldsymbol{S_{BH}}}$, between two successive iterations, becomes negligible.

Having found the final value for the power $\boldsymbol{P}_{\boldsymbol{UAV}}^{\boldsymbol{S_{BH}}}$ for S_{BH} subbands, in order to test the smaller value $(S_{BH}-1)$, the subband s_h requiring the highest access power value is removed from the backhaul subband pool before proceeding with the same described method.

These steps are summarized in Algorithm 9. Upon testing all possible subband values, the value of S_{BH} requiring the minimum access power, S_{BH}^f, is retained. In case the total access power is smaller than $P_{UAV}^{\max}$, to maximize the intersection region, problem (5.35) is solved for the value of S_{BH}^f, as stated in step 8. Algorithm 9 also yields the backhaul subband assignment $\boldsymbol{b}$ and the intersection region $X_{S_{BH}^f}$. In the opposite case, i.e., when the total needed access power is larger than $P_{UAV}^{\max}$, the intersection region is chosen as the coverage disk of the MBS to guarantee the backhaul constraint.

The complexity of Algorithm 9 stems from solving the convex problems (5.35), (5.39) and (5.43). For problem (5.39), a closed form solution with a complexity in the order of $\mathcal{O}(S)$ is found. Problems (5.35) and (5.43) are solved using the interior point method, having $\mathcal{O}(S^{2.5})$ as complexity [148]. Moreover, problem (5.35) is solved for the retained value of S_{BH} requiring the minimum needed power, i.e., for S_{BH}^f. However, problems (5.39) and (5.43) are solved for all values of $S_{BH} \in \mathcal{S}_{BH}^{\text{poss}}$. Since $|\mathcal{S}_{BH}^{\text{poss}}| \leq S$, the complexity of Algorithm 9 is hence upper bounded by $\mathcal{O}(S^{3.5})$.

5.4.2.3 Finding the optimal UAV position and the final power levels

Having decided on the assignment of backhaul subbands and found the acceptable horizontal region in which the UAV can be located to serve users, in this section, the exact UAV position as well as the power values in the backhaul and access links are optimized.

To minimize the UAV transmit power, problem (5.7) is reformulated as follows:

$$\min_{\substack{\boldsymbol{P_{UAV}},\boldsymbol{P_{MBS}},\\ \boldsymbol{z_{UAV}}}} \sum_{k\in\mathcal{K}}\sum_{s\in\mathcal{S}} a_{k,s}\, P_{UAV,k,s} \tag{5.44}$$

$$\begin{aligned}
\text{such that} \quad & \sum_{s\in\mathcal{S}} a_{k,s}\, R_{k,s} = R_k^{\text{req}},\ \forall\, k\in\mathcal{K}, && \text{(5.44a)}\\
& \sum_{k\in\mathcal{K}}\sum_{s\in\mathcal{S}} a_{k,s}\, R_{k,s} \leq \sum_{s\in\mathcal{S}} b_s\, R_{UAV,s}, && \text{(5.44b)}\\
& \sum_{s\in\mathcal{S}} b_s\, P_{MBS,s} \leq P_{MBS}^{\max}, && \text{(5.44c)}\\
& H_{\min} \leq H \leq H_{\max}, && \text{(5.44d)}\\
& (x_{UAV}, y_{UAV}) \in X_{S_{BH}^f}. && \text{(5.44e)}
\end{aligned}$$

Constraint (5.44e) states that the 2D or horizontal position of the UAV should be in the simultaneous coverage region of the users and the MBS.

Algorithm 9 Deciding on the number and choice of backhaul subbands

Output: $S_{BH}^{f}, \boldsymbol{b}, X_{S_{BH}^{f}}$.
Initialization: $S_{BH} = S,\ \mathcal{S}_{S_{BH}} = \{1, \ldots, S\}$.
1: **for** $S_{BH} = S \rightarrow S_{BH,\min}$ **do**
Repeat:
2: Solve problem (5.39) to find the backhaul power $\boldsymbol{P}_{\boldsymbol{MBS}}^{\boldsymbol{S_{BH}}}$.
3: Solve problem (5.43) to find the access power $\boldsymbol{P}_{\boldsymbol{UAV}}^{\boldsymbol{S_{BH}}}$.
Until Convergence
4: $s_h = \underset{s \in \mathcal{S}_{S_{BH}}}{\operatorname{argmax}} P_{UAV,k,s}^{S_{BH}}$,
$\mathcal{S}_{S_{BH}-1} = \mathcal{S}_{S_{BH}} \setminus \{s_h\}$.
5: **end for**
6: $S_{BH}^{f} = \underset{S_{BH} \in \mathcal{S}_{BH}^{\text{poss}}}{\operatorname{argmin}} \sum_{s \in \mathcal{S}} \sum_{k \in \mathcal{K}} P_{UAV,k,s}^{S_{BH}}$.
$b(s) = 1, \forall s \in \mathcal{S}_{S_{BH}^{f}}$.
7: **if** $\sum P_{UAV,k,s}^{S_{BH}^{f}} \leq P_{UAV}^{\max}$ **then**
8: Solve problem (5.35) to find the power values in the access link, $\boldsymbol{P}_{\boldsymbol{UAV}}^{\boldsymbol{S_{BH}^{f}}}$, that maximize the intersection region.
$X_{S_{BH}^{f}} \leftarrow$ intersection region for S_{BH}^{f}.
9: **else**
10: $X_{S_{BH}^{f}} = D_{MBS,S_{BH}^{f}}$.
11: **end if**

To minimize the needed access power, the BI should be minimized while meeting the backhaul rate requirement. To this end, constraint (5.44b) should be met with equality.

From Eq. (5.3), the value of P_{UAV,k,s_k} satisfying constraint (5.44a) for each user k is given by:

$$P_{UAV,k,s_k} = A_1(k) \frac{N_0 B_c + a_{s_k} P_{MBS,s_k} G_{k,s_k}^2}{G_{UAV,k}^2}, \tag{5.45}$$

where $A_1(k) = \left(2^{R_k^{\text{req}}/B_c} - 1\right)$. Moreover, P_{MBS,s_k} is the backhaul power used by the MBS on subband s_k to meet the backhaul rate R_{UAV,s_k}. The expression of P_{MBS,s_k} is given by:

$$P_{MBS,s_k} = \left(2^{\frac{R_{UAV,s_k}}{B_c}} - 1\right) \frac{N_0 B_c + C_{SI} P_{UAV,k,s_k}}{G_{UAV,MBS}^2}. \tag{5.46}$$

Note that in Eq. (5.45) and (5.46), the actual value of the elevation angle is used for the calculation of the channel gains G, i.e., the elevation angle of user k and that of the MBS are respectively given by:

$$\theta_k = \arctan\left(H/r_k\right), \tag{5.47}$$

$$\theta_{MBS} = \arctan\left(H/r_{MBS}\right). \tag{5.48}$$

After replacing P_{MBS,s_k} in Eq. (5.45) by its expression from Eq. (5.46) , P_{UAV,k,s_k} can be expressed as:

$$P_{UAV,k,s_k} = \frac{A_1(k) N_0 B_c \times \{G_{UAV,MBS}^2 + a_{s_k} A_2(s_k) G_{k,s_k}^2\}}{G_{UAV,MBS}^2 G_{UAV,k}^2 - a_{s_k} G_{k,s_k}^2 A_1(k) A_2(s_k) C_{SI}}, \tag{5.49}$$

where $A_2(s_k) = \left(2^{R_{UAV,s_k}/B_c} - 1\right)$.

Although the value of $\boldsymbol{A_1}$ is known for all users, the value of $\boldsymbol{A_2}$ is not since it depends on the backhaul rate per subband which, in its turn, depends on the power in the backhaul and access links. Having expressed the values of the power in the access and the backhaul links in terms of $\boldsymbol{A_1}$, $\boldsymbol{A_2}$, and the channel gain values, solving problem (5.44) can be achieved by solving the following, equivalent, problem:

$$\min_{\boldsymbol{z_{UAV}}, \boldsymbol{A_2}} \sum_{k \in \mathcal{K}} P_{UAV,k,s_k} \tag{5.50}$$

$$\text{such that } \sum_{k \in \mathcal{K}} \sum_{s \in \mathcal{S}} a_{k,s} R_{k,s} = \sum_{s \in \mathcal{S}} b_s R_{UAV,s}, \tag{5.50a}$$

$$\sum_{s \in \mathcal{S}} b_s P_{MBS,s} \leq P_{MBS}^{\max}, \tag{5.50b}$$

$$H_{min} \leq H \leq H_{max}, \tag{5.50c}$$

$$(x_{UAV}, y_{UAV}) \in X_{S_{BH}^f}, \tag{5.50d}$$

where P_{UAV,k,s_k} is given by Eq. (5.49).

The solution of problem (5.50) is obtained iteratively. At the first iteration, $\boldsymbol{A_2}$ is initialized with an equal rate repartition in the backhaul link, leading to $A_2(s_k)$ taking the following value:

$$A_2(s_k) = \left(2^{\frac{R^{\text{req}}}{S_{BH}^f B_c}} - 1\right), \ \forall\, s_k \in \mathcal{S} \text{ such that } b(s_k) = 1. \tag{5.51}$$

With $\boldsymbol{A_2}$ known, problem (5.50) is solved numerically for the UAV position. Then, with the computed position, the values of $\boldsymbol{A_2}$ minimizing the UAV transmit power are found. The two steps are iterated until convergence, reached when the change in the values of $\boldsymbol{z_{UAV}}$ and $\boldsymbol{A_2}$, between two consecutive iterations, becomes negligible. Upon convergence, the power values in the backhaul and access links are computed according to Eq. (5.46) and Eq. (5.49), respectively. The complexity of solving (5.50) is upper bounded by $\mathcal{O}(S^{2.5})$ [148].

It should be noted that, in practice, once the UAV has been positioned, it can exchange the channel state information (CSI) with its users and with the MBS. Hence, in reality, the final power values can be calculated based on the actual channel state values instead of the mean channel state obtained by the probabilistic path loss model (i.e., Eq. (5.2)). It should also be noted that finding the UAV position by solving problem (5.50) and enforcing the horizontal position of the UAV to be in the intersection region of the coverage disks of all users as in constraint (5.50d) does not guarantee that the power needed by the UAV is lower than its power budget with probability one. In fact, it may not be possible to find a value of the UAV altitude H simultaneously guaranteeing the rate requirements of all users. In that case, the actual achieved elevation angles differ significantly from the optimal value θ_{opt}, leading to a higher needed power. If that should happen, the technique introduced next, in Section 5.4.3, is employed to maximize the achieved rates while respecting the UAV power budget.

5.4.3 NOMA Pairing to Account for the Case where the Rate Requirements Cannot Be Met with $P_{UAV}^{\max}$

If the user rate requirements are very high, the UAV power budget might not be sufficient to ensure them. In this case, either the feasibility test in (5.23) fails to find power values respecting the UAV budget, or Algorithm 9 fails to find a value of S_{BH} respecting this budget. Moreover, in a small percentage of the cases, a value of the UAV altitude simultaneously guaranteeing the user rate requirements might not be achievable. If that should happen, the UAV position and the power values are still found according to the analysis of Section 5.4.2.3. However, the solution of problem (5.50) violates constraint (5.7c). Therefore, additional steps, described hereinafter, are performed to enhance user satisfaction. More precisely, after finding the maximum achievable user rates with the UAV power budget, a NOMA pairing step is conducted to increase, to the extent of the possible, the rates of users not yet reaching their target value.

5.4.3.1 Finding the Maximum Achievable Rates with $P_{UAV}^{\max}$

To find the maximum achievable rates respecting the UAV power budget, problem (5.52) is formulated:

$$\max_{\boldsymbol{P_{UAV}}} \sum_{k\in\mathcal{K}} R_{k,s_k} \tag{5.52}$$

$$\text{such that } \sum_{k\in\mathcal{K}} P_{UAV,k,s_k} \leq P_{UAV}^{\max}, \tag{5.52a}$$

$$\sum_{k\in\mathcal{K}} P_{MBS,s_k} \leq P_{MBS}^{\max}, \tag{5.52b}$$

$$R_{k,s_k} \leq R_k^{\text{req}},\ \forall\ k \in \mathcal{K}, \tag{5.52c}$$

where R_{k,s_k} is the rate of user k over its assigned subband s_k, given by Eq. (5.3). Constraint (5.52c) ensures that none of the users exceed their rate requirement. Variable P_{MBS,s_k} is expressed using Eq. (5.46), with the value of R_{UAV,s_k} found by the solution of problem (5.50). The optimization problem formulated in (5.52) is convex, hence can be solved efficiently by the interior point method.

5.4.3.2 Finding the Candidate Subbands for NOMA Pairing

The solution of problem (5.52) yields the data rates that users can obtain from orthogonal multiple access (OMA) scheduling. To bring users closer to their requested rates, a NOMA pairing step is conducted to use the allocated power of unsatisfied users more efficiently.

5.4.3.2.1 NOMA Rate Expressions and Constraints

With NOMA scheduling, the signals of up to N_n users are superposed and transmitted over subband n. This multiplexing results in co-channel interference between the collocated users. Hence, user j, scheduled over subband n, performs SIC to remove the interference of some of the other collocated users before demodulating its own signal. Assuming perfect SIC, the achievable rate in this case can be written as:

$$R_{j,n} = B_c \log_2 \left(1 + \frac{P_{UAV,j,n}\, G_{UAV,j}^2}{I_{j,n} + b_n\, P_{MBS,n}\, G_{j,n}^2 + N_0 B_c} \right), \tag{5.53}$$

where $I_{j,n} = \sum_{k \in \mathcal{I}_{j,n}} P_{UAV,k,n} G^2_{UAV,j}$ represents the co-channel interference experienced by user j from users in the set $\mathcal{I}_{j,n}$, i.e., from users whose interference could not be removed. Because of the complexity resulting from SIC decoding, the maximum value of N_n is restricted to 2, for every subband $n \in \mathcal{S}$. When $b_n = 1$, i.e., when subband n is used in the backhaul link, due to the presence of the BI (second term in the denominator of Eq. (5.53)), user j, called first user, can remove the interference of the other user k collocated on n, if their channel gains verify the following condition, proved in [149]:

$$G_{UAV,k} G_{j,n} < G_{UAV,j} G_{k,n} \tag{5.54}$$

User k, called second user, decodes its signal directly while considering the signal of the first user j as noise.

To guarantee SIC stability [49, 81], i.e., successful decoding at the level of user j, the signal of user k must be received at user j with an amount of power superior to that of user j, added to the BI power at the level of user j. Indeed, as shown in [30], in case this power condition is not satisfied, the users outage probabilities will always be one. When $b_n = 1$, the power multiplexing condition is written as [149]:

$$P_{UAV,k,s_k} > P_{UAV,j,n} + P_{MBS,n} \, G^2_{j,n} / G^2_{UAV,j}. \tag{5.55}$$

When subband n is not used in the backhaul link, leading to $b_n = 0$, conditions (5.54) and (5.55) become respectively (5.56) and (5.57) as in the classical NOMA case [81, 49].

$$G_{UAV,k} < G_{UAV,j}, \tag{5.56}$$

$$P_{UAV,k,s_k} > P_{UAV,j,n}. \tag{5.57}$$

5.4.3.3 Determining the Eligible Subbands for Each User

The set of users being considered for NOMA pairing, $\mathcal{K}_{NOMA}$, consists of users that have not achieved their requested data rates, i.e.,:

$$\mathcal{K}_{NOMA} = \{k \in \mathcal{K}, | R_k < R_k^{\text{req}}\}. \tag{5.58}$$

The set of subbands considered for NOMA pairing, $\mathcal{S}_{NOMA}$, consists of subbands belonging to users having achieved their required data rates, i.e.,:

$$\mathcal{S}_{NOMA} = \{s \in \mathcal{S} |\ a_{k,s} = 1, R_k = R_k^{\text{req}}\}. \tag{5.59}$$

To bring a user $k \in \mathcal{K}_{NOMA}$ closer to its requested data rate, the power allocated to k is divided between its originally allocated subband s_k and a subband n on which k can be paired using NOMA such that its throughput is increased. To avoid penalizing users having achieved their rate requirements and to avoid having a chain of power modifications on each subband n considered for NOMA pairing, we make the following assumptions:

1. unsatisfied users are paired as second users via NOMA, i.e., as users not performing SIC,

2. the access power of the user already scheduled on subband n (i.e., first user) is kept constant after pairing,

3. the backhaul power on subband n is kept constant.

Assumption 1 ensures that the satisfied first user, initially scheduled on subband n, performs SIC in order to decode its message without the interference from the newly paired NOMA user. Together with assumptions 2 and 3, assumption 1 guarantees that the rate of the initially scheduled user is not penalized.

Let k be an unsatisfied user initially scheduled as OMA user on subband s_k. Among the set of available subbands for NOMA pairing $\mathcal{S}_{NOMA}$, k can be scheduled via NOMA on a subset of subbands satisfying certain conditions. The subset of subbands on which user k can be scheduled using NOMA is denoted by $\mathcal{S}^k_{NOMA}$ and satisfies $\mathcal{S}^k_{NOMA} \subseteq \mathcal{S}_{NOMA}$. Let $n \in \mathcal{S}^k_{NOMA}$ and let j be the user already scheduled on subband n. Subband n should satisfy the following conditions:

1. condition (5.56) when $b_n = 0$, or condition (5.54) when $b_n = 1$,
2. condition (5.57) when $b_n = 0$, or condition (5.55) when $b_n = 1$,
3. scheduling user k on subband $n \in \mathcal{S}^k_{NOMA}$ can increase the achieved data rate of user k,
4. scheduling user k on subband $n \in \mathcal{S}^k_{NOMA}$ does not penalize the backhaul rate already achieved on subbands n and s_k.

Condition 1 ensures that user k is scheduled as second user on subband n and condition 2 ensures SIC stability [49, 81]. In fact, since the power previously allocated to user k, P_{UAV,k,s_k} is to be divided between multiple subbands, condition 2 ensures k has enough power to divide between s_k and n in a SIC stable manner.

Finding subbands ensuring conditions 1 and 2 is straightforward. To guarantee conditions 3 and 4, the power on subbands s_k and n necessary for user k to reach its previously achieved rate on subband s_k alone, R_{UAV,s_k}, while maintaining constant the backhaul rate on subbands s_k and n, is found. If the total needed power of user k on subbands s_k and n is lower than the previously allocated power to user k, the surplus of power can be used towards increasing the achieved rate of user k. To find whether subband n ensures conditions 3 and 4, the following problem is solved:

$$\min_{R'_{k,s_k}, R'_{k,n}} \quad P'_{UAV,k,s_k} + P'_{UAV,k,n} \tag{5.60}$$

$$\begin{aligned}
\text{such that } & R'_{k,s_k} + R'_{k,n} = R_{k,s_k}, && \text{(5.60a)}\\
& P'_{UAV,k,n} > P'_{UAV,j,n} \text{ if } b_n = 0 \text{ or} \\
& P'_{UAV,k,n} > P_{UAV,j,n} + P_{MBS,n}\, G^2_{j,n} \,/\, G^2_{UAV,j} \text{ if } b_n = 1, && \text{(5.60b)}\\
& P'_{MBS,s_k} \leq P_{MBS,s_k}, && \text{(5.60c)}\\
& R'_{UAV,s_k} + R'_{UAV,n} = R_{UAV,s_k} + R_{UAV,n}. && \text{(5.60d)}
\end{aligned}$$

In problem (5.60), P'_{UAV,k,s_k} and $P'_{UAV,k,n}$ are the power values relative to the signal of user k on subbands s_k and n respectively whereas R'_{k,s_k}($R'_{k,n}$ respectively) is the updated rate of user k on subband s_k (on subband n respectively). Constraint (5.60a) ensures that the achieved rate when partitioning the power allocated to user k between subbands s_k and n

is equal to the previously achieved rate of user k on subband s_k alone. Constraint (5.60b) guarantees SIC stability on subband n and constraint (5.60c) guarantees that P'_{MBS,s_k}, the updated backhaul power on s_k, does not exceed its initial value P_{MBS,s_k}. Constraint (5.60d) guarantees that the sum backhaul rate on subbands s_k and n remains constant, $R'_{UAV,s_k}$and $R'_{UAV,n}$ being the updated backhaul rates on subbands s_k and n, respectively.

Solving problem (5.60) yields the minimum power values necessary to ensure the achievable rate of user k (found by Problem (5.52)), when paired over subband n as second user. If this solution satisfies:

$$\Delta P = P_{UAV,k,s_k} - \left(P'_{UAV,k,s_k} + P'_{UAV,k,n}\right) \geq 0, \tag{5.61}$$

i.e., if a surplus of power remains, user k can use the excess power ΔP to increase its achieved rate. To solve problem (5.60), the cases where subband n is used in the backhaul link or not are discussed separately in Appendix A.2.

5.4.3.4 NOMA Pairing and Power Optimization

In this section, the assignment of unsatisfied users to the candidate subbands, as well as the power optimization subsequent to pairing are discussed.

To conduct the NOMA pairing step, a metric $M(k,n)$ is associated with each pair $\left(k \in \mathcal{K}_{NOMA}, n \in \mathcal{S}^k_{NOMA}\right)$, depending on whether subband n is used in the backhaul link or not. Metric $M(k,n)$ is formulated as:

$$M(k,n) = \begin{cases} \frac{P_{UAV,j,n} G^2_{UAV,k} + P_{MBS,n} G^2_{k,n}}{G^2_{UAV,k}}, & \text{if } b_n = 1, \\ -G^2_{UAV,k}, & \text{otherwise.} \end{cases} \tag{5.62}$$

When $b_n = 1$, $M(k,n)$ reflects the BI as well as the NOMA interference suffered by user k if paired on subband n (cf. Eq. (5.53)). If $b_n = 0$, the metric takes into account only the channel gain of user k. By minimizing $M(k,n)$, the achieved rate is maximized by minimizing the experienced interference when $b_n = 1$, and by maximizing the channel gain of the paired user when $b_n = 0$.

Once this metric is calculated, the Hungarian algorithm [146] is used to find the NOMA assignment of unsatisfied users to candidate subbands. Then, the following optimization problem is solved to find the new power values and deduce the achieved user rate for each paired user separately:

$$\max_{\substack{P'_{UAV,k,s_k}, \\ P'_{UAV,k,n}}} \quad R'_{k,s_k} + R'_{k,n} \tag{5.63}$$

$$\begin{aligned}
\text{such that } & P'_{UAV,k,s_k} + P'_{UAV,k,n} \leq P_{UAV,k,s_k}, && \text{(5.63a)} \\
& P'_{UAV,k,n} > P'_{UAV,j,n} \quad \text{if } b_n = 0 \text{ or} \\
& P'_{UAV,k,n} > P_{UAV,j,n} + P_{MBS,n}\, G^2_{j,n} \,/\, G^2_{UAV,j} \quad \text{if } b_n = 1, && \text{(5.63b)} \\
& P'_{MBS,s_k} \leq P_{MBS,s_k}, && \text{(5.63c)} \\
& R'_{UAV,s_k} + R'_{UAV,n} = R_{UAV,s_k} + R_{UAV,n}. && \text{(5.63d)} \\
& R'_{k,s_k} + R'_{k,n} \leq R^{\text{req}}_k. && \text{(5.63e)}
\end{aligned}$$

Problem (5.63) divides the total power allocated to user k, P_{UAV,k,s_k}, between subbands s_k and n, such that its achieved rate is maximized.

The steps described in this section are repeated until either all users are satisfied, or no more NOMA pairings can take place. Algorithm 10 summarizes the additional steps described in section 5.4.3.

Algorithm 10 NOMA pairing to maximize the achieved rates

Output: $\boldsymbol{P_{UAV}}, \boldsymbol{P_{MBS}}, \boldsymbol{R}, \boldsymbol{R_{UAV}}$.

1: Solve problem (5.52) to find the maximum achievable rates in an OMA setting with the current power budget.
2: For unsatisfied users, find the candidate subbands for NOMA pairing, by solving problem (5.60) for each subband $n \in \mathcal{S}_{NOMA}$.
Repeat:
3: Find the assignment of second users to candidate subbands using the Hungarian Algorithm.
4: Solve problem (5.63) for each paired user.
5: Update the achieved rates.
Until $\mathcal{K}_{NOMA} = \emptyset \mid \mathcal{S}^k_{NOMA} = \emptyset, \ \forall \ k \in \mathcal{K}_{NOMA}$

Remark 3. *In this chapter, we assume that the SIC decoding process is perfect. In practice, the SIC process might be imperfect due to hardware impairments. If one were to consider an imperfect SIC, the effect of the latter would be only perceived at the level of the user having a higher channel gain on subband n considered for NOMA pairing, as the number of users non-orthogonally multiplexed over each channel is limited to two. The user with a higher channel gain on subband n is referred to as first user and denoted by j. Accounting for the interference brought by the imperfect SIC, the rate achieved by user j on subband n can be formulated as:*

$$R_{j,n} = B_c \log_2 \left(1 + \frac{P'_{UAV,j,n} G^2_{UAV,j}}{I^{SIC}_{j,n} + b_n P_{MBS,n} G^2_{j,n} + N_0 B_c} \right). \tag{5.64}$$

In Eq. (5.64), $I^{SIC}_{j,n}$ is the interference experienced by user j due to the imperfect SIC cancellation given by [150]:

$$I^{SIC}_{j,n} = \epsilon \, P'_{UAV,k,n} G^2_{UAV,j}, \tag{5.65}$$

where ϵ denotes the uncanceled fraction of the interfering signal power and $P'_{UAV,k,n}$ is the power allocated to the second user, user k, on subband n.

In the proposed solution, when performing NOMA pairing while considering an imperfect SIC, we need to ensure that the rate of the initially scheduled user j is not penalized by its collocation with user k on subband n. Hence, to oppose the effect of the imperfect SIC, the power allocated to user j must be sufficiently increased. However, the system is power-limited, and the only power available for re-distribution is the one allocated to user k. Hence, in addition to partitioning its power between the new subband n and the old one s_k, user k must lend a part of its power to user j to counterbalance the effect of the added interference. Therefore, when checking whether the available power of user k is sufficient to increase its data rate in problem (5.60), we should also check whether a surplus of the power of user k, sufficient to make up for the added interference at the level of user j, remains. If the answer is positive, when solving problem (5.63) to find the new power variables, the power needed to be added to the signal of user j must also be found.

5.4.4 Summary of the Proposed Solution

The flowchart in Fig. 5.4 summarizes the overall proposed method.

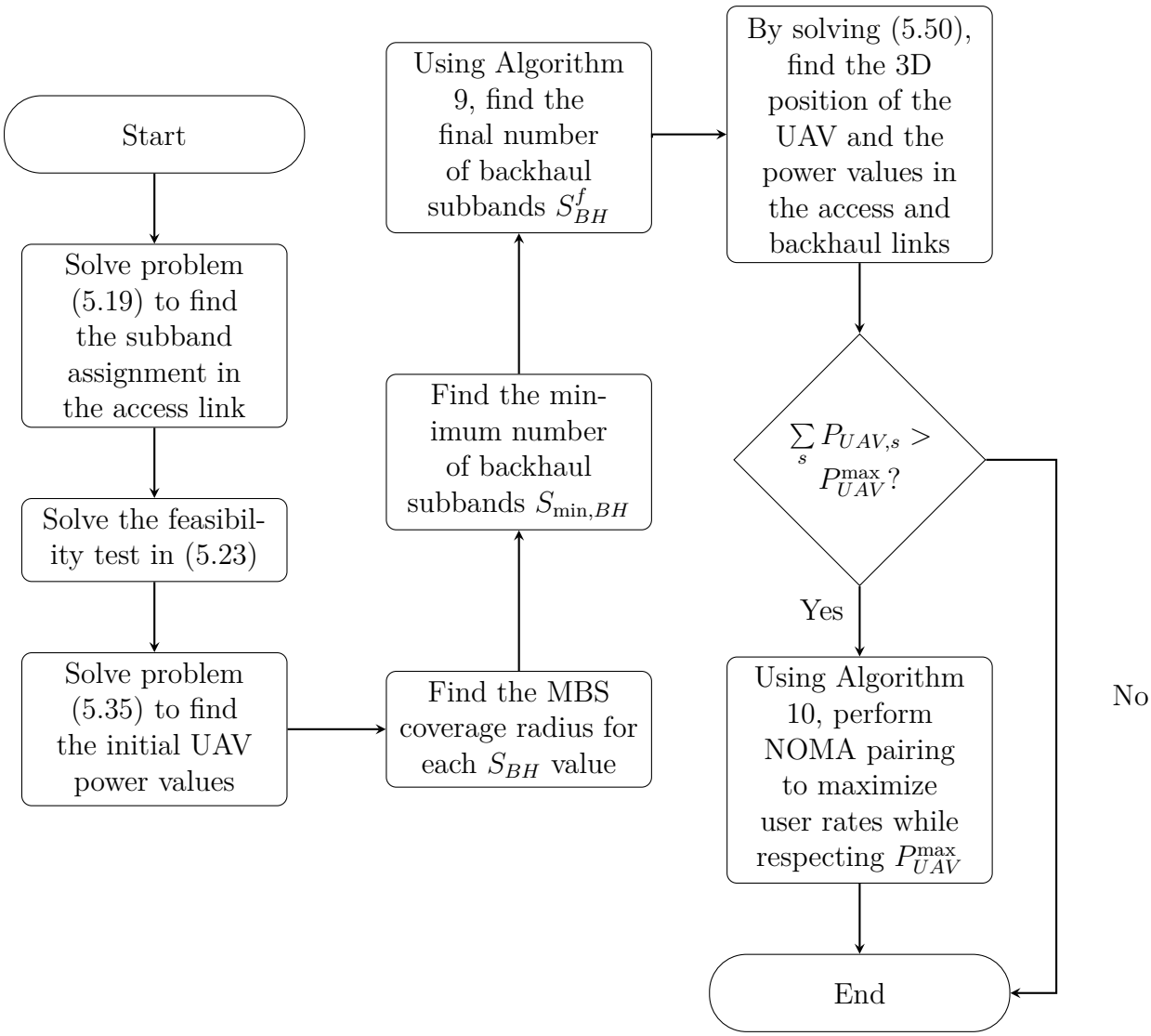


Figure 5.4 – Flow chart of the proposed method.

5.5 Simplified Solution to the Positioning and Resource Allocation Problems in UAV-Enabled Networks

In this section, a simplified version of the solution proposed in Section 5.4 is presented.

5.5.1 Preliminaries

The simplified solution to the UAV positioning and the resource allocation problems is based on three main assumptions:

1. First, the subband assignment in the access link between the UAV and the end users in not considered. It is supposed to be known in advance and taken arbitrarily.

2. Second, an equal initial power allocation is assumed in both the access and the backhaul links to simplify the solution.

3. Third and most importantly, the simplified solution assumes that the available UAV power budget is sufficient to guarantee user rate requirements. Hence, this simplified solution is not applicable to systems where user rate requirements are relatively high.

Building on these assumptions, similarly to Section 5.4, each user k is associated with a 2D coverage region guaranteeing its rate requirement defined based on constraint (5.7a). For each user $k \in \mathcal{K}$, the 2D coverage region of user k is found according to Section 5.3 yielding a coverage disk D_k centered at user k with a coverage radius C_k. For each value of the number of subbands in the backhaul link S_{BH}, the maximum coverage radius of the MBS is also found according to the analysis of Section 5.3. For each value $S_{BH} \in \mathcal{S}$, Section 5.3 yields a coverage disk $D_{MBS,S_{BH}}$ centered at the MBS and having $C_{MBS,S_{BH}}$ as radius.

5.5.2 Proposed Solution

5.5.2.1 Deciding on the Subband Assignment in the Backhaul Link

As in Section 5.4.2.1, the minimum number of required backhaul subbands is first found. Having found the coverage regions of users and of the MBS in the same manner as in Section 5.4, Proposition 6 always holds. Hence, the minimum number of backhaul subbands $S_{BH,\min}$ is the smallest value of $S_{BH} \in \mathcal{S}$ yielding a coverage disk $D_{MBS,S_{BH}}$ intersecting with the largest simultaneous coverage region of all users $D_{\text{int}}^{\max}$, i.e., satisfying Eq. (5.38). Value $S_{BH,\min}$ is found using bisection search.

To calculate the total needed power in the backhaul link for each $S_{BH} \in \mathcal{S}_{BH,min}$, we first note that, as stated in Proposition 5, the needed coverage radius of the MBS is upper bounded by the value $C_{MBS,max}$. For all values of the number of backhaul subbands $S_{BH} \in \{S_{BH,min}, \ldots, S\}$ yielding a coverage radius $C_{MBS,S_{BH}}$ lower than the maximum value, the needed power in the backhaul link is the total MBS power budget. However, for S_{BH} values leading to a coverage radius larger than the maximum needed value, the coverage radius $C_{MBS,S_{BH}}$ is substituted with $C_{MBS,max}$ and the needed power in the backhaul link is reduced with respect to the MBS power budget. In this case, the power per subband in the backhaul link is given by:

$$P_{MBS,s} = \left(2^{\frac{R_{UAV,s}}{Bc}} - 1\right) \frac{N_0 B_c + C_{SI} P_{UAV,s}}{10^{-L_{MBS,\max}/10}}, \tag{5.66}$$

where $L_{MBS,\max}$ is the path loss value associated with $C_{MBS,max}$. The total needed MBS power when S_{BH} subbands are used in the backhaul link is hence given by:

$$P_{MBS,S_{BH}}^{\text{needed}} = S_{BH} \times P_{MBS,s}. \tag{5.67}$$

Having found the minimum required number of backhaul subbands and the backhaul power needed for every potential value of backhaul subbands number, we now turn our attention to the choice of these subbands. Clearly, this choice should be done carefully

since it has an undeniable impact on the interference levels experienced by the users. To this end, a metric is introduced for the choice of the backhaul subbands.

As shown in Eq. (5.9), every user has an upper bound on its acceptable path loss, given by L_k^{th}. The latter depends on the requested data rate of user k, the BI and the power of the UAV on the subband assigned to k. To tend towards the best achievable performance (i.e., lowest interference levels), the upper bound of each user's path loss should be as high as possible, since a larger upper bound reflects in a larger coverage region. Therefore, the backhaul subbands should be chosen to guarantee that the resulting upper bounds are as large as possible. For user k, L_k^{th} increases when the term inside the brackets in Eq. (5.9) decreases. Therefore, each user k is associated with the following metric:

$$MU(k) = \left(2^{\frac{R_k^{\text{req}}}{B_c}} - 1\right)\left(N_0 B_c + P_{MBS,s} G_{k,s_k}^2\right). \tag{5.68}$$

Note that $P_{UAV,s}$ is removed from Eq. (5.68) since it is common to all users (considering equal initial power allocation).

To optimize system performance, when S_{BH} subbands are used in the backhaul link, the subbands belonging to the S_{BH} users having the lowest metric given by Eq. (5.68) are chosen.

Once the assignment of backhaul subbands is known, the path loss upper bounds taking into account the BI at the user side can be calculated for all users according to Eq. (5.9), for all potential values of backhaul subbands numbers, i.e., for $S_{BH} \geq S_{\text{min,BH}}$. Let the set of path loss upper bounds be denoted by $L^{\text{ub}} \in \mathbb{R}^{K \times |\mathcal{S}_{BH}^{\text{poss}}|}$, where $\mathcal{S}_{BH}^{\text{poss}}$ is the set containing the potential values of backhaul subbands numbers, i.e., $\mathcal{S}_{BH}^{\text{poss}} = \{S_{BH,\min}, S_{BH,\min}+1, \ldots, S\}$. Next, the value of S_{BH} ensuring the best system performance is found.

The minimum number of backhaul subbands $S_{BH,\min}$ was found without considering BI. When accounting for this interference, the coverage regions of users decrease in size. Therefore, it is not guaranteed that all subbands numbers in $\mathcal{S}_{BH}^{\text{poss}}$ ensure that the coverage regions of users intersect with that of the MBS on the one hand, and with each other on the other. Hence, an algorithm to find the value of the number of backhaul subbands resulting in the best performance is introduced in the following.

First, to guarantee the intersection of the MBS coverage region with that of all users, the following observation is made.

Proposition 7. *For S_{BH} backhaul subbands, the coverage region of the MBS intersects with that of user k if:*

$$L_{k,S_{BH}}^{ub} \geq L_{k,S_{BH}}^{min}. \tag{5.69}$$

Proof. The coverage region of the MBS intersects with that of user k if:

$$C_{k,S_{BH}} \geq C_{k,S_{BH}}^{\min} = \sqrt{x_k^2 + y_k^2} - C_{MBS,S_{BH}}, \tag{5.70}$$

where $C_{k,S_{BH}}$ is the coverage radius of user k accounting for the BI when S_{BH} subbands are used in the backhaul link.

From $C_{k,S_{BH}}^{\min}$, we can find the minimum accepted path loss $L_{k,S_{BH}}^{\min}$ which results in the intersection of the two coverage regions. The value of $L_{k,S_{BH}}^{\min}$ relates to the maximum accepted rate requirement for user k in order for the UAV to be able to serve him, when S_{BH} subbands are used in the backhaul link. Having L^{ub} as the upper bound on achieved path loss for all potential values $S_{BH} \in \mathcal{S}_{BH}^{\text{poss}}$, the coverage region of user k intersects with that of the MBS if Eq. (5.69) holds. □

Any potential value of S_{BH} that does not verify condition (5.69) for all users is removed from $\mathcal{S}_{BH}^{\text{poss}}$.

Although all values of S_{BH} remaining in $\mathcal{S}_{BH}^{\text{poss}}$ guarantee the intersection of the coverage region of the MBS with the coverage regions of all users, not all of these values guarantee that the coverage regions of users intersect with each other. In other words, not all of these potential values guarantee that the users can be served simultaneously while receiving the necessary information from the MBS. Therefore, an algorithm that finds the number of backhaul subbands resulting in the best coverage of all users simultaneously is introduced next.

Algorithm 11 Finding the final number of backhaul subbands

Output: $S_{BH}^{f}, X_{S_{BH}^{f}}$.

Initialization: $size \in \mathbb{R}^{|\mathcal{S}_{BH}^{\text{poss}}|\times 1}$.

1: **for** $i = 1 : |\mathcal{S}_{BH}^{\text{poss}}|$ **do**
2: $S_{BH} = \mathcal{S}_{BH}^{\text{poss}}(i)$.
3: $X_{S_{BH}} = D_{MBS,S_{BH}}$. // $D_{MBS,S_{BH}}$ is the coverage disk of the MBS for S_{BH} backhaul subbands
4: **for** $k = 1 : K$ **do**
5: Find radius C_k from Eq. (5.11) after setting $L_k^{\text{th}} = L_{k,S_{BH}}^{\text{ub}}$.
6: $X_{S_{BH}}$ =check_intersection$(X_{S_{BH}}, x_k, y_k, C_k)$.
7: **end for**
8: $size(S_{BH})$= area$(X_{S_{BH}})$.
9: **end for**
10: $S_{BH}^{f} = \underset{S_{BH}\in\mathcal{S}_{BH}^{\text{poss}}}{\text{argmax}}\ size(S_{BH})$.

Algorithm 11 finds, for each value of $S_{BH} \in \mathcal{S}_{BH}^{\text{poss}}$, the area of the intersection region of all users (steps 3 through 8). The intersection region is first initialized as the coverage disk of the MBS for the tested S_{BH} value (step 3). Then, for each user $k \in \mathcal{K}$, the function "check_intersection" finds the resulting intersection region with the coverage disk of user k, i.e., the disk centered at user k (i.e., centered at (x_k, y_k)) and having C_k as radius (step 6). The values of the number of backhaul subbands guaranteeing simultaneous service to all users are those having $size > 0$. After finding the intersection area for every $S_{BH} \in \mathcal{S}_{BH}^{\text{poss}}$, the value of S_{BH} that maximizes the simultaneous coverage region, S_{BH}^{f}, as well as this coverage region $X_{S_{BH}^{f}}$, are retained.

5.5.2.2 Finding the Optimal UAV Position and the Final Power Levels

With the subband assignment in the backhaul link found in Section 5.5.2.1, the final UAV position and the power levels in the access and the backhaul links are determined according to the solution proposed in Section 5.4.2.3.

5.6 Simulation Results

Extensive simulations of the proposed framework were conducted, where users are randomly located within a squared urban area of size 1 km×1 km. The MBS is located at

the bottom left corner, as shown in Fig. 5.1. The simulation parameters are listed in Table 5.1, with the urban environment constants set as in [62]. Users are divided into two different QoS classes. For $K = 32$ users, the first class corresponds to rate requirements ranging between 4 and 4.8 Mbps, with 0.2 Mbps increment. The second class of users consists of rates requirements between 9 and 9.8 Mbps, with 0.2 Mbps increment.

Table 5.1 – Simulation Parameters

Parameter	**Value Range**	**Parameter**	**Value Range**
(α, β)	$(9.61, 0.16)$	(η_L, η_{NL})	$(1, 20)$ dB
f_c	2 GHz	BW	20 MHz
$P_{MBS}^{\max}$	2 to 8 W	$H_{\min}$	100 m
$P_{UAV}^{\max}$	0.5 to 3 W	$H_{\max}$	800 m
K	8, 16, 32, 64	$R^{\text{req}} = \sum_k R_k^{\text{req}}$	208 to 233.6 Mbps
User-MBS Path Loss	$128.1 + 37.6\times \log_{10}(d[Km])$	Root-mean-square (RMS) delay spread	500 ns
$1/C_{SI}$	130 dB	N_0	-174 dBm/Hz

5.6.1 Compared Methods

The method proposed in Section 5.4 of this chapter is denoted by OptPInit&MinP-NOMA. The performance of a variant of this method that does not perform the NOMA pairing step detailed in Algorithm 10 is also shown, and is denoted by OptPInit&MinP-OMA. OptPInit&MinP-OMA solves for the UAV placement, channel and power assignment according to the proposed solution of Sections 5.4.1 and 5.4.2. If the used power in the access link exceeds the UAV power budget, problem (5.52) is solved to maximize the achieved rates, in the OMA setting, while enforcing the UAV power budget.

For comparison, two different methods are also simulated:

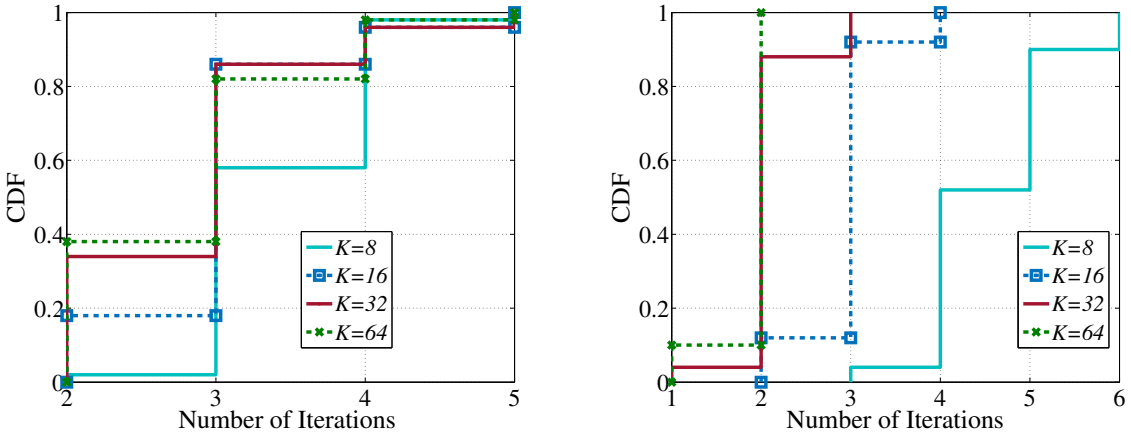
- OBA-PSO: this solution is based on the study of [25]. In OBA-PSO, the available frequency band is divided orthogonally between the access and the backhaul links to avoid inter-link interference. The spectrum division is optimized to minimize the spectrum portion used solely in the backhaul link. In more detail, the amount of spectrum needed to achieve the backhaul rate requirement is first determined. Then, the remaining amount of bandwidth is equally divided into K subbands, before assigning one subband to each user. The 3D placement of the UAV is then conducted using the particle swarm optimization (PSO) algorithm.

- EqPInit&MaxInt: this is the simplified solution presented in Section 5.5 of this chapter. As EqPInit&MaxInt does not optimize the subband assignment in the access link, for a fairer comparison with OptPInit&MinP-NOMA, we assume that the subband assignment in the access link is determined according to Section 5.4.1.1, i.e., similarly to OptPInit&MinP-NOMA. Recall that EqPInit&MaxInt assumes that the available UAV power budget is sufficient to guarantee user rate requirements, making

it unsuitable for systems with high user rate requirements. To compare the performance of EqPInit&MaxInt with OptPInit&MinP-NOMA, when EqPInit&MaxInt fails to find a solution that respects the UAV power budget, the rate requirement of all users is gradually reduced by the amount $\delta R = 0.25$ Mbps, until a solution is found.

5.6.2 Convergence of the Proposed Technique

First, the convergence of both the iterative method of Algorithm 9 and problem (5.50) is analyzed. Fig. 5.4 shows the number of needed iterations before convergence is reached in both Algorithm 9 and problem (5.50).



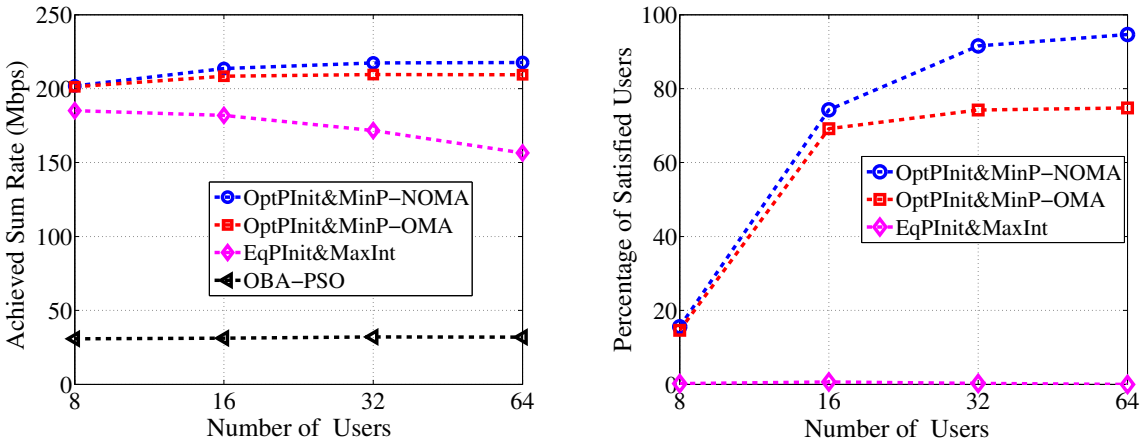
(a) (b)

Figure 5.5 – Number of needed iterations for the convergence of (a) problem (5.50), (b) Algorithm 9.

In Fig. 5.5a, the cumulative distribution function (CDF) of the number of needed iterations for the convergence of problem (5.50) is plotted for various values of the number of users K. Fig. 5.5a shows that problem (5.50) converges within a very small number of iterations (less than 5) for $K = 8, 16, 32$ and 64 users, hence, fairly quickly. Fig. 5.5b plots the CDF of the number of needed iterations before Algorithm 9 converges for various values of K. Fig. 5.5b also shows that the iterative approach of Algorithm 9 converges within a very small number of iterations (less than 6) for all K values.

5.6.3 Performance of the Proposed Technique

Fig. 5.6 compares the performance of the different methods for a number of users K ranging between 8 and 64. It was assumed that 75% of the users have a sum rate requirement of 132 Mbps while the remaining 25% have a sum rate requirement of 88 Mbps, resulting in $R^{\text{req}} = 220$ Mbps, for all values of K. Fig. 5.6a shows that OBA-PSO results in the lowest achieved sum rate for all considered values of K. In fact, because of the orthogonal spectrum division between the access and the backhaul links, the amount of bandwidth assigned to each user is inherently smaller than the one in our proposed method. Constrained by the UAV power budget, users cannot be served with a sufficient data rate.



(a) (b)

Figure 5.6 – (a) Achieved sum rate, and (b) percentage of satisfied users, in terms of K for $R^{\text{req}} = 220$ Mbps, $P_{UAV}^{\max} = 1W$ and $P_{MBS}^{\max} = 4W$.

In contrast, our proposed methods based on IBFD wireless backhauling achieve much higher data rates. However, EqPInit&MaxInt achieves the lowest data rate among these methods because of the initial equal power repartition and the reduction of data rates when the UAV power budget $P_{UAV}^{\max}$ is not sufficient to satisfy user requirements. As K increases, the achieved throughput of EqPInit&MaxInt deteriorates because the data rate of each user is reduced by a constant amount of δR until a solution satisfying $P_{UAV}^{\max}$ is found. On the other hand, the achieved throughput of both OptPInit&MinP-OMA and OptPInit&MinP-NOMA increases with K because these methods can better exploit multi-user diversity, with OptPInit&MinP-NOMA outperforming its OMA counterpart. For $K = 64$, OptPInit&MinP-NOMA outperforms OptPInit&MinP-OMA, EqPInit&MaxInt and OBA-PSO by 8, 61 and 186 Mbps, respectively. When it comes to the percentage of users having received their data rate, i.e., the percentage of satisfied users, Fig. 5.6b shows that EqPInit&MaxInt can hardly satisfy any user. The performance of OBA-PSO is not shown since the satisfaction percentage is equal to zero for all values of K. OptPInit&MinP-NOMA achieves an average satisfaction percentage of 94% for $K = 64$ users, outperforming OptPInit&MinP-OMA by almost 20%, which shows the benefit of the proposed NOMA pairing.

Remark 4. *Regarding the low value of the satisfaction percentage at $K = 8$ users, recall that the number of subbands S is assumed to be equal to the number of users K. For a small number of subbands S, the amount of BI present at the quasi-totality of subbands prevents most users from achieving their target. Moreover, since $R^{req} = 220$ Mbps for every considered value of K, the rate requirement per user is higher for smaller values of K, rendering difficult the satisfaction of users.*

In Fig. 5.7, the needed UAV transmit power is shown in terms of the number of users K. The power for EqPInit&MaxInt is a decreasing function of K because the achieved rate also decreases with K for this method. Both OptPInit&MinP-OMA and OBA-PSO consume the total budget to maximize the achieved rate. On the other hand, the needed power for OptPInit&MinP-NOMA is lower than that of its OMA counterpart

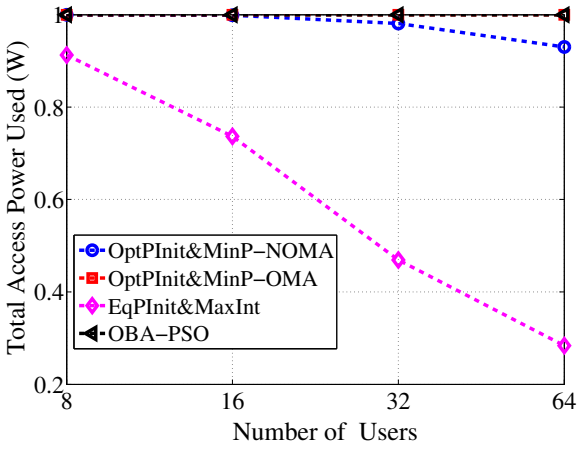


Figure 5.7 – Used access power in terms of K for $R^{\text{req}} = 220$ Mbps, $P_{UAV}^{\max} = 1W$ and $P_{MBS}^{\max} = 4W$.

when $K = 32$ or 64 users. Hence, the NOMA pairing step is not only able to increase the achieved data rate, but can do so while consuming less transmit power.

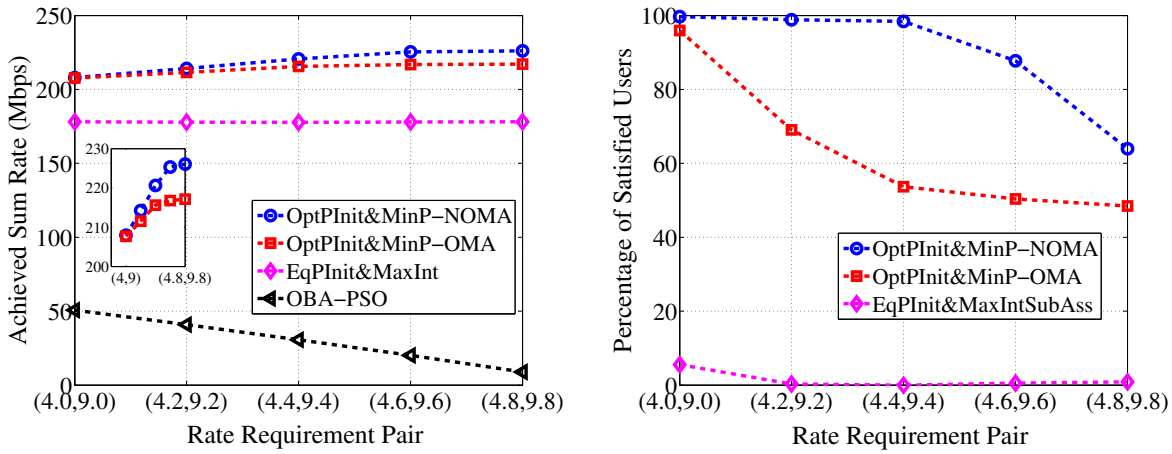
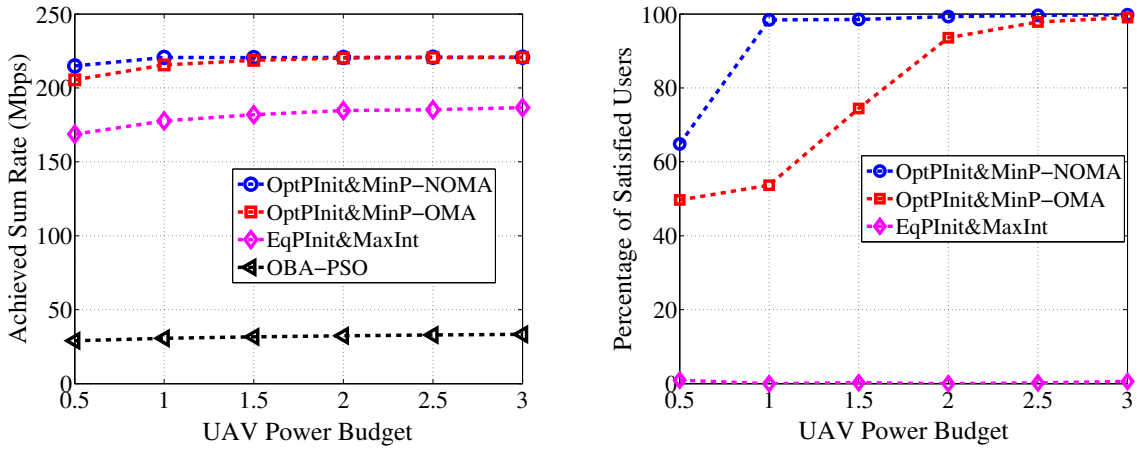


Figure 5.8 – (a) Achieved sum rate, and (b) percentage of satisfied users, in terms of R_k^{req} for $K = 32$ users, $P_{UAV}^{\max} = 1W$ and $P_{MBS}^{\max} = 4W$.

Fig. 5.8 compares the performance of the different methods for $K = 32$ users with a varying data rate requirement. The K users are equally partitioned into 2 classes based on their data rate requirements, shown on the x-axis of Fig. 5.8a and 5.8b. As the target rate increases, the achieved rate of OBA-PSO decreases as more bandwidth is needed in the backhaul link to meet the backhaul rate requirement. OBA-PSO is also unable to satisfy any user. OptPInit&MinP-OMA and OptPInit&MinP-NOMA achieve the best performance, with OptPInit&MinP-NOMA outperforming OptPInit&MinP-OMA and EqPInit&MaxInt by up to 9 and 48 Mbps, respectively, in terms of data rate. In terms of percentage of users having reached their rate requirements, OptPInit&MinP-NOMA outperforms OptPInit&MinP-OMA and EqPInit&MaxInt by up to 45% and 98%,

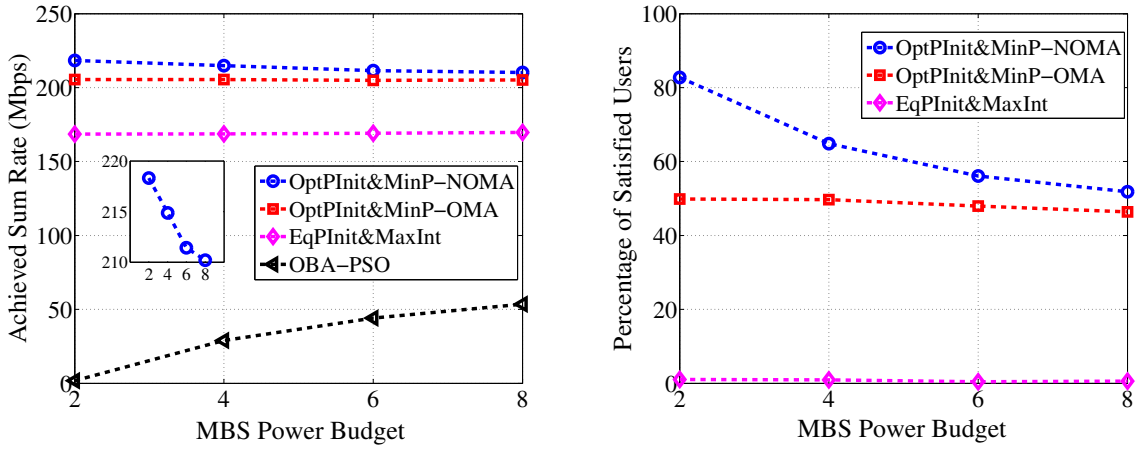
respectively.



(a) (b)

Figure 5.9 – (a) Achieved sum rate, and (b) percentage of satisfied users, in terms of $P_{UAV}^{\max}$ for $K = 32$ users and $P_{MBS}^{\max} = 4W$.

Fig. 5.9 compares the performance of the different methods for a varying power budget of the UAV, $P_{UAV}^{\max}$. The number of users is $K = 32$, half of which have a rate requirement of 4.4 Mbps, while the other half requires 9.4 Mbps to be satisfied. Although the rate achieved by EqPInit&MaxInt increases with $P_{UAV}^{\max}$, the percentage of satisfied users for EqPInit&MaxInt remains very low as shown in Fig. 5.9b. Regarding the OptPInit&MinP-NOMA method, it outperforms the OMA version when $P_{UAV}^{\max}$ is low, however the gap diminishes when the $P_{UAV}^{\max}$ increases as OMA scheduling has a higher chance of satisfying all users for a larger UAV power budget.



(a) (b)

Figure 5.10 – Achieved sum rate (a), and percentage of satisfied users (b), in terms of $P_{MBS}^{\max}$ for $K = 32$ and $P_{UAV}^{\max} = 0.5W$.

Fig. 5.10 compares the performance of the different methods in terms of the MBS power budget, $P_{MBS}^{\max}$. The system consists of $K = 32$ users with the same rate requirements as

in Fig. 5.9. Fig. 5.10a shows that the rate achieved by OBA-PSO increases with $P_{MBS}^{\max}$. In fact, for a larger $P_{MBS}^{\max}$ value, the amount of needed spectrum for the backhaul link decreases, allowing users a larger bandwidth in the access link. However, OBA-PSO is still unable to satisfy any user. On the other hand, the achieved rate of OptPInit&MinP-NOMA decreases with $P_{MBS}^{\max}$. This is due to NOMA being less beneficial when a higher BI, resulting from a higher $P_{MBS}^{\max}$, is experienced. Nonetheless, OptPInit&MinP-NOMA still outperforms OptPInit&MinP-OMA, EqPInit&MaxInt and OBA-PSO by up to 13, 50, and 217 Mbps respectively, in terms of throughput as shown in Fig. 5.10a. In terms of users having reached their required rates, Fig. 5.10b shows that OptPInit&MinP-NOMA outperforms OptPInit&MinP-OMA, EqPInit&MaxInt and OBA-PSO by up to 33, 81, and 83% respectively.

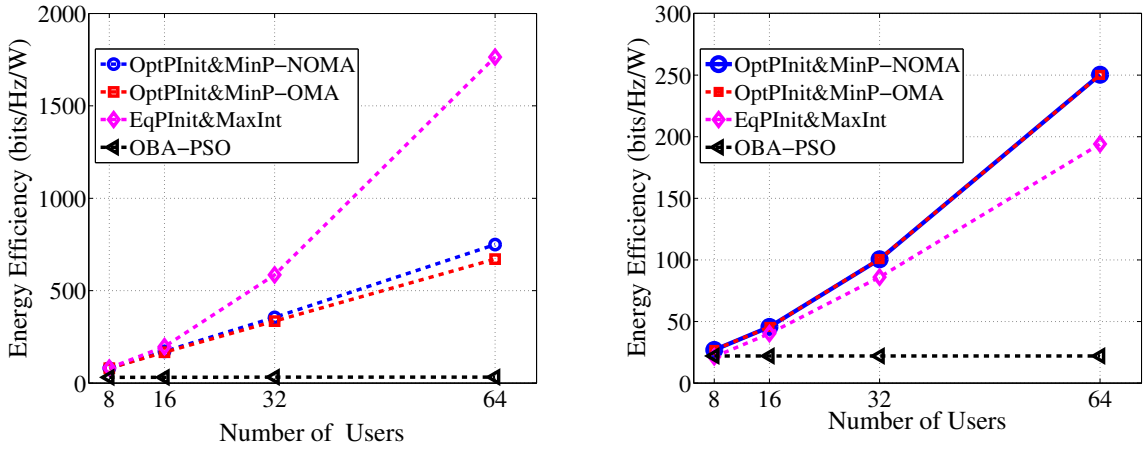


Figure 5.11 – Achieved energy efficiency in terms of the number of users K, for $R^{\text{req}} = 220$Mbps, $P_{UAV}^{\max} = 1$W, and $P_{MBS}^{\max} = 4$W, (a) at the UAV level, and (b) at the MBS level.

Finally, in Fig. 5.11, the energy efficiency (EE) achieved by the different considered methods in terms of the number of users K, at both the UAV and the MBS levels, is plotted. The system configuration considered for Fig. 5.11 consists of K users with a sum rate requirement R^{req} equal to 220 Mbps. Users are partitioned into two classes based on their rate requirement where 75% of the users have a sum rate requirement equal to 132 Mbps while the remaining 25% have a sum rate requirement of 88 Mpbs. The UAV power budget is $P_{UAV}^{\max} = 1$W, while the MBS power budget is chosen to be $P_{MBS}^{\max} = 4$W.

At the UAV level, the EE is given by:

$$EE_{UAV} = \frac{\sum\limits_{k,s} R_{k,s}}{\sum\limits_{k,s} P_{UAV,k,s}}. \tag{5.71}$$

For $K = 8$ and 16 users, EqPInit&MaxInt, OptPInit&MinP-OMA, and OptPInit&MinP-NOMA achieve almost the same EE at the UAV level as shown in Fig. 5.11a. When the number of users increases reaching $K = 32$ and 64 users, EqPInit&MaxInt outperforms both OptPInit&MinP-OMA and OptPInit&MinP-NOMA in terms of EE. However, as shown by Fig. 5.6, where the same simulation parameters are adopted, EqPInit&MaxInt is not able to guarantee the user rate requirements for any value of K. Moreover, the total

rate achieved by EqPInit&MaxInt decreases with K, as validated by Fig. 5.6a. This rate decrease is due to the reduction of the rate requirements by a constant value equal to δR until a feasible solution is found. The decrease in the achieved rate is met with a decrease in the power needed by EqPInit&MaxInt, as shown in Fig. 5.7. In fact, when K increases, since the total rate requirement of all users is kept constant at 220 Mbps, the rate required by each user decreases. This translates into a larger decrease in the used power for higher K values because of the logarithmic relationship between the achieved rate and the consumed power. Hence, the EE at the UAV level increases for EqPInit&MaxInt.
On the other hand, the achieved rate of both OptPInit&MinP-OMA and OptPInit&MinP-NOMA is larger than that of EqPInit&MaxInt and the gap in the achieved rate increases with K. To achieve the higher data rates, OptPInit&MinP-OMA and OptPInit&MinP-NOMA require a higher power level than EqPInit&MaxInt, as shown in Fig. 5.7. Moroever, both OptPInit&MinP-OMA and OptPInit&MinP-NOMA outperform EqPInit&MaxInt in terms of the percentage of satisfied users as shown in Fig. 5.6b. Hence, the gain in EE obtained with EqPInit&MaxInt towards OptPInit&MinP-OMA and OptPInit&MinP-NOMA is explained by the superior performance of the two latter methods in terms of increasing the user achieved rate and the level of user satisfaction, the primary objectives of the proposed solution in this chapter.

Regarding the EE at the MBS level, it is given by:

$$EE_{MBS} = \frac{\sum_{k,s} R_{UAV,s}}{\sum_{s} P_{MBS,s}}. \tag{5.72}$$

We can see from Fig. 5.11b that the EE achieved by OBA-PSO at the MBS level is constant in terms of the number of users K. The constant EE is due to OBA-PSO aiming to minimize the amount of spectrum used in the backhaul link. This translates into OBA-PSO requiring the whole MBS power budget to achieve the backhaul capacity, which is constant for any value of K, regardless of the number of users. Moreover, we can see that both OptPInit&MinP-OMA and OptPInit&MinP-NOMA achieve a similar performance in terms of K, one that is superior to the performance of EqPInit&MaxInt. In fact, the UAV position yielded by EqPInit&MaxInt is closer to the users, hence farther from the MBS. Therefore, to achieve the backhaul capacity, EqPInit&MaxInt requires a higher amount of power in the backhaul link, leading to a decrease in the achieved EE at the MBS level. On the other hand, both OptPInit&MinP-OMA and OptPInit&MinP-NOMA succeed in finding a balanced UAV position between the users and the MBS. Hence, the power in the backhaul link needed by these methods is smaller than the one needed by EqPInit&MaxInt, resulting in a better EE at the MBS level for OptPInit&MinP-OMA and OptPInit&MinP-NOMA.

To conclude the EE analysis, the simplified proposed method, EqPInit&MaxInt, outperforms the more elaborate proposed method, OptPInit&MinP-NOMA, in terms of the EE achieved at the UAV level. However, OptPInit&MinP-NOMA outperforms EqPInit&MaxInt in terms of the EE achieved at the MBS level, and at the same time greatly increases both the user rates and the percentage of users having reached their rate requirements.

5.7 Summary

In this chapter, we proposed a novel framework for optimizing UAV-enabled wireless networks while taking into account the wireless backhaul constraint of UAVs. In particular, we considered a geographical area where a terrestrial BS is missing and a FD-UAV-BS is deployed to serve users. An optimization problem aiming at minimizing the power used by the UAV to serve users with their rate requirements was proposed. To do so, the 3D position of the UAV, the subband assignment and the power allocation in both the access and the backhaul links were optimized. NOMA scheduling was also proposed to enhance system performance. A simplified version of the proposed method was also introduced. The simplified version does not optimize the subband assignment in the access link, assumes that the UAV power budget is sufficient to guarantee user rate requirements and does not adopt NOMA scheduling. Through simulations, the proposed solution was shown to converge within a small number of iterations. Moreover, simulation results showed that the proposed method yields significant performance enhancement, when compared with the simplified proposed solution and a previous method from the literature.

5.8 Appendix A

A.1 Proof of Proposition 4

As seen from Eq. (5.16), the MBS coverage radius $C_{MBS,S_{BH}}$ increases when the value of $L^{\text{th}}_{S_{BH}}$, given by Eq. (5.36), does. In addition, $L^{\text{th}}_{S_{BH}}$ increases when the term inside the brackets in Eq. (5.36) decreases. In other words, $L^{\text{th}}_{S_{BH}}$ increases when (5.73) decreases:

$$F(S_{BH}) = \left(2^{\frac{R^{\text{req}}}{S_{BH}\times B_c}} - 1\right) \times S_{BH}/P^{\max}_{MBS}. \tag{5.73}$$

The first derivative of $F(S_{BH})$ with respect to S_{BH} is given by:

$$\frac{\partial F}{\partial S_{BH}} = \frac{1}{P^{\max}_{MBS}} \left(2^{\frac{R^{\text{req}}}{S_{BH}\times B_c}} \left(1 - \frac{R^{\text{req}}\log(2)}{S_{BH}\times B_c}\right) - 1\right). \tag{5.74}$$

Hence, F decreases with the increase of S_{BH} if:

$$g(S_{BH}) = 2^{\frac{R^{\text{req}}}{S_{BH}\times B_c}} \left(1 - \frac{R^{\text{req}}\log(2)}{S_{BH}\times B_c}\right) < 1. \tag{5.75}$$

The first derivative of $g(S_{BH})$ with respect to S_{BH} is given by:

$$\frac{\partial g}{\partial S_{BH}} = \frac{(R^{\text{req}})^2\log(2)^2}{(S_{BH})^3(B_c)^2} \times 2^{\frac{R^{\text{req}}}{S_{BH}\times B_c}} > 0, \tag{5.76}$$

i.e., $g(S_{BH})$ is a strictly increasing function of S_{BH}. Moreover,

$$\lim_{S_{BH}\to\infty} g(S_{BH}) = 1. \tag{5.77}$$

Hence, the condition of Eq. (5.75) is always guaranteed and F is a decreasing function of S_{BH}. Thus, $C_{MBS,S_{BH}}$ increases with S_{BH}.

A.2 Solution to Problem (5.60)

A.2.1 Subband n is not used in the backhaul link

In this case, constraint (5.60d) comes down to ensuring:

$$R'_{UAV,s_k} = R_{UAV,s_k}. \tag{5.78}$$

If $b_{s_k} = 0$, i.e., if subband s_k is not used in the backhaul link, Eq. (5.78) and hence constraint (5.60d) are automatically satisfied. On the other hand, if $b_{s_k} = 1$, with the decrease of the value of P'_{UAV,k,s_k} resulting from the division of P'_{UAV,k,s_k} between subbands s_k and n, the SI experienced by the UAV on subband s_k decreases. To keep constant the rate of the UAV in the backhaul link, R_{UAV,s_k}, hence ensuring constraint (5.60d), P'_{MBS,s_k} should decrease as per Eq. (5.5). Therefore, constraint (5.60c) is always met when the total access power allocated to user k decreases. According to Eq. (5.49), P'_{UAV,k,s_k} can be formulated as:

$$P'_{UAV,k,s_k} = \frac{(2^{R'_{k,s_k}/B_c} - 1)\, J_1}{J_2 - J_3\,(2^{R'_{k,s_k}/B_c} - 1)}. \tag{5.79}$$

where

$$\begin{cases} J_1 = N_0 B_c \left(G^2_{UAV,MBS} + a_{k,s}\, A_2(s_k)\, G^2_{k,s_k} \right), \\ J_2 = G^2_{UAV,MBS}\, G^2_{UAV,k}, \\ J_3 = a_{k,s}\, G^2_{k,s_k}\, A_2(s_k)\, C_{SI}. \end{cases} \tag{5.80}$$

From Eq. (5.53), when $b_n = 0$, the power needed on subband n to achieve a rate equal to $R'_{k,n}$ is found according to:

$$P'_{UAV,k,n} = (2^{R'_{k,n}/B_c} - 1) J_4, \tag{5.81}$$

where $J_4 = N_0 B_c \,/\, G^2_{UAV,k} + P_{UAV,j,n}$.

Let $(2^{R'_{k,s_k}/B_c} - 1) = x$. In order to ensure constraint (5.60a), $P'_{UAV,k,n}$ should take the following value:

$$P'_{UAV,k,n} = (\frac{2^{R_{k,s_k}/B_c}}{x+1} - 1) J_4. \tag{5.82}$$

Then, the total needed power to achieve a rate equal to R_{k,s_k}, on subbands s_k and n, is given by:

$$\begin{aligned} P'_k(x) &= P'_{UAV,k,s_k}(x) + P'_{UAV,k,n}(x) \\ &= \frac{x J_1}{J_2 - x J_3} + (\frac{2^{R_{k,s_k}/B_c}}{x+1} - 1) J_4. \end{aligned} \tag{5.83}$$

Function P'_k is convex in x as its second derivative $\frac{\partial^2 P'_k}{\partial^2 x}$ is positive. The value of x that minimizes P'_k is found by setting:

$$\frac{\partial P'_k}{\partial x} = 0. \tag{5.84}$$

As a result, variable x takes the following value:

$$x^* = \frac{J_2 \sqrt{J_4\, 2^{R_{k,s_k}}} - \sqrt{J_1\, J_2}}{J_3 \sqrt{J_4\, 2^{R_{k,s_k}}} + \sqrt{J_1\, J_2}}. \tag{5.85}$$

Having x^*, the minimum updated individual power values on s_k and n as well as the total needed power by user k are found using Eq. (5.79), Eq. (5.82) and Eq. (5.83), respectively. If these power values satisfy the SIC stability constraint in (5.60b), then subband n is a valid candidate for NOMA pairing of user k. In the opposite case, x_{SIC}, the value of x that enforces constraint (5.60b) at equality is found. Subband n is a valid candidate in this case if:

$$P'_k(x_{SIC}) < P_{UAV,k,s_k}. \tag{5.86}$$

In fact, in this case, we may have $x_{SIC} < x^*$ (in the decreasing part of $P'_k(x)$) or $x_{SIC} > x^*$ (in the increasing part). But in both cases, $P'_k(x_{SIC}) > P'_k(x^*)$.

A.2.2 Subband n is used in the backhaul link

The backhaul power on subband n is kept constant, as per assumption 3. Therefore, when user k is scheduled on subband n, the SI term in the denominator of Eq. (5.5) increases, leading to:

$$R'_{UAV,n} < R_{UAV,n}. \tag{5.87}$$

To ensure constraint (5.60d), R'_{UAV,s_k} must be increased with respect to R_{UAV,s_k} while respecting constraint (5.60c). Hence, in this case, R'_{UAV,s_k} is not constant, and by extension neither is term $A_2(s_k)$ in Eq. (5.79). In fact, from Eq. (5.5), $R'_{UAV,n}$ is a decreasing function of $P'_{UAV,k,n}$ leading to R'_{UAV,s_k} and $A_2(s_k)$ being increasing functions of $P'_{UAV,k,n}$. This leads to a set of non-linear expressions for R'_{UAV,s_k} and $R'_{UAV,n}$. Obtaining closed-form expressions of R'_{UAV,s_k} and $R'_{UAV,n}$ that minimize the total needed power is hence cumbersome. Therefore, problem (5.60) is solved numerically to find whether pairing k on subband n can increase its total achieved rate when subband n is used in the backhaul link.

It should be noted that due to the decrease of the rate $R'_{UAV,n}$ when subband n is used in the backhaul link, only subbands s_k that are used in the backhaul link, i.e., that satisfy $b_{s_k} = 1$, are considered.

Chapter 6

Conclusions and Future Work

In this thesis, we addressed several challenging resource management problems for next-generation wireless networks. These problems can be summarized as: 1) resource allocation for mixed-traffic systems benefiting from a distributed antenna configuration and non-orthogonal multiple access (NOMA) scheduling, 2) uncoordinated channel access for grant-free communications, 3) uncoordinated channel access and power control for self-organizing networks (SONs) benefiting from NOMA, 4) resource allocation and positioning for unmanned aerial vehicle (UAV)-aided communication networks benefiting from NOMA and with backhaul consideration. Next, we present a summary of the research that was carried out in this dissertation.

6.1 Summary

6.1.1 Resource Allocation for Mixed Traffic Systems

In Chapter 3, we studied a wireless communication system consisting of users having heterogeneous traffic characteristics. The first user type consists of best-effort (BE) users aiming to maximize their achieved rates while preserving system fairness. The second user type consists of real-time (RT) users aiming to receive a certain quantity of data bits before the expiration of their latency limit. To enhance system performance, a distributed antenna system (DAS) benefiting from NOMA scheduling was considered. To solve the resource allocation problem, we first proposed a low-complexity greedy algorithm that aims at first satisfying RT users, and when possible, serving BE users. We then focused on the subband and antenna assignment problem and proposed a solution based on matching theory to solve this problem. In addition to finding the complexity of the matching theory-based method, we proved its convergence and stability. Using simulation results, we showed that both proposed solutions greatly enhance system performance especially when it comes to increasing the satisfaction level of RT users, outperforming conventional methods.

6.1.2 Uncoordinated Spectrum Access using Multi-Armed Bandits

In Chapter 4, we studied uncoordinated spectrum access in wireless networks. We first studied a SON and proposed a novel solution for the uncoordinated channel and power

allocation between deployed access points (APs). We considered a NOMA setting for channel access, leading multiple APs to be able to access the same channel and receive a non-zero reward. We modeled and solved the uncoordinated channel access problem using the multi-player multi-armed bandits (MP-MAB) framework with varying channel rewards across APs, multiple plays and non-zero reward on collision. The power control problem was modeled and solved using the MP-MAB framework with varying channel rewards across APs. We then focused on grant-free communications where multiple users aim to organize their transmissions on multiple chosen channels, without any coordination or communication between them. To this effect, we proposed an algorithm based on the MP-MAB framework with multiple plays. The case of varying channel rewards across users was considered, resulting in a MP-MAB model with varying channel rewards and multiple plays. The aim of the proposed algorithm is to optimize system performance while reducing transmission collisions. Using theoretical derivations, we proved that the proposed methods, in both studies, result in sub-linear regret. We further validated the performance of the proposed algorithms and the theoretical results using numerical simulations.

6.1.3 UAV-Enabled Communication Networks

In Chapter 5, we studied a UAV-aided communication network. A UAV was deployed to serve users that cannot be covered by a terrestrial base station (BS). As a UAV does not have a wired backhaul link to the core network, it must rely on a wireless backhaul link. In this chapter, we gave particular attention to this wireless backhaul link between the UAV and a macro base station (MBS) serving as gateway to the core network. To increase system performance and spectral efficiency, we considered an in-band backhaul, full-duplex capabilities at the UAV and NOMA for multiple access. To optimize the UAV position and solve the resource and power allocation problems in the backhaul and access links, we formulated and solved an optimization problem that minimizes the UAV transmit power. The solved optimization problem takes into account the user rate requirement, the backhaul capacity constraint, and the transmit power budgets for the UAV and the MBS. Using simulation results, we validated the superior performance of the proposed solutions, when compared to previously proposed techniques in the literature.

6.2 Future Work

The proposed solutions in this thesis addressed some of the challenges that arise in the context of resource allocation, spectrum management and the use of UAVs as aerial BSs in next-generation cellular networks. Next, we discuss some open problems that can be addressed as an extension to this work:

- **Intelligent resource allocation algorithms for mixed traffic systems:** Chapter 3 of this thesis proposed two low-complexity solutions for the mixed traffic system. However, the number of users considered in this chapter was limited. The extension of this study to large-scale networks in very interesting since future communication networks are expected to provide service to billions of connected machine-type devices (MTDs) [9], in addition to human users. In large-scale networks, introducing intelligent entities in the network would greatly enhance system performance. In

fact, these intelligent entities can leverage machine learning algorithms to optimize resource allocation, ensuring the quality of service (QoS) requirements of different devices are met.

- **Low-complexity resource allocation algorithms for grant-free communications:** In Chapter 4, using the multi-armed bandits (MAB) framework, we presented a game-theoretic solution for the uncoordinated spectrum access. However, this solution is not applicable in large-scale networks. For Internet of Things (IoT) networks requiring grant-free communications, coming up with low-complexity algorithms to solve the resource allocation problem is of utmost importance. Such algorithms can benefit from the reinforcement learning (RL) framework. Moreover, these algorithms should aim at optimizing different performance measures such as fairness among users, communication latency and reliability.

- **Interference mitigation for UAV-enabled communication networks:** The ability of UAV-BSs to establish line-of-sight (LOS) connectivity with multiple ground users can lead to substantial inter-cell interference. Hence, there is a need to design efficient cell coordination and scheduling techniques to mitigate interference between aerial and terrestrial BSs in UAV-enabled cellular networks.

- **UAV trajectory optimization:** In Chapter 5, the hovering position of the UAV-BS was optimized and NOMA was used whenever the UAV power budget was not sufficient to guarantee rate requirements. However, the simulation results showed that NOMA cannot always guarantee meeting user QoS requirements at all times. An interesting extension of this work would be to optimize the UAV trajectory in a way to serve all users with their rate requirements while minimizing the needed energy.

- **Multi-UAV systems:** A second extension to the work of Chapter 5 would be to deploy multiple UAV-BSs to serve the needed area. A multi-UAV system is much more challenging to design than a cellular system with only one BS. In fact, the needed number of UAV-BSs must first be found, before optimizing the positions of all UAV-BSs. Moreover, the resource allocation algorithms must pay particular attention to the inter-cell interference between the UAV-BSs in a multi-UAV communication system.

Bibliography

[1] P. Popovski, "Ultra-reliable communication in 5g wireless systems," *Proc. Int. Conf. on 5G for Ubiquitous Connectivity (5GU)*, Feb. 2014.

[2] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[3] C. Bockelmann, N. K. Pratas, G. Wunder, S. Saur, M. Navarro, D. Gregoratti, G. Vivier, E. De Carvalho, Y. Ji, C. Stefanovic, P. Popovski, Q. Wang, M. Schellmann, E. Kosmatos, P. Demestichas, M. Raceala-Motoc, P. Jung, S. Stanczak, and A. Dekorsy, "Towards Massive Connectivity Support for Scalable mMTC Communications in 5G Networks," *IEEE Access*, vol. 6, pp. 28 969–28 992, May 2018.

[4] C. Bockelmann, N. Pratas, H. Nikopour, K. Au, T. Svensson, C. Stefanovic, P. Popovski, and A. Dekorsy, "Massive machine-type communications in 5g: physical and mac-layer solutions," *IEEE Communications Magazine*, vol. 54, no. 9, pp. 59–65, Sept. 2016.

[5] P. Popovski, K. F. Trillingsgaard, O. Simeone, and G. Durisi, "5G Wireless Network Slicing for eMBB, URLLC, and mMTC: A Communication-Theoretic View," *IEEE Access*, vol. 6, pp. 55 765–55 779, Sept. 2018.

[6] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, June 2014.

[7] S. Vanka, S. Srinivasa, Z. Gong, P. Vizi, K. Stamatiou, and M. Haenggi, "Superposition coding strategies: Design and experimental evaluation," *IEEE Trans. Wireless Commun.*, vol. 11, no. 7, pp. 2628–2639, July 2012.

[8] J. G. Andrews, "Interference cancellation for cellular systems: a contemporary overview," *IEEE Wireless Commun.*, vol. 12, no. 2, pp. 19–29, April 2005.

[9] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5G wireless networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1617–1655, 2016.

[10] *Game Theory in Wireless and Communication Networks: Theory, Models, and Applications*. Cambridge: Cambridge University Press, 2011.

[11] T. Lattimore and C. Szepesvári, *Bandit Algorithms*. Cambridge Univ. Press, 2020.

[12] A. Anandkumar, N. Michael, A. K. Tang, and A. Swami, "Distributed algorithms for learning and cognitive medium access with logarithmic regret," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 4, pp. 731–745, Apr. 2011.

[13] K. Liu and Q. Zhao, "Distributed learning in multi-armed bandit with multiple players," *IEEE Trans. Signal Process.*, vol. 58, no. 11, pp. 5667–5681, Nov. 2010.

[14] Z. Tian, J. Wang, J. Wang, and J. Song, "Distributed NOMA-based multi-armed bandit approach for channel access in cognitive radio networks," *IEEE Wireless Commun. Lett.*, vol. 8, no. 4, pp. 1112–1115, 2019.

[15] M. Bande and V. V. Veeravalli, "Multi-user multi-armed bandits for uncoordinated spectrum access," 2018. [Online]. Available: arxiv:1807.00867

[16] Y. Lin, T. Wang, and S. Wang, "UAV-Assisted Emergency Communications: An Extended Multi-Armed Bandit Perspective," *IEEE Commun. Lett.*, vol. 23, no. 5, pp. 938–941, March 2019.

[17] D. Gale and L. S. Shapley, "College admissions and the stability of marriage," *The American Mathematical Monthly*, vol. 69, no. 1, pp. 9–15, 1962. [Online]. Available: https://doi.org/10.1080/00029890.1962.11989827

[18] B. Zhang, X. Mao, J. Yu, and Z. Han, "Resource allocation for 5G heterogeneous cloud radio access networks with D2D communication: A matching and coalition approach," *IEEE Trans. Veh. Commun.*, vol. 67, no. 7, pp. 5883–5894, July 2018.

[19] W. Liang, Z. Ding, Y. Li, and L. Song, "User pairing for downlink non-orthogonal multiple access networks using matching algorithm," *IEEE Trans. Commun.*, vol. 65, no. 12, pp. 5319–5332, Dec. 2017.

[20] B. Di, S. Bayat, L. Song, and Y. Li, "Radio resource allocation for downlink non-orthogonal multiple access (NOMA) networks using matching theory," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2015, pp. 1–6.

[21] J. Zhao, Y. Liu, K. K. Chai, Y. Chen, and M. Elkashlan, "Joint subchannel and power allocation for NOMA enhanced D2D communications," *IEEE Trans. Commun.*, vol. 65, no. 11, pp. 5081–5094, Nov 2017.

[22] O. Semiari, W. Saad, S. Valentin, M. Bennis, and B. Maham, "Matching theory for priority-based cell association in the downlink of wireless small cell networks," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Florence, Italy, May 2014, pp. 444–448.

[23] C. He, G. Y. Li, F. Zheng, and X. You, "Energy-efficient resource allocation in OFDM systems with distributed antennas," *IEEE Trans. Veh. Technol.*, vol. 63, no. 3, pp. 1223–1231, March 2014.

[24] R. Jain, D. Chiu, and W. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer systems," *DEC Technical Report 301*, Sept. 1984.

[25] E. Kalantari, I. Bor-Yaliniz, A. Yongacoglu, and H. Yanikomeroglu, "User association and bandwidth allocation for terrestrial and aerial base stations with backhaul considerations," in *Proc. IEEE Annual Int. Symp. on Personal, Indoor, and Mobile Radio Commun. (PIMRC)*, Montreal, QC, Canada, Oct. 2017, pp. 1–6.

[26] A. Azari, P. Popovski, G. Miao, and C. Stefanovic, "Grant-free radio access for short-packet communications over 5G networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2017, pp. 1–7.

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

[28] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "System-level performance evaluation of downlink non-orthogonal multiple access (NOMA)," in *Proc. IEEE Annual Symp. on Personal, Indoor, and Mobile Radio Commun. (PIMRC)*, Sept 2013, pp. 611–615.

[29] A. Benjebbour, Y. Saito, Y. Kishiyama, A. Li, A. Harada, and T. Nakamura, "Concept and practical considerations of non-orthogonal multiple access (NOMA) for future radio access," in *Int. Symp. on Intelligent Signal Process. and Commun. Syst.*, Nov 2013, pp. 770–774.

[30] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, "On the performance of non-orthogonal multiple access in 5G systems with randomly deployed users," *IEEE Signal Process. Lett.*, vol. 21, no. 12, pp. 1501–1505, Dec. 2014.

[31] H. Nikopour and H. Baligh, "Sparse code multiple access," in *Proc. IEEE Annual Int. Symp. on Personal, Indoor, and Mobile Radio Commun. (PIMRC)*, Sept., 2013, pp. 332–336.

[32] S. Chen, B. Ren, Q. Gao, S. Kang, S. Sun, and K. Niu, "Pattern division multiple access—a novel nonorthogonal multiple access for fifth-generation radio networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 4, pp. 3185–3196, Apr. 2017.

[33] Li Ping, Lihai Liu, Keying Wu, and W. K. Leung, "Interleave division multiple-access," *IEEE Trans. Wireless Commun.*, vol. 5, no. 4, pp. 938–947, Apr. 2006.

[34] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, "A survey on 3GPP heterogeneous networks," *IEEE Wireless Commun.*, vol. 18, no. 3, pp. 10–21, June 2011.

[35] J. G. Andrews, H. Claussen, M. Dohler, S. Rangan, and M. C. Reed, "Femtocells: Past, present, and future," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 497–508, Apr. 2012.

[36] A. Ghosh, N. Mangalvedhe, R. Ratasuk, B. Mondal, M. Cudak, E. Visotsky, T. A. Thomas, J. G. Andrews, P. Xia, H. S. Jo, H. S. Dhillon, and T. D. Novlan, "Heterogeneous cellular networks: From theory to practice," *IEEE Commun. Mag.*, vol. 50, no. 6, pp. 54–64, June 2012.

[37] Y. Li, M. Sheng, X. Wang, Y. Shi, and Y. Zhang, "Globally optimal antenna selection and power allocation for energy efficiency maximization in downlink distributed antenna systems," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2014, pp. 3856–3861.

[38] X. Gu, X. Ji, Z. Ding, W. Wu, and M. Peng, "Outage probability analysis of non-orthogonal multiple access in cloud radio access networks," *IEEE Commun. Lett.*, vol. 22, no. 1, pp. 149–152, Jan 2018.

[39] Q. Vien, N. Ogbonna, H. X. Nguyen, R. Trestian, and P. Shah, "Non-orthogonal multiple access for wireless downlink in cloud radio access networks," in *21th European Wireless Conf.*, May 2015, pp. 1–6.

[40] A. Feki, V. Capdevielle, and E. Sorsy, "Self-organized resource allocation for lte pico cells: A reinforcement learning approach," in *Proc. IEEE Veh. Techn. Conf. Spring (VTC)*, Yokohama, Japan, May 2012, pp. 1–5.

[41] M. N. Esfahani and B. S. Ghahfarokhi, "Improving spectrum efficiency in fractional allocation of radio resources to self-organized femtocells using learning automata," in *Int. Symp. on Telecommun.*, Sept.

[42] M. Mozaffari, W. Saad, M. Bennis, Y. Nam, and M. Debbah, "A tutorial on UAVs for wireless networks: Applications, challenges, and open problems," *IEEE Commun. Surveys Tuts.*, pp. 1–28, Mar. 2019.

[43] *Reinforcement Learning: An Introduction.* Cambridge, MA: MIT Press, 1998.

[44] D. Duchemin, J.M. Gorce, and C. Goursaud, "Low complexity Detector for massive uplink random access with NOMA in IoT LPWA networks," in *Proc. IEEE Wireless Commun. and Networking Conf. Wkshp*, Marrakech, Morocco, Apr. 2019, pp. 1–6.

[45] Z. Ding, X. Lei, G. K. Karagiannidis, R. Schober, J. Yuan, and V. K. Bhargava, "A survey on non-orthogonal multiple access for 5G networks: Research challenges and future trends," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2181–2195, Oct. 2017.

[46] X. Fu, M. Pischella, and D. Le Ruyet, "On Gaussian approximation algorithms for SCMA," in *Int. Symp. Wireless Commun. Syst. (ISWCS)*, Oulu, Finland, Aug. 2019, pp. 155–160.

[47] M. Egan, L. Clavier, C. Zheng, M. de Freitas, and J.M. Gorce, "Dynamic interference for uplink SCMA in large-scale wireless networks without coordination," *EURASIP J. Wireless Commun. and Networking*, Aug. 2018.

[48] N. Otao, Y. Kishiyama, and K. Higuchi, "Performance of non-orthogonal access with SIC in cellular downlink using proportional fair-based resource allocation," in *Int. Symp. Wireless Commun. Syst. (ISWCS)*, Aug. 2012, pp. 476–480.

[49] J. Zhu, J. Wang, Y. Huang, S. He, X. You, and L. Yang, "On optimal power allocation for downlink non-orthogonal multiple access systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 12, pp. 2744–2757, Dec. 2017.

[50] H. Zuo and X. Tao, "Power allocation optimization for uplink non-orthogonal multiple access systems," in *Int. Conf. on Wireless Commun. Sig. Process. (WCSP)*, Oct. 2017, pp. 1–5.

[51] M. Al-Imari, P. Xiao, M. A. Imran, and R. Tafazolli, "Uplink non-orthogonal multiple access for 5G wireless networks," in *Int. Symp. Wireless Commun. Syst. (ISWCS)*, Aug. 2014, pp. 781–785.

[52] J. Choi, "NOMA-Based Random Access With Multichannel ALOHA," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 12, pp. 2736–2743, Dec. 2017.

[53] S. Blanchard and B. Kearney, "Application of IRIDIUM(R) system technology to UAV based PCS services to the warfighter," in *IEEE Military Commun. Conf. Proc*, vol. 3, Oct. 1998, pp. 814–817 vol.3.

[54] A. Sanjab, W. Saad, and T. Başar, "Prospect theory for enhanced cyber-physical security of drone delivery systems: A network interdiction game," in *Proc. Int. Conf. on Communications (ICC)*, May 2017, pp. 1–6.

[55] H. Kim and J. Ben-Othman, "A Collision-Free Surveillance System Using Smart UAVs in Multi Domain IoT," *IEEE Commun. Lett.*, vol. 22, no. 12, pp. 2587–2590, Dec. 2018.

[56] A. Zajiè, *Mobile-to-mobile wireless channels*. Artech House, 2012.

[57] Y. Zheng, Y. Wang, and F. Meng, "Modeling and simulation of pathloss and fading for air-ground link of haps within a network simulator," in *Int. Conf Cyber-Enabled Dist. Comp. and Knowledge Discovery*, Beijing, China, Oct. 2013, pp. 421–426.

[58] Q. Feng, E. K. Tameh, A. R. Nix, and J. McGeehan, "Wlcp2-06: Modelling the likelihood of line-of-sight for air-to-ground radio propagation in urban environments," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, San Francisco, CA, USA, Nov. 2006, pp. 1–5.

[59] K. Daniel, M. Putzke, B. Dusza, and C. Wietfeld, "Three dimensional channel characterization for low altitude aerial vehicles," in *Int. Symp. Wireless Commun. Syst.*, Sep. 2010, pp. 756–760.

[60] K. Sasloglou, I. A. Glover, V. Gazis, P. Kikiras, K. Mathioudakis, and I. Andonovic, "Empirical channel models for optimized communications in a network of unmanned ground vehicles," in *IEEE Int. Symp. Sig. Process. Inf. Tech.*, Dec. 2013, pp. 113–118.

[61] E. Yanmaz, R. Kuschnig, and C. Bettstetter, "Channel measurements over 802.11a-based UAV-to-ground links," in *Proc. IEEE Global Commun. Conf. (GLOBECOM) Workshops*, Dec. 2011, pp. 1280–1284.

[62] A. Al-Hourani, S. Kandeepan, and A. Jamalipour, "Modeling air-to-ground path loss for low altitude platforms in urban environments," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Austin, Tx, USA, Dec. 2014, pp. 2898–2904.

[63] A. Al-Hourani, S. Kandeepan, and S. Lardner, "Optimal LAP altitude for maximum coverage," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 569–572, Dec. 2014.

[64] ITU-R, "Rec. p.1410-2 propagation data and prediction methods for the design of terrestrialbroadband millimetric radio access systems," *Series, Radiowave propagation*, 2003.

[65] *Game Theory: Analysis of Conflict.* Harvard University Press, 1991.

[66] P. Pardalos, A. Migdalas, and L. Pitsoulis, *Pareto Optimality, Game Theory and Equilibria.* New York, USA: Springer, 2008, vol. 17.

[67] D. S. G. R. Grimmett, *Probability and Random Processes.* New York, USA: Oxford University Press.

[68] H. P. Young, "The evolution of conventions," *Econometrica*, vol. 61, no. 1, pp. 57–84, 1993. [Online]. Available: http://www.jstor.org/stable/2951778

[69] K.-M. Chung, H. Lam, Z. Liu, and M. Mitzenmacher, "Chernoff-Hoeffding bounds for Markov chains: Generalized and simplified," *29th Symp. Theor. Aspects of Comput. Sci.*, pp. 124–135, Feb. 2012.

[70] M. B. Booth, V. Suresh, N. Michelusi, and D. J. Love, "Multi-armed bandit beam alignment and tracking for mobile millimeter wave communications," *IEEE Commun. Lett.*, vol. 23, no. 7, pp. 1244–1248, July 2019.

[71] A. Garivier and O. Cappé, "The kl-ucb algorithm for bounded stochastic bandits and beyond," in *Proc. Annual Conf. Learning Theory*, vol. 19, Budapest, Hungary, June 2011, pp. 359–376.

[72] S. Agrawal and N. Goyal, "Analysis of thompson sampling for the multi-armed bandit problem," in *Proc. Annual Conf. Learning Theory*, ser. Proceedings of Machine Learning Research, vol. 23, Edinburgh, Scotland, June 2012, pp. 1–26.

[73] P. Auer, N. Cesa-Bianchi, and P. Fischer, "Finite-time analysis of the multi-armed bandit problem," *Machine Learning*, vol. 47, p. 235–256, 2002.

[74] Y. Gu, W. Saad, M. Bennis, M. Debbah, and Z. Han, "Matching theory for future wireless networks: fundamentals and applications," *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 52–59, May 2015.

[75] A. Benjebbour, A. Li, Y. Saito, Y. Kishiyama, A. Harada, and T. Nakamura, "System-level performance of downlink NOMA for future LTE enhancements," in *Proc. IEEE Global Commun. Conf. (GLOBECOM) Workshops*, Dec. 2013, pp. 66–70.

[76] M. Youssef, J. Farah, C. A. Nour, and C. Douillard, "Waterfilling-based resource allocation techniques in downlink non-orthogonal multiple access (NOMA) with single-user MIMO," in *IEEE Symp. on Comput. and Commun. (ISCC)*, Jul. 2017, pp. 499–506.

[77] W. Xu, K. Niu, Z. He, and W. Wu, "Resource allocation in multiuser ofdm distributed antenna systems," in *Proc. IEEE Veh. Techn. Conf. Fall (VTC)*, Baltimore, MD, USA, Setp. 2007, pp. 1797–1801.

[78] Y. Yao and M. Mehmet-Ali, "Optimal resource allocation in distributed broadband wireless communication systems," in *IEEE Annual Int. Symp. on Personal, Indoor and Mobile Radio Commun. - (PIMRC)*, Sydney, NSW, Australia, Sept. 2012, pp. 291–297.

[79] D. Boviz, C. S. Chen, and S. Yang, "Effective design of multi-user reception and fronthaul rate allocation in 5G cloud RAN," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 8, pp. 1825–1836, June 2017.

[80] K. N. Pappi, P. D. Diamantoulakis, and G. K. Karagiannidis, "Distributed uplink-NOMA for cloud radio access networks," *IEEE Commun. Lett.*, vol. 21, no. 10, pp. 2274–2277, Oct 2017.

[81] M. J. Youssef, J. Farah, C. A. Nour, and C. Douillard, "Resource allocation for mixed traffic types in distributed antenna systems using NOMA," in *Proc. IEEE Veh. Techn. Conf. Fall (VTC)*, Chicago, USA, Aug. 2018.

[82] M. J. Youssef, J. Farah, C. A. Nour, and C. Douillard, "Resource allocation in NOMA systems for centralized and distributed antennas with mixed traffic using matching theory," *IEEE Trans. Commun.*, vol. 68, no. 1, pp. 414–428, Jan. 2020.

[83] M. Katoozian, K. Navaie, and H. Yanikomeroglu, "Utility-based adaptive radio resource allocation in OFDM wireless networks with traffic prioritization," *IEEE Trans. Wireless Commun.*, vol. 8, no. 1, pp. 66–71, Jan 2009.

[84] W. Chung, C. Chang, and L. Wang, "An Intelligent Priority Resource Allocation Scheme for LTE-A Downlink Systems," *IEEE Wireless Commun. Lett.*, vol. 1, no. 3, pp. 241–244, June 2012.

[85] Y. Chung and C. Chang, "A balanced resource scheduling scheme with adaptive priority thresholds for OFDMA downlink systems," *IEEE Trans. Veh. Technol.*, vol. 61, no. 3, pp. 1276–1286, March 2012.

[86] R. Balakrishnan and B. Canberk, "Traffic-aware QoS provisioning and admission control in OFDMA hybrid small cells," *IEEE Trans. Veh. Technol.*, vol. 63, no. 2, pp. 802–810, Feb 2014.

[87] M. Pischella and J. Belfiore, "Resource allocation for QoS-aware OFDMA using distributed network coordination," *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 1766–1775, May 2009.

[88] Q. Liao, P. Baracca, D. Lopez-Perez, and L. G. Giordano, "Resource scheduling for mixed traffic types with scalable TTI in dynamic TDD systems," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2016, pp. 1–7.

[89] Y. Sun, D. W. K. Ng, Z. Ding, and R. Schober, "Optimal joint power and subcarrier allocation for full-duplex multicarrier non-orthogonal multiple access systems," *IEEE Trans. Commun.*, vol. 65, no. 3, pp. 1077–1091, Mar. 2017.

[90] S. Timotheou and I. Krikidis, "Fairness for non-orthogonal multiple access in 5G systems," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1647–1651, Oct. 2015.

[91] L. Lei, D. Yuan, and P. Värbrand, "On power minimization for non-orthogonal multiple access (NOMA)," *IEEE Commun. Lett.*, vol. 20, no. 12, pp. 2458–2461, Dec. 2016.

[92] J. Farah, E. Sfeir, C. A. Nour, and C. Douillard, "New resource allocation techniques for base station power reduction in orthogonal and non-orthogonal multiplexing systems," in *Proc. Int. Conf. on Commun. (ICC) Workshops*, May 2017, pp. 618–624.

[93] A. Brighente and S. Tomasin, "Power allocation for non-orthogonal millimeter wave systems with mixed traffic," *IEEE Trans. Wireless Commun.*, vol. 18, no. 1, pp. 432–443, Jan. 2019.

[94] Q. Vien, T. A. Le, C. V. Phan, and M. O. Agyeman, "An energy-efficient NOMA for small cells in heterogeneous CRAN under QoS constraints," in *23th European Wireless Conf.*, May 2017, pp. 1–6.

[95] J. Farah, A. Kilzi, C. A. Nour, and C. Douillard, "Power minimization in distributed antenna systems using non-orthogonal multiple access and mutual successive interference cancellation," *IEEE Trans. Veh. Technol.*, vol. 67, no. 12, pp. 11 873–11 885, Dec. 2018.

[96] E. Okamoto, "An improved proportional fair scheduling in downlink non-orthogonal multiple access system," *Proc. IEEE Veh. Techn. Conf. Fall (VTC)*, pp. 1–5, Sept 2015.

[97] A. Benjebbour, A. Li, Y. Saito, Y. Kishiyama, A. Harada, and T. Nakamura, "System-level performance of downlink NOMA for future LTE enhancements," *IEEE Globecom Workshops*, pp. 66–70, Dec 2013.

[98] M. S. A. E. Roth, "Two-sided matching, a study in game-theoretic modeling and analysis," *Econometric Society Monographs. Cambridge University Press,*, 1990.

[99] T. Quint, "The core of an m-sided assignment game," *Games and Economic Behavior*, vol. 3, 1990.

[100] Y. L. Lee, L. Wang, T. C. Chuah, and J. Loo, "Joint resource allocation and user association for heterogeneous cloud radio access networks," in *28th Int. Teletraffic Congress (ITC 28)*, vol. 01, Sept 2016, pp. 87–93.

[101] B. Di, L. Song, and Y. Li, "Sub-channel assignment, power allocation, and user scheduling for non-orthogonal multiple access networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7686–7698, Nov 2016.

[102] M. Hojeij, J. Farah, C. A. Nour, and C. Douillard, "Resource allocation in downlink non-orthogonal multiple access (NOMA) for future radio access," in *Proc. IEEE Veh. Techn. Conf. Spring (VTC)*, May 2015, pp. 1–6.

[103] 3GPP, "TR25-814 (V7.1.0), Physical Layer Aspects for Evolved Universal Terrestrial Radio Access (UTRA)," 2006.

[104] M. R. Palattella, M. Dohler, A. Grieco, G. Rizzo, J. Torsner, T. Engel, and L. Ladid, "Internet of things in the 5G era: Enablers, architecture, and business models," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 3, pp. 510–527, March 2016.

[105] H. Elsawy, E. Hossain, and D. I. Kim, "HetNets with cognitive small cells: user offloading and distributed channel access techniques," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 28–36, June 2013.

[106] S. Sesia, I. Toufik, and M. Baker, *LTE: The UMTS Long Term Evolution, From Theory to Practice*. Wiley, Feb. 2009.

[107] R. Razavi and H. Claussen, "Urban small cell deployments: Impact on the network energy consumption," in *Proc. IEEE Wireless Commun. and Networking Conf. (WCNC)*, Paris, France, Apr. 2012, pp. 47–52.

[108] J. Farah, A. Kilzi, C. Abdel Nour, and C. Douillard, "Power Minimization in Distributed Antenna Systems Using Non-Orthogonal Multiple Access and Mutual Successive Interference Cancellation," *IEEE Trans. on Veh. Technol.*, vol. 67, no. 12, pp. 11 873–11 885, Dec. 2018.

[109] M. Rahman and H. Yanikomeroglu, "Enhancing cell-edge performance: a downlink dynamic interference avoidance scheme with inter-cell coordination," *IEEE Trans. Wireless Commun.*, vol. 9, no. 4, pp. 1414–1425, Apr. 2010.

[110] M. J. Youssef, J. Farah, C. A. Nour, and C. Douillard, "Full-duplex and backhaul-constrained UAV-enabled networks using NOMA," *IEEE Trans. Veh. Technol., Early Access*, June 2020.

[111] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[112] M. J. Youssef, V. V. Veeravalli, J. Farah, and C. A. Nour, "Stochastic multi-player multi-armed bandits with multiple plays for uncoordinated spectrum access," *Accepted for publication in Proc. IEEE Annual Int. Symp. on Personal, Indoor, and Mobile Radio Commun. (PIMRC)*, Sept. 2020.

[113] M. J. Youssef, V. V. Veeravalli, J. Farah, C. A. Nour, and C. Douillard, "Resource allocation in NOMA-based self-organizing networks using stochastic multi-armed bandits," *Submitted to IEEE Trans. Veh. Technol.*

[114] A. Bin Sediq, R. Schoenen, H. Yanikomeroglu, and G. Senarath, "Optimized distributed inter-cell interference coordination (ICIC) scheme using projected subgradient and network flow optimization," *IEEE Trans. Commun.*, vol. 63, no. 1, pp. 107–124, Jan. 2015.

[115] J. Yun and K. G. Shin, "Distributed coordination of co-channel femtocells via inter-cell signaling with arbitrary delay," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 6, pp. 1127–1139, 2015.

[116] M. Yassin, Y. Dirani, M. Ibrahim, S. Lahoud, D. Mezher, and B. Cousin, "A novel dynamic inter-cell interference coordination technique for lte networks," in *Proc. IEEE Annual Int. Symp. on Personal, Indoor, and Mobile Radio Commun. (PIMRC)*, Hong Kong, China, Sept. 2015, pp. 1380–1385.

[117] S. K. Sharma and X. Wang, "Collaborative distributed Q-learning for RACH congestion minimization in cellular IoT networks," *IEEE Commun. Lett.*, vol. 23, no. 4, pp. 600–603, Apr. 2019.

[118] O. Naparstek and K. Cohen, "Deep multi-user reinforcement learning for distributed dynamic spectrum access," *IEEE Trans. Wireless Commun.*, vol. 18, no. 1, pp. 310–323, Jan. 2019.

[119] J. Zhang, X. Tao, H. Wu, N. Zhang, and X. Zhang, "Deep reinforcement learning for throughput improvement of uplink grant-free NOMA system," *IEEE Internet Things J., Early Access*, pp. 1–11, Feb. 2020.

[120] A. Feki and V. Capdevielle, "Autonomous resource allocation for dense lte networks: A multi armed bandit formulation," in *Proc. IEEE Annual Int. Symp. on Personal, Indoor, and Mobile Radio Commun. (PIMRC)*, Toronto, ON, Canada, Sept. 2011, pp. 66–70.

[121] P. Coucheney, K. Khawam, and J. Cohen, "Multi-armed bandit for distributed inter-cell interference coordination," in *Proc. Int. Conf. on Communications (ICC)*, London, UK, June 2015, pp. 3323–3328.

[122] A. Magesh and V. V. Veeravalli, "Multi-player multi-armed bandits with non-zero rewards on collisions for uncoordinated spectrum access," 2019. [Online]. Available: arXiv:1910.09089

[123] S. Ali, A. Ferdowsi, W. Saad, N. Rajatheva, and J. Haapola, "Sleeping multi-armed bandit learning for fast uplink grant allocation in machine type communications," *IEEE Trans. Commun., Early Access*, Apr. 2020.

[124] M. A. Adjif, O. Habachi, and J. Cances, "Joint channel selection and power control for NOMA: A multi-armed bandit approach," in *Proc. IEEE Wireless Commun. and Networking Conf. (WCNC)*, 2019, pp. 1–6.

[125] S. Maghsudi and S. Stańczak, "Joint channel selection and power control in infrastructureless wireless networks: A multiplayer multiarmed bandit framework," *IEEE Trans. Veh. Technol.*, vol. 64, no. 10, pp. 4565–4578, Oct. 2015.

[126] M. J. Youssef, C. A. Nour, J. Farah, and C. Douillard, "Backhaul-constrained resource allocation and 3d placement for UAV-enabled networks," in *Proc. IEEE Veh. Techn. Conf. Fall (VTC)*, Honolulu, Hi, USA, Sept. 2019, pp. 1–7.

[127] I. Bistritz and A. Leshem, "Distributed multi-player bandits - a game of thrones approach," in *32nd Proc. Int. Conf. on Neural Inf. Process. Syst.*, ser. NIPS'18, Montreal, Canada, 2018, pp. 7222–7232.

[128] J. R. Marden, H. P. Young, and L. Y. Pao, "Achieving Pareto optimality through distributed learning," in *SIAM J. Control Optim.*, no. 5, 2014, pp. 2753–2770.

[129] J. Rosenski, O. Shamir, and L. Szlak, "Multi-player bandits-a musical chairs approach," in *Int. Conf. on Mach. Learn.*, 2016, pp. 155–163.

[130] W. Hoeffding, "Probability inequalities for sums of bounded random variables," *J. Amer. Statist. Assoc.*, vol. 58, no. 301, pp. 13–30, 1963.

[131] Y. Zeng, R. Zhang, and T. J. Lim, "Wireless communications with unmanned aerial vehicles: opportunities and challenges," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 36–42, May 2016.

[132] E. Kalantari, M. Z. Shakir, H. Yanikomeroglu, and A. Yongacoglu, "Backhaul-aware robust 3D drone placement in 5G+ wireless networks," in *Proc. Int. Conf. on Commun. (ICC) Workshops*, Paris, France, May 2017, pp. 109–114.

[133] M. Heino, D. Korpi, T. Huusari, E. Antonio-Rodriguez, S. Venkatasubramanian, T. Riihonen, L. Anttila, C. Icheln, K. Haneda, R. Wichman, and M. Valkama, "Recent advances in antenna design and interference cancellation algorithms for in-band full duplex relays," *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 91–101, May 2015.

[134] U. Siddique, H. Tabassum, E. Hossain, and D. I. Kim, "Wireless backhauling of 5G small cells: challenges and solution approaches," *IEEE Wireless Commun.*, vol. 22, no. 5, pp. 22–31, Oct. 2015.

[135] Y. Choi and H. Shirani-Mehr, "Simultaneous transmission and reception: Algorithm, design and system level performance," *IEEE Trans. Wireless Commun.*, vol. 12, no. 12, pp. 5992–6010, Dec. 2013.

[136] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, June 2014.

[137] K. Yang, N. Yang, N. Ye, M. Jia, Z. Gao, and R. Fan, "Non-orthogonal multiple access: Achieving sustainable future radio access," *IEEE Commun. Mag.*, vol. 57, no. 2, pp. 116–121, Feb. 2019.

[138] L. Wang, B. Hu, and S. Chen, "Energy efficient placement of a drone base station for minimum required transmit power," *IEEE Wireless Commun. Lett., Early Access*, Feb. 2018.

[139] M. Alzenad, A. El-Keyi, and H. Yanikomeroglu, "3-D placement of an unmanned aerial vehicle base station for maximum coverage of users with different QoS requirements," *IEEE Wireless Commun. Lett.*, vol. 7, no. 1, pp. 38–41, Feb. 2018.

[140] L. Zhang and N. Ansari, "On the number and 3-D placement of in-band full-duplex enabled drone-mounted base-stations," *IEEE Wireless Commun. Lett.*, vol. 8, no. 1, pp. 221–224, Feb. 2019.

[141] T. M. Nguyen, W. Ajib, and C. Assi, "A novel cooperative NOMA for designing UAV-assisted wireless backhaul networks," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 11, pp. 2497–2507, Nov. 2018.

[142] Y. Liu, Z. Qin, Y. Cai, Y. Gao, G. Y. Li, and A. Nallanathan, "UAV communications based on non-orthogonal multiple access," *IEEE Wireless Commun.*, vol. 26, no. 1, pp. 52–57, Feb. 2019.

[143] A. A. Nasir, H. D. Tuan, T. Q. Duong, and H. V. Poor, "UAV-Enabled Communication Using NOMA," *IEEE Trans. Commun.*, vol. 67, no. 7, pp. 5126–5138, Jul. 2019.

[144] N. Zhao, X. Pang, Z. Li, Y. Chen, F. Li, Z. Ding, and M. Alouini, "Joint trajectory and precoding optimization for UAV-Assisted NOMA networks," *IEEE Trans. Commun.*, vol. 67, no. 5, pp. 3723–3735, 2019.

[145] M. Alzenad, A. El-Keyi, F. Lagum, and H. Yanikomeroglu, "3-D Placement of an Unmanned Aerial Vehicle Base Station (UAV-BS) for Energy-Efficient Maximal Coverage," *IEEE Wireless Commun. Lett.*, vol. 6, no. 4, pp. 434–437, Aug. 2017.

[146] J. Munkres, "Algorithms for the assignment and transportation problems," in *Journal of the Society for Industrial and Applied Mathematics*, vol. 5, 1957, pp. 32–38.

[147] J. Eckhoff, "Helly, Radon, and Carathéodory type theorems," in *Handbook of Convex Geometry*, North-Holland, 1993, pp. 389–448.

[148] P. M. Vaidya, "Speeding-up linear programming using fast matrix multiplication," in *Annual Symp. Found. Comput. Sci.*, Oct. 1989, pp. 332–337.

[149] J. Farah, J. Akiki, and E. P. Simon, "Energy-efficient techniques for combating the influence of reactive jamming using non-orthogonal multiple access and distributed antenna systems," *Wireless Telecom. Symp. (WTS)*, Apr. 2019.

[150] X. Wang, R. Chen, Y. Xu, and Q. Meng, "Low-Complexity Power Allocation in NOMA Systems With Imperfect SIC for Maximizing Weighted Sum-Rate," *IEEE Access*, vol. 7, pp. 94 238–94 253, July 2019.

DOCTORAT BRETAGNE LOIRE / MATHSTIC

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Titre : Nouvelles approches pour l'allocation des ressources dans les réseaux de communication futurs utilisant NOMA et aidés par les drones

Mots clés : Accès multiple non orthogonal, trafic hétérogène, accès non coordonné au spectre, réseaux auto-organisés, drones.

Résumé : Avec des prévisions de milliards d'appareils connectés, l'Internet des objets (IoT) est le moteur de l'évolution des réseaux de communication sans fil. Cette augmentation exponentielle du nombre d'appareils connectés s'accompagne d'une prolifération d'applications hétérogènes et de nouveaux cas d'utilisation sans fil très différents des services multimédias classiques. Par rapport aux systèmes de communication classiques, les systèmes de communication sans fil de prochaine génération devraient offrir des débits de données très élevés, une grande fiabilité, une faible latence, une amélioration de la qualité de service (QoS) perçue par les utilisateurs ainsi qu'une augmentation du nombre d'utilisateurs pris en charge. Pour répondre à ces exigences, certains des éléments clés que les futurs systèmes de communication doivent exploiter comprennent de nouvelles techniques d'accès au spectre telles que l'accès multiple non orthogonal (NOMA) et l'accès au spectre non coordonné, les réseaux auto-organisés (SON) et les réseaux de communication assistés par des drones (UAV). L'objectif principal de cette thèse est d'exploiter ces éléments clés pour fournir de nouvelles solutions d'allocation de ressources et de configuration de réseaux qui visent à optimiser l'utilisation des ressources radio disponibles dans ces systèmes de prochaine génération. Diverses configurations sont considérées, comprenant les systèmes sans fil avec trafic hétérogène, l'accès non coordonné au spectre dans les SON et les systèmes de communication assistés par des drones. Pour chaque configuration, une solution exploitant la technique NOMA est proposée. Les résultats ainsi obtenus montrent que les solutions proposées surpassent les techniques de l'état de l'art.

Title : New approaches for resource allocation in future communication networks using NOMA and UAVs

Keywords : Non-orthogonal multiple access, mixed traffic, uncoordinated spectrum access, self-organized networks, unmanned aerial vehicles.

Abstract : With a forecasted number of billions of connected devices, the Internet of Things (IoT) is driving the evolution of wireless communication networks. This exponential increase in the number of connected devices is accompanied by a proliferation of heterogeneous IoT applications, resulting in the emergence of new wireless use cases that greatly differ from conventional multimedia services. When compared to previous communication systems, next generation wireless communication systems are expected to provide very high data rates, high reliability, low latency, improvement in the quality of service (QoS) perceived by users and an increase in the number of supported users. To meet these requirements, some of the key elements future communication systems must leverage include novel spectrum access techniques such as non-orthogonal multiple access (NOMA) and uncoordinated spectrum access, self-organized networks (SON) and unmanned aerial vehicles (UAV)-aided communication networks. The main objective of this thesis is to exploit these key elements to provide novel resource allocation and net-work design solutions that aim at optimizing the use of available radio resources in next generation wireless communication networks. Different settings are considered, spanning wireless systems with heterogeneous mobile traffic requirements, uncoordinated spectrum access in SONs and UAV-aided communication systems. For each setting, a solution leveraging NOMA scheduling is proposed. The obtained results of the proposed solutions are promising, where these solutions are shown to outperform existing techniques from the literature.