New approaches for interference management in future generation networks for 5G and beyond using NOMA

Antoine Kilzi

UNIVERSITE 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
Antoine KILZI

Nouvelles approches de gestions des interférences pour les réseaux de communication 5G et futurs utilisant la NOMA

New approaches for interference management in future generation networks for 5G and beyond using NOMA (in English)

Thèse présentée et soutenue à Brest, le 10 février 2021
Unité de recherche : Lab-STICC
Thèse N° : 2021IMTA0238

Rapporteurs avant soutenance :

Didier Le Ruyet	Professeur des universités, Conservatoire National des Arts et Métiers (CNAM)
Georges Kaddoum	Professeur, École de Technologie Supérieure (ÉTS)

Composition du Jury :

Président :	Jean-Marie Gorce	Professeur des universités, INSA de Lyon
Rapporteurs :	Didier Le Ruyet	Professeur des universités, Conservatoire National des Arts et Métiers (CNAM)
	Georges Kaddoum	Professeur, École de Technologie Supérieure (ÉTS)
Directrices de thèse :	Catherine Douillard	Professeure, IMT Atlantique Brest
	Joumana Farah	Professeure, Université Libanaise Roumieh
Encadrant de thèse :	Charbel Abdel Nour	Maître de conférences, IMT Atlantique Brest

Sous le sceau de l'Université Bretagne Loire

IMT Atlantique

École Doctorale MATHSTIC

New Approaches for Interference Management in Future Generation Networks for 5G and Beyond using NOMA

Thèse de Doctorat

Spécialité : Télécommunications

Présentée par **Antoine Kilzi**

Département : Électronique
Laboratoire : Lab-STICC

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

Encadrant : **Charbel Abdel Nour**

Abstract

New Approaches for Interference Management in Future Generation Networks for 5G and Beyond using NOMA

Antoine KILZI

Electronics Department, IMT Atlantique

The ever-increasing demand for higher data rates, greater data volumes, more connected devices, and lower latency requirements have pushed the International Telecommunication Union (ITU) to redefine the requirements for International Mobile Telecommunications (IMT) for 2020 and beyond, its three main pillars being enhanced Mobile BroadBand (eMBB), Ultra-Reliable Low-Latency Communications (URLLC) and massive Machine-Type Communications (mMTC). Various techniques have been proposed by the academia and the industry in order to satisfy the aforementioned tight requirements and to address the challenges of future generation networks. Examples of such solutions are Non-Orthogonal Multiple Access (NOMA), small cells, Distributed Antenna Systems (DAS) and Cloud Radio Access Network (C-RAN) systems, Coordinated Multipoint (CoMP), Unmanned Aerial Vehicules (UAV), Device to Device (D2D) and Full Duplex (FD) communications. In all these techniques, the underlying problematic is that of interference management. In fact, the broader problem of interference punctuates the entire mobile communications field from the design of multiple access schemes (NOMA), to redefining network architectures (DAS) and coordination frameworks (CoMP), to the paradigm shifts represented by emerging solutions (UAVs and D2Ds). This thesis revolves around the interference management problem for various communication scenarios consisting of the combination of NOMA with one or several of the aforementioned techniques. The interference cancellation properties provided by the NOMA receivers are heavily investigated for the dual contexts of system power minimization and throughput maximization.

In Chapter 2, we tackle the problem of downlink power minimization in a single cell environment with DAS and using NOMA. First, an existing heuristic for the joint user-subcarrier-antenna and power assignment with user-rate requirements is extended from the Centralized Antenna Systems (CAS) context to DAS. Several complexity reduction techniques are proposed as well as novel Power Allocation (PA) schemes. Then, the inherent potentials of combining NOMA with DAS are investigated. The main contribution of Chapter 2 is the proposition of a new NOMA serving scheme termed mutual SIC, where paired users are able to mutually cancel their interferences thanks to the powering of multiplexed signals from different distributed antennas. The information theoretic conditions enabling mutual SIC are therefore studied, and as a result, the power minimization heuristics are reshaped to take advantage of the unveiled potentials of DAS with mutual SIC NOMA.

In Chapter 3, the proposed approaches of Chapter 2 are ported to the context of Hybrid Distributed Antenna Systems (HDAS) where a subset of the distributed antennas might face transmit power limitations. Under these practical considerations, meeting the user-rate requirements is no longer guaranteed, and particular attention is required to design Resource Allocation (RA) schemes satisfying the problem constraints. Therefore, optimal PA for HDAS is derived first and its specificities and divergence from DAS are

thoroughly investigated. This enables the proposal of a simple criterion to guarantee the existence of viable RA schemes. Afterwards, different strategies are proposed to tackle the power minimization problem in HDAS, capturing the characteristics of the HDAS scenario and enabling an efficient recourse to NOMA mutual SIC techniques which proved their efficiency in reducing the system power.

Chapter 4 is devoted to the generalization of the concepts of mutual SIC to arbitrary NOMA cluster sizes and for broader transmission scenarios, namely Joint Transmission (JT) in Coordinated Multipoint (CoMP) systems. A particular care is given to the decoding orders at the users level that are shown to greatly impact both the Power Multiplexing Constraint (PMC) and the rate constraints for the application of mutual SIC. Therefore, a new mathematical formalism for the derivation of the generalized mutual SIC conditions is proposed accounting for both the users decoding orders on one hand, and JT serving on the other hand. Then, the generalized mutual SIC concept is applied for two-user clusters (Dual Mutual SIC (DMSIC)) and three-user clusters (Triple Mutual SIC (TMSIC)). The proposed DMSIC and TMSIC solutions are shown to outperform existing CoMP NOMA schemes, achieving higher system Spectral Efficiency (SE), while providing greater fairness among served users.

In Chapter 5, the context of UAV-assisted networks is considered, where a UAV is dispatched to support a two-cell system with a saturated antenna. Inspired by the advantages of TMSIC from Chapter 4, the UAV positioning problem is formulated to maximize the chances of applying TMSIC. To that end, a novel mathematical framework is introduced to model the problem of UAV positioning with TMSIC feasibility in mind. The presented probabilistic framework captures the randomness of Air-to-Ground channel link characteristics and enables the formulation of TMSIC-seeking UAV placement problems in probabilistic terms. Several positioning strategies are proposed based on various network optimization metrics related to the proposed model. The trade-offs between the proposed strategies are highlighted and the use case scenarios for every positioning technique are discussed.

In the last chapter, the advantages of mutual SIC are studied for integration in the context of Device to Device (D2D) inband underlay communication systems, targeting maximum D2D sum-throughput, and using both Half Duplex and Full Duplex scenarios. The conditions allowing for mutual SIC between D2D devices and Cellular Users (CUs) are derived, and necessary and sufficient channel conditions taking into account transmit power limits and PMCs are identified. A geometrical approach is introduced to efficiently solve the optimal PA problem with a reduced complexity, enabling optimal global RA including D2D-CU channel pairing and PA. The implementation of mutual SIC is shown to provide great complementarity with D2D applications as the interference cancellation configurations of mutual SIC take advantage of the near-far effect to extend the realm of application scenarios of classical D2D following interference avoidance schemes.

Keywords: Non-Orthogonal Multiple Access, mutual Successive Interference Cancellation, Power Multiplexing Constraint, Power Allocation, Resource Allocation, Waterfilling, Spectral Efficiency, Distributed Antenna Systems, Coordinated Multipoint, Unmanned Aerial Vehicules, Device to Device, Full Duplex.

Contents

List of Figures

List of Tables

Acronyms

3GPP third generation parthership project.

A2G Air-to-Ground.

APAT Alternative Power Allocation Technique.

BBU Baseband Unit.

BS Base Station.

C-RAN Cloud Radio Access Network.

CAPEX CAPital EXpenditure.

CAS Centralized Antenna Systems.

CB Coordinated Beamforming.

CDMA Code Division Multiple Access.

CoMP Coordinated Multipoint.

CS Coordinated Scheduling.

CSI Channel State Information.

CSP Constraint Satisfaction Problem.

CU Cellular User.

D2D Device to Device.

DAS Distributed Antenna Systems.

DMSIC Dual Mutual SIC.

DPS Dynamic Point Selection.

DPS-CoMP Dynamic Point Selection CoMP.

EE Energy Efficiency.

eICIC enhanced ICI Coordination.

eMBB enhanced Mobile BroadBand.

FD Full Duplex.

FDD Frequency Division Duplex.

FDMA Frequency Division Multiple Access.

FTPA Fractional Transmit Power Allocation.

GUE Ground User Equipments.

HD Half Duplex.

HDAS Hybrid Distributed Antenna Systems.

HetNet Heterogeneous Network.

ICI Inter-Cell Intereference.

IoT Internet of Things.

JP Joint Processing.

JT Joint Transmission.

JT-CoMP Joint Transmission CoMP.

KKT Karush-Kuhn-Tucker.

LoS Line-of-Sight.

LPO Local Power Optimization.

LTE Long Term Evolution.

MA Multiple Access.

MBS Mobile Base Station.

MIMO-NOMA Multiple-Input Multiple-Output NOMA.

mMTC massive Machine-Type Communications.

NAIC Network-Assisted Interference Cancellation and Suppression.

NLoS non-LoS.

NOMA Non-Orthogonal Multiple Access.

OFDMA Orthogonal Frequency Division Multiple Access.

OMA Orthogonal Multiple Access.

OPAd Optimal Power Adjustment.

OPEX OPerating EXpenditure.

PA Power Allocation.

PD-NOMA Power Domain NOMA.

PMC Power Multiplexing Constraint.

QoS Quality of Service.

RA Resource Allocation.

RRH Remote Radio Head.

RSI Residual Self Interference.

RSS Received Signal Strength.

SA Subcarrier Assignment.

SE Spectral Efficiency.

SI Self Interference.

SIC Successive Interference Cancellation.

SINR Signal to Interference and Noise Ratio.

SNR Signal-to-Noise Ratio.

TDD Time Division Duplex.

TDMA Time Division Multiple Access.

TMSIC Triple Mutual SIC.

TP Transmission Point.

TPs Transmission Points.

UAV Unmanned Aerial Vehicules.

UE User Equipment.

UL UpLink.

UPT UAV Positioning Technique.

URLLC Ultra-Reliable Low-Latency Communications.

Résumé étendu de la thèse en français

Introduction générale

De nos jours, la place qu'occupent les communications mobiles dans la société moderne va bien au-delà de la communauté des experts, puisqu'elles ont contribué à façonner la société actuelle de manière inédite. L'interaction entre l'offre et la demande suit la loi de Say, où l'offre génère des applications par lesquelles une plus grande demande surgit, nécessitant ainsi une offre supplémentaire [1]. Les progrès techniques, qui offrent une plus grande facilité d'utilisation et des services plus étendus, ont pénétré dans la vie quotidienne des consommateurs et ont considérablement modifié les activités humaines. Cet effet, associé à la concurrence féroce pour obtenir des parts de marché plus importantes, pousse les opérateurs de réseaux mobiles (MNO) à faire davantage de battage publicitaire et, par conséquent, conditionne la société à en attendre toujours plus. Le côté avide étant déclenché, chaque nouvelle percée technologique (par exemple, la naissance de l'iPhone en 2007) génère une nouvelle gamme d'applications, qui s'insère dans les habitudes des sociétés et se transforme en besoins réels, justifiant ainsi une demande supplémentaire à laquelle l'offre doit faire face. Cette rétroaction auto-renforcée a entraîné une demande sans cesse croissante de débits de données plus élevés, de volumes de données plus importants, de dispositifs plus connectés, d'exigences de latence plus faibles pour des plans de données moins chers [2]. Parallèlement, l'émergence de l'Internet des objets (IoT), des communications de machine à machine et de véhicule à véhicule et d'autres technologies complexifie considérablement les profils de trafic, imposant aux MNO la contrainte d'une plus grande flexibilité pour répondre aux diverses demandes des réseaux de générations actuelle et futures. L'union internationale des télécommunications (ITU) a défini les besoins en matière de télécommunications mobiles internationales (IMT) pour 2020 et au-delà, les trois principaux piliers étant les communications mobiles large bande améliorées (eMBB: diffusion de vidéos 4K, réalité virtuelle et augmentée, etc.), les communications ultra-fiables à faible latence (URLLC: par exemple, chirurgie à distance, sécurité des transports) et les communications massives de type machine (mMTC : par exemple, compteurs intelligents, détection de réseau).

Diverses techniques ont été proposées par le monde universitaire et l'industrie afin de satisfaire aux exigences strictes susmentionnées et de relever les défis des réseaux de la future génération. Parmi ces solutions, on peut citer l'accès multiple non orthogonal (NOMA), les petites cellules, les systèmes d'antennes distribuées (DAS) et les réseaux d'accès radio de type *Cloud-RAN* (C-RAN), les communications multipoints coordonnées (CoMP), les véhicules aériens sans pilote (UAV), les communications de dispositif à dis-

positif (D2D) et les communications en duplex intégral (FD). Pour toutes ces techniques, la problématique sous-jacente est celle de la gestion des interférences. En fait, le problème plus large de l'interférence ponctue tout le domaine des communications mobiles, de la conception des schémas d'accès multiples (ex. NOMA), à la redéfinition des architectures de réseau (ex. DAS) et des cadres de coordination (ex. CoMP), jusqu'aux changements de paradigme représentés par les solutions émergentes (ex. UAVs et D2Ds).

Cette thèse s'articule autour du problème de la gestion des interférences pour divers scénarios de communication impliquant la combinaison de NOMA avec une ou plusieurs des techniques mentionnées ci-dessus. Dans un premier temps, les propriétés d'annulation des interférences des récepteurs NOMA sont étudiées dans le contexte de systèmes d'antennes distribuées, ce qui donne lieu à un moyen d'annulation totale de l'interférence que nous baptisons *mutual SIC*. Il s'en suit des applications de minimisation de la puissance dans les cellules qui sont étudiées dans les chapitres 2 et 3, ou de maximisation des débits dans les systèmes CoMP, les systèmes assistés par des UAV, ainsi que les systèmes moyennant de la communication D2D qui sont étudiés dans une seconde partie de la thèse.

Chapitre 1 : Contexte général

Nous présentons dans ce chapitre une vue d'ensemble des principaux schémas d'accès multiple, des architectures de réseau et des techniques de communication qui sont abordés dans la thèse. Nous discutons d'abord de la façon dont le nombre croissant de dispositifs connectés pousse à l'adoption de la technique NOMA, puis nous présentons les principes de cette technique appliquée dans le domaine de puissance, en mettant en évidence ses avantages et en soulignant ses conditions d'application théoriques et pratiques. Ensuite, nous expliquons les changements de paradigme lors du passage des systèmes d'antennes centralisées aux architectures distribuées densifiées telles que les DAS et les réseaux C-RAN. Les techniques spécifiques aux DAS sont présentées du point de vue de l'allocation des ressources. La densification des réseaux étant limitée par les interférences intercellulaires qu'elle génère, les principes de la CoMP, technique actuellement la plus avancée pour la coordination des interférences intercellulaires, sont présentés par la suite. Enfin, le contexte des communications D2D est décrit. Ses capacités à répondre à la demande diversifiée et à décharger le trafic de données du cœur du réseau vers ses dispositifs frontaux sont expliquées. En outre, la relation symbiotique que le D2D entretient avec les communications en duplex intégral (FD) est détaillée.

Principe de la technique 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 un même bloc de ressources temps-fréquence. Du côté de l'émetteur, les signaux des différents utilisateurs se voient attribuer des niveaux de puissance différents, et le codage par superposition est utilisé pour transmettre les signaux combinés des utilisateurs. Nous désignons par x_1 et x_2 les signaux multiplexés des utilisateurs UEs 1 et 2, avec pour puissances respectives P_1 et P_2, et des gains de canal h_1 and h_2 avec $|h_1| > |h_2|$. Dans le contexte du NOMA, UE 1 est qualifié d'utilisateur *fort* alors que UE 2 est l'utilisateur *faible*. Le signal superposé et transmis par la station de base est donné par $x = x_1 + x_2$, et les signaux reçus au niveau de UE 1 et UE 2 sont donnés par: $y_1 = xh_1 + n_1$ et $y_2 = xh_2 + n_2$, où n_i représente le bruit

blanc gaussien perçu au niveau de UE i, ayant pour variance σ^2. Au niveau de UE 1, un récepteur à annulation successive d'interférence (SIC) est appliqué pour extraire x_1 de y_1. Il procède par une détection, une démodulation et un décodage du signal x_2 pour ensuite le ré-encoder et le soustraire au signal reçu comme montré dans la Fig.1.

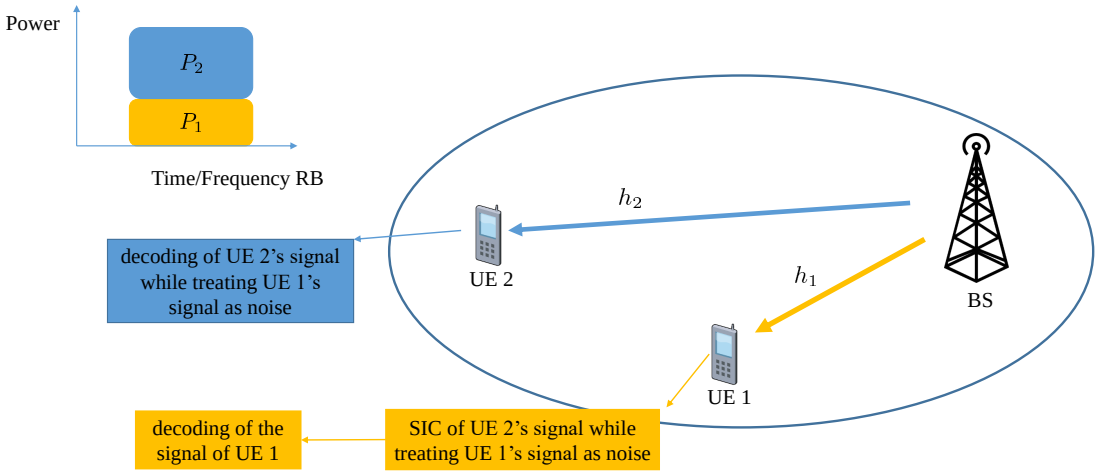


Figure 1 – Représentation d'un système NOMA à deux utilisateurs où UE 1 est l'utilisateur fort implémentant un récepteur SIC.

Par conséquent, x_1 peut être décodé sans interférence à un débit théorique donné par la capacité de Shannon:

$$R_1 = \log_2\left(1 + \frac{P_1|h_1|^2}{\sigma^2}\right).$$

Au niveau de l'utilisateur faible, l'interférence de UE 1 s'ajoute au bruit blanc, et le débit atteignable dans ce cas est donné par:

$$R_2 = \log_2\left(1 + \frac{P_2|h_2|^2}{P_1|h_2|^2 + \sigma^2}\right).$$

Densification de réseau et système d'antennes distribuées

L'idée de base de la densification du réseau est de rapprocher les nœuds d'accès au réseau des utilisateurs finaux en répartissant plusieurs points de transmission (TP) dans la cellule au lieu de les regrouper au même endroit comme pour un système centralisé (CAS). Cela permet d'améliorer la couverture de la cellule et d'accroître sa capacité en améliorant la qualité de la liaison grâce à la réduction de l'affaiblissement sur le trajet et à la diversité spatiale supplémentaire favorisant la communication en ligne de visée (LoS). En outre, la densification du réseau augmente la réutilisation par unité de surface du spectre disponible, ce qui améliore considérablement la capacité du réseau.

Sur la densification distribuée ou centralisée

La densification des réseaux peut être classée en densification distribuée et densification centralisée. La densification distribuée correspond au déploiement géographique de petites cellules dans des zones où un trafic important est généré. Les petites cellules, les picocellules et les femtocellules sont des BS entièrement fonctionnelles, capables de remplir

toutes les fonctions des macrocellules (bande de base et traitement radio) mais avec une puissance moindre et des zones de couverture plus petites. Chaque petite cellule ayant sa propre connexion de liaison de retour, la coordination entre elles n'est pas simple et des protocoles de gestion des interférences distribués sont nécessaires [3, 4].

D'autre part, lorsque l'unité de traitement en bande de base d'une BS est découplée de ses unités radio, il est possible de réaliser une densification centralisée du réseau dans un système DAS en déployant des têtes radio distantes (RRHs) dans toute la cellule, tout en les connectant à une unité de traitement centrale appelée Baseband Unit (BBU) par des fibres optiques à haut débit et à faible latence. Les RRH sont responsables de la conversion numérique-analogique, de la conversion analogique-numérique, de l'amplification de puissance et du filtrage, tandis que la BBU se charge de tout le traitement en bande de base et des procédures de niveau supérieur telles que la programmation des utilisateurs, le contrôle d'accès au support et la gestion des ressources radio (RRM). Cette architecture en étoile permet une coordination complète entre les RRH. Les différences entre les DAS et les petites cellules sont illustrées en Fig. 2.

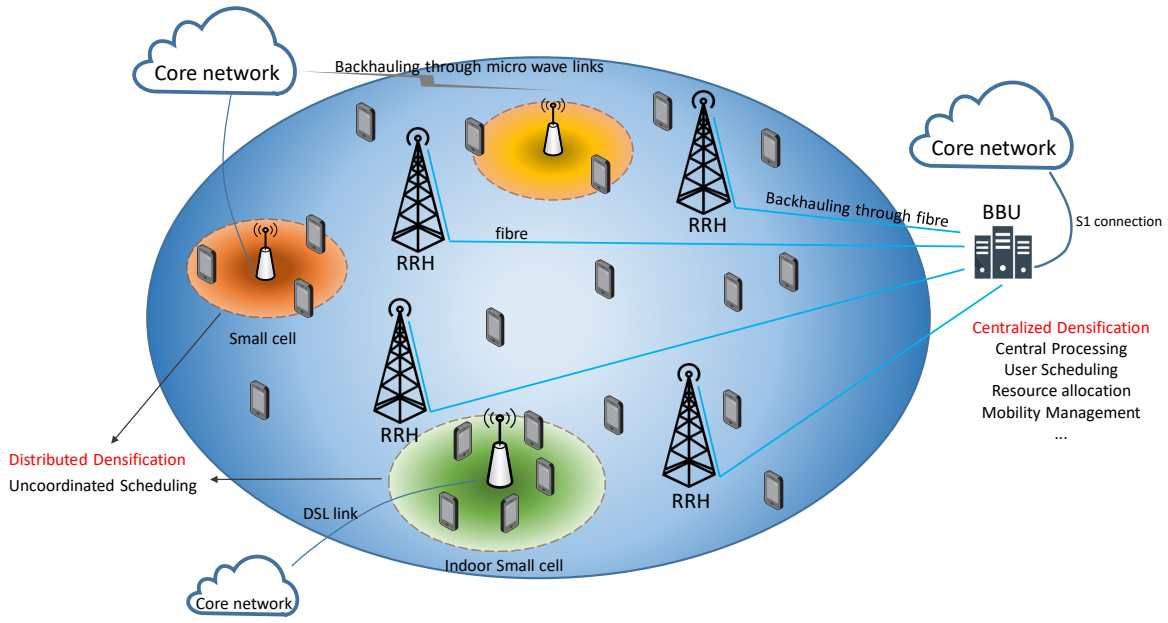


Figure 2 – Schéma d'un réseau hétérogène densifié composé de petites cellules autonomes avec connexion de raccordement individuelle, et de RRH distribuées contrôlées par une seule entité BBU.

Dans toute la littérature, une distinction a été faite entre le déploiement d'antennes pour améliorer la couverture et l'augmentation de la capacité. Les systèmes de petites cellules sont généralement considérés comme des amplificateurs de capacité, capables de fournir des gains de capacité importants pour de petites régions à forte activité de réseau en réutilisant la fréquence de la cellule. Dans ce scénario, le fait de disposer d'une petite zone de couverture permet de créer une zone de haute capacité localisée qui ne crée pas d'interférences excessives sur les sites voisins. D'autre part, le renforcement de la couverture était l'objectif principal des premiers déploiements de DAS : les signaux étaient diffusés simultanément sur toutes les antennes pour couvrir la zone de couverture. Bien que

raisonnable du point de vue de la couverture pure, cette approche présente l'inconvénient de provoquer d'importantes interférences hors cellule par rapport aux petites cellules et aux CAS. En outre, des études telles que [5, 6] ont montré que les utilisateurs peuvent être servis de manière plus efficace grâce à la diversité de sélection, où l'un des RRH est sélectionné pour transmettre le signal de l'utilisateur. Il est démontré que cette approche offre une plus grande capacité et un service aux utilisateurs plus efficace en termes de puissance. En outre, grâce à la densification centralisée des DAS, la programmation BBU peut fonctionner de telle sorte que certains RRH réutilisent tout le spectre tandis que d'autres RRH partagent dynamiquement la fréquence de la cellule. Pour toutes ces raisons, les potentialités des DAS nous semblent plus attrayantes que celles des petites cellules, notamment du point de vue de l'allocation de ressources. C'est pourquoi, dans cette thèse, une grande importance a été accordée à la configuration DAS avec diversité de sélection dans les schémas d'allocation de ressources (RA) proposés.

Transmission en multipoints coordonnés CoMP

La limite fondamentale de la densification des réseaux réside dans l'interférence croissante causée par la diminution de la distance inter-sites. Il a été démontré dans [7] que lorsque la densité de petites cellules augmente au-delà d'un certain seuil, le rapport signal-sur-interférence-plus-bruit SINR diminue car les signaux interférants passent d'une propagation sans visibilité (NLoS) à une propagation LoS, ce qui dégrade les performances du réseau. Pour atténuer les interférences intercellulaires (ICI), 3GPP a proposé dans la version 9 [8], puis a adopté dans la version 11 [9], la technique CoMP pour améliorer les performances des utilisateurs sujets aux interférences et les performances globales du réseau. Le principe consiste à appliquer une coordination entre les cellules adjacentes, soit pour atténuer les interférences au bord de la cellule sans restreindre l'utilisation des ressources du réseau, soit pour tirer intelligemment parti des interférences.

Plusieurs classifications des techniques de CoMP existent dans la littérature. Dans cette thèse nous traiterons des techniques CoMP de selection dynamique du point de transmission (DPS) et de la technique de transmission conjointe par points multiples de transmission (JT).

Sélection dynamique du point de transmission DPS

Dans le DPS, les données relatives à un UE sont transmises par un seul nœud d'émission pour une ressource temps/fréquence donnée. Cela requiert, en plus de l'échange d'information relative à l'état du canal (CSI), la disponibilité des données d'utilisateurs pour tous les émetteurs coopérants, ce qui permet au point sélectionné de changer dynamiquement d'un intervalle de temps de transmission à un autre. Par conséquent, le RRH présentant la perte de chemin la plus faible pour l'UE est toujours sélectionné.

Transmission conjointe par points multiples JT-CoMP

Avec la transmission JT-CoMP, des TP coopérant transmettent simultanément le signal du même utilisateur sur la même ressource temps-fréquence. Le traitement conjoint des données permet leur précodage sur les multiples nœuds d'émission afin qu'elles soient combinées de manière cohérente au niveau de chaque UE. La JT-CoMP est la technique de coordination la plus prometteuse, mais aussi la plus difficile à mettre en œuvre au vu

des contraintes strictes de synchronisation qu'elle impose. Les techniques CoMP DPS et JT sont présentées en Fig. 3.

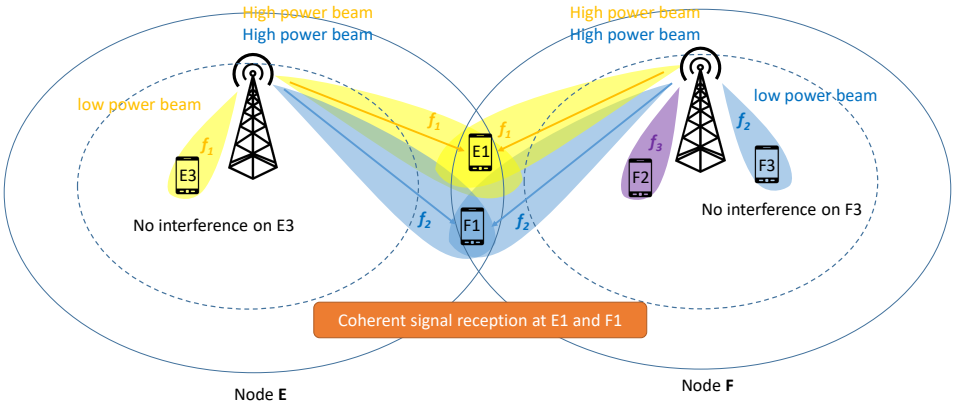


Figure 3 – Transmission en JT aux utilisateurs E1 et F1, et transmission en DPS pour F2.

Communications *device-to-device* (D2D)

L'idée de base des communications D2D est de permettre la communication directe entre des terminaux proches au lieu de faire transiter l'information par les stations de base et le cœur du réseau. La communication D2D décharge le réseau du trafic de liaison montante et descendante, ce qui libère de la capacité et des ressources énergétiques pour servir d'autres utilisateurs. En outre, grâce à la proximité des terminaux, un canal D2D efficace peut être établi, ce qui permet d'obtenir des débits de données élevés avec des puissances d'émission minimales et une latence très faible. Cela améliore l'efficacité énergétique du système et limite la zone d'interférences, permettant une meilleure réutilisation du spectre [10], [11]. De nombreux services peuvent bénéficier du D2D, comme l'illustre la Fig. 4. On peut notamment citer les applications de partage de contenu pour l'échange de vidéos et de photos, les jeux en réseau, les services de diffusion en continu avec mise en cache, les relais d'extension de couverture, les communications de véhicule à véhicule (V2V) nécessitant des contraintes de latence strictes, etc.

Concernant les communications D2D, la classification suivante peut être faite [12] :

- **Communications D2D en *outband*:** la communication D2D prend place sur une bande non licenciée du spectre sans affecter le réseau cellulaire.
- **Communication D2D en *inband*:** le canal D2D est alloué sur le spectre du réseau cellulaire. La communication D2D peut être soit en *overlay* ou en *underlay*.
  - **Overlay:** Des canaux dédiés du spectre cellulaire sont alloués aux communications D2D, ce qui empêche l'interférence co-canal entre système D2D et réseau cellulaire.

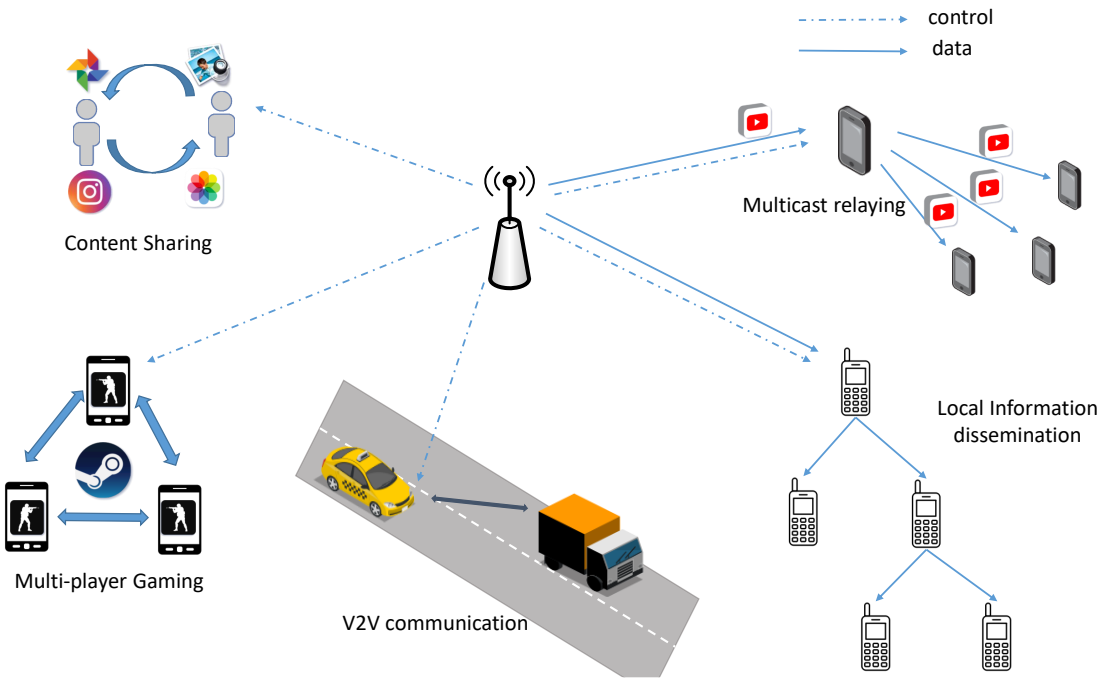


Figure 4 – Un aperçu des applications possibles du D2D.

– **Underlay:** Dans ce cas, le spectre du réseau cellulaire est réutilisé par les équipements D2D et le défi réside en une gestion efficace des interférences entre les réseaux D2D et cellulaires.

En raison de la nature stochastique de la bande sans licence et des difficultés à coordonner la communication sur deux bandes différentes (puisque la communication outband nécessite une deuxième interface radio et utilise d'autres technologies sans fil telles que WiFi Direct [13]), la transmission inband a suscité beaucoup d'intérêt au sein de la communauté des chercheurs [14, 15]. De plus, en raison de l'augmentation prévue du nombre de dispositifs connectés, dédier des bandes cellulaires au D2D ne sera pas une solution viable, c'est pourquoi la plupart des recherches se concentrent sur le D2D en bande sous-jacente ou underlay [16–19].

Duplex intégral FD

Une technologie très prometteuse à appliquer en conjonction avec le D2D est la communication FD. Le FD permet à un même UE de transmettre et de recevoir des informations en même temps et en utilisant la même fréquence [20]. Les systèmes de communication précédents impliquaient soit une transmission et une réception simultanées, mais en utilisant des fréquences distinctes dans le cas du FDD (Frequency Division Duplex), soit une transmission et une réception dans le même canal, mais en utilisant des intervalles de temps orthogonaux pour le TDD (Time Division Duplex), communément appelé communication half-duplex (HD). Les gains obtenus par le FD peuvent aller jusqu'à un doublement virtuel de l'efficacité spectrale (ES) par rapport aux systèmes TDD et FDD. En contrepartie, une auto-interférence (SI) est observée en raison du retour en boucle du signal transmis dans le récepteur, ce qui limite son intérêt par rapport au HD. Le défi de la

conception d'un équipement FD consiste à annuler la SI de sorte que l'auto-interférence résiduelle (RSI) soit comparable au bruit de fond. Aujourd'hui, les améliorations apportées à l'architecture des antennes et aux circuits des émetteurs-récepteurs permettent de réduire considérablement la RSI [21–23], ce qui plaide en faveur de l'utilisation du FD dans les futures normes de communication.

La plupart des analyses de haut niveau sur les gains de capacité du FD [24–26] modélisent la RSI comme une variable aléatoire gaussienne complexe de moyenne nulle et de variance ηP_{tx}, où η est la capacité d'annulation de la SI du dispositif FD et P_{tx} sa puissance de transmission. Ainsi, la puissance de la RSI, P_{RSI}, est donnée par:

$$P_{RSI} = \eta P_{tx}. \tag{1}$$

Le facteur d'annulation η peut varier entre 0 et 1, avec $\eta = 0$ correspondant à une annulation parfaite de la SI, et $\eta = 1$ se référant au cas où aucune annulation n'est appliquée. Dans la thèse, les valeurs effectives de η varient entre -80 dB et -130 dB. Par conséquent, la RSI est directement liée à la puissance du signal transmis, ce qui rend le FD plus adapté aux applications à faible puissance comme dans les réseaux D2D. L'intérêt croissant pour la combinaison de la communication FD avec la technologie D2D a donné naissance à de nouvelles applications et de nouveaux scénarios D2D, comme le montre la Fig. 5.

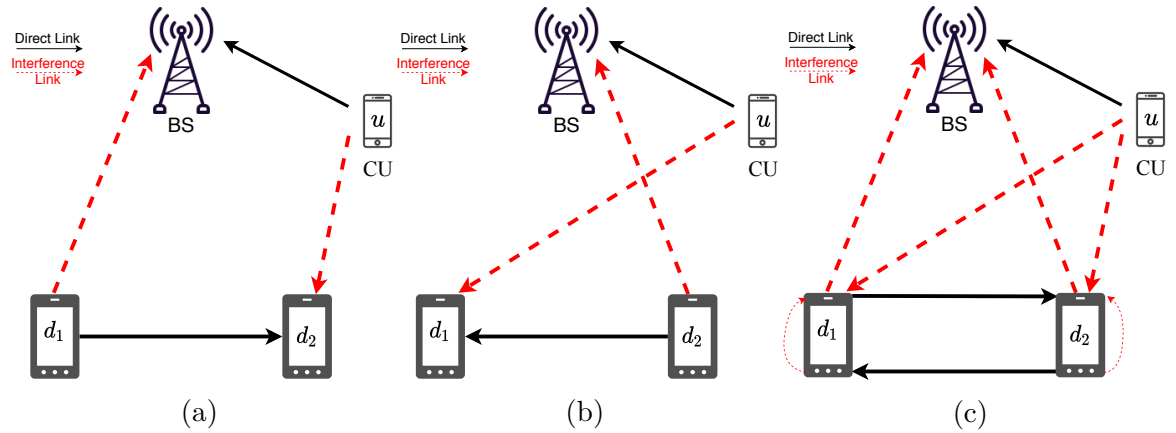


Figure 5 – Transmission D2D sous-jacente au réseau cellulaire (a) transmission en HD, premier demi-cycle, d_1 transmet à d_2. (b) transmission en HD, second demi-cycle, d_2 transmet à d_1. (c) Transmission en FD, d_1 et d_2 transmettent en même temps.

Dans cette thèse, nous nous intéresserons à la topologie dite bidirectionnelle FD-D2D présentée dans la figure 5c. Dans ce cas d'utilisation, un système D2D est sous-jacent au réseau cellulaire. Les dispositifs D2D cherchent à échanger des informations, d'où la topologie bidirectionnelle, tout en bénéficiant de la technologie FD au niveau des deux dispositifs d_1 et d_2. Dans ce cas, les dispositifs D2D vont provoquer des interférences sur le signal de l'utilisateur cellulaire au niveau de la station de base, et le signal de l'utilisateur cellulaire va interférer avec les deux dispositifs. La version HD de cette topologie est également présentée en Fig. 5 : dans la Fig. 5a, d_1 transmet des informations à d_2 pendant que d_2 reçoit, et dans la Fig. 5b, d_2 transmet des informations à d_1 pendant que d_1 reçoit.

Chapitre 2 : NOMA *Mutual SIC* pour la minimisation de puissance dans les systèmes d'antennes distribuées

Dans ce chapitre, nous considérons un système DAS pour servir en voie descendante les utilisateurs d'une cellule ayant chacun une contrainte de débit cible à respecter. Le but est de déterminer l'allocation de ressources en termes d'antennes de liaison, de sous-bandes allouées par utilisateur, et de puissance allouée par sous-bande minimisant la puissance totale du système. Nous proposons une résolution en deux étapes distinctes: la première moyennant une communication orthogonale (OMA), et la seconde utilisant l'accès non-orthogonal au spectre par le NOMA. Nous montrons comment la combinaison du NOMA et du DAS donne lieu au concept de *mutual SIC* où les deux utilisateurs appairés parviennent à annuler leurs interférences.

Algorithmes proposés

En OMA, le partage du spectre entre les utilisateurs est réalisé de manière itérative, conjointement à une allocation de puissance optimale basée sur le concept de *waterfilling*. Après une phase d'initialisation où chaque utilisateur est servi par sa meilleure sous-porteuse, l'algorithme opère de la manière suivante:

- L'utilisateur consommant le plus de puissance est sélectionné,
- Le couple (antenne, sous-porteuse) présentant le meilleur gain de canal est alloué à l'utilisateur sélectionné,
- La puissance totale consommée par l'utilisateur est mise à jour ainsi que le classement des utilisateurs consommant le plus de puissance.

Ces étapes sont répétées jusqu'à l'allocation de tout le spectre aux utilisateurs. Il s'ensuit l'étape itérative d'appariement NOMA des utilisateurs :

- L'utilisateur consommant le plus de puissance est sélectionné pour être appairé comme second utilisateur en NOMA.
- La sous-porteuse conduisant à la plus grande réduction de puissance de l'utilisateur en question est sélectionnée. La puissance totale de l'utilisateur est minimisée par une optimisation locale de puissance (LPO).
- La puissance totale consommée par l'utilisateur est mise à jour par un waterfilling appliqué aux sous-porteuses qui lui sont exclusivement allouées.

Mutual SIC NOMA

Lorsque les signaux des utilisateurs appairés en NOMA sont émis par des antennes différentes, il devient possible d'appliquer le mutual SIC où l'interférence entre utilisateurs est éliminée au niveau des deux utilisateurs simultanément. Pour ce faire, les conditions de canal à vérifier et les conditions de multiplexage de puissance (PMC) à respecter sont données par:

$$h_{1,2}h_{2,1} > h_{1,1}h_{2,2} \tag{2}$$

$$\frac{h_{2,2}}{h_{2,1}} < \frac{P_1}{P_2} < \frac{h_{1,2}}{h_{1,1}} \tag{3}$$

où $h_{i,j}$ représente le gain de canal de l'utilisateur i avec l'antenne j. Les débits alors atteignables par les deux utilisateurs en bit par seconde par Hz sont donnés par:

$$R_1 = \log_2(1 + P_1 h_{1,1}^2/\sigma^2), \quad R_2 = \log_2(1 + P_2 h_{2,2}^2/\sigma^2)$$

Grâce au mutual SIC, l'allocation de puissance optimale consiste en un simple waterfilling mais en veillant à respecter les conditions de PMC de (3). Si ce n'est pas le cas, un ajustement de puissance est requis ; ce dernier pouvant porter sur P_2 uniquement dans le cas de l'ajustement direct de puissance (DPA), ou sur P_1 et P_2 conjointement dans les cas de l'ajustement optimal et semi-optimal (OPAd) et (SOPAd).

Exemples de résultats

La performance des techniques proposées est évaluée par le biais de simulations numériques dans les contextes DAS et CAS. Pour le DAS, les méthodes basées sur le NOMA classique (une même antenne d'émission) sont désignées par "SRRH", et celles employant le mutual SIC sont désignées par "MutSIC". Les variantes du SRRH diffèrent par la méthode d'allocation de puissance choisie : le "FTPA" [27, 28], le "LPO" que nous avons proposé, et le "OPA" qui opère une optimisation de puissance globale mais est bien plus complexe en termes de temps de calcul [29].

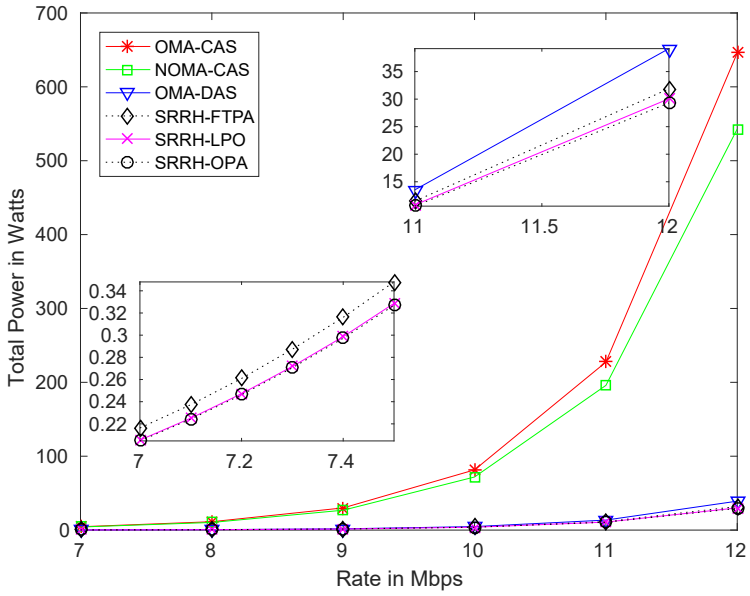


Figure 6 – Puissance consommée en fonction du débit requis par utilisateur $R_{k,req}$ pour les contextes CAS et DAS, en signalisation OMA et NOMA-SRRH.

Dans la Fig. 6, la puissance totale des différentes techniques est représentée en fonction du débit cible des utilisateurs pour un total de 15 utilisateurs dans la cellule. Les résultats montrent que la configuration DAS surpasse largement le CAS avec une réduction de puissance d'un facteur 16 environ. À un débit cible de 12 Mbps, la puissance totale requise en utilisant SRRH-FTPA, SRRH-LPO et SRRH-OPA est respectivement inférieure de 17,6 %, 24,5 % et 26,1 % à celle de la configuration OMA-DAS. Cela montre un avantage net du NOMA classique sur l'OMA même dans le contexte DAS. En

outre, l'application de la LPO permet une réduction de la puissance de 7,7 % par rapport à la FTPA, avec une charge de calcul similaire. La pénalité de performance de la LPO par rapport à l'allocation de puissance optimale n'est que de 2 % à 12 Mbps, mais avec une complexité considérablement réduite. La Fig. 7 compare dans le contexte

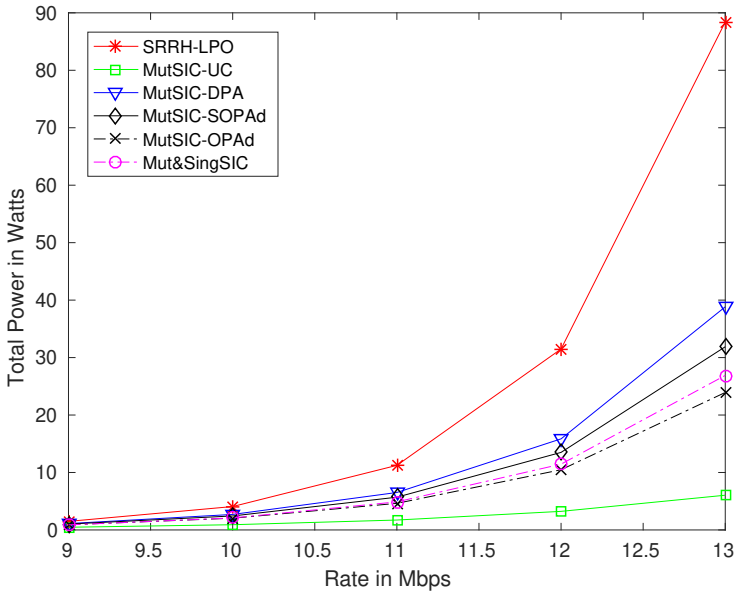


Figure 7 – Puissance totale en fonction de $R_{k,req}$ pour les schémas NOMA-DAS proposés.

DAS le NOMA classique avec une seule antenne au NOMA mutual SIC. Les méthodes MutSIC-DPA, MutSIC-SOPAd et MutSIC-OPAd dépassent de loin les performances de la SSRH-LPO avec des réductions de puissance respectives de 56.1 %, 63.9 % et 72.9 % pour un débit $R_{k,req}$ = 13 Mbps. Les gains significatifs du OPAd par rapport au SOPAd, et du SOPAd par rapport au DPA sont obtenus au prix d'une augmentation de la complexité de l'allocation de puissance, d'où un compromis entre performance et complexité. Finalement, en combinant les techniques SRRH pour certaines sous-porteuses au NOMA mutual SIC pour d'autres sous-porteuses, il devient alors possible de réduire encore la puissance. C'est le cas du Mut&SingSIC qui combine le SOPAd et le LPO pour aboutir à un gain supplémentaire de 15.6 % par rapport au MutSIC-SOPAd.

Chapitre 3 : NOMA mutual SIC pour la minimisation de puissance dans les systèmes d'antennes hybrides distribuées

Dans ce chapitre, nous prolongeons l'étude du problème de minimisation de puissance d'une cellule DAS en voie descendante au cas pratique où certaines antennes sont contraintes en puissance de transmission (systèmes DAS hybrides ou HDAS). En présence de contraintes de puissance limite des antennes, il n'est plus possible de satisfaire les demandes de débits cibles pour n'importe quelle association d'utilisateurs et d'antennes par la simple biais de l'allocation de puissance. Nous proposons deux méthodes distinctes et

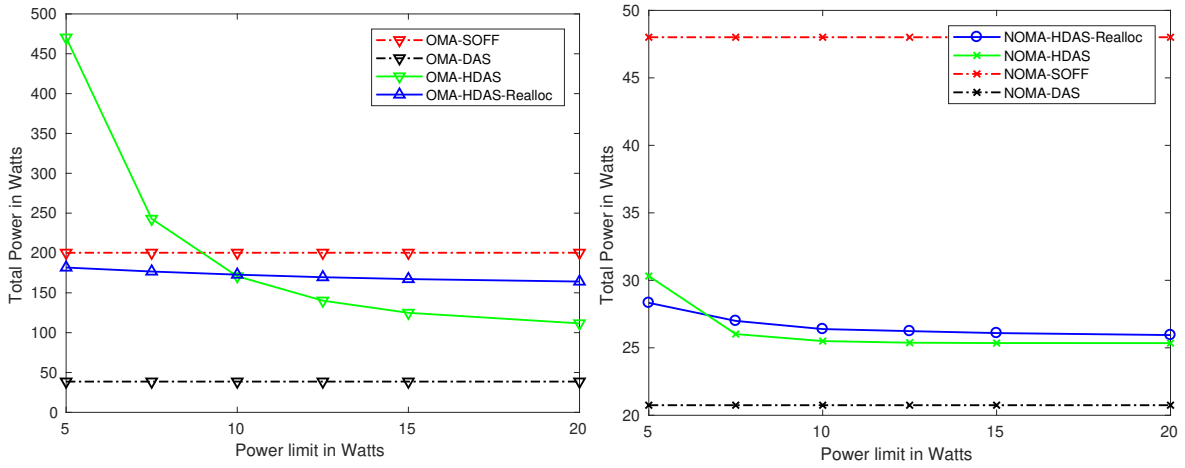
complémentaires pour garantir la faisabilité du problème et le résoudre. Les principales contributions de ce chapitre peuvent être résumées comme suit :

- Nous fournissons une analyse approfondie de l'allocation de puissance optimale en OMA pour le HDAS en y soulignant les propriétés uniques qui le différencient du cas classique du DAS,
- Nous déterminons un ensemble de conditions suffisantes pour que l'allocation de canal choisie et l'association des utilisateurs aux antennes sélectionnée garantissent l'existence d'une solution qui satisfasse les contraintes de débits cibles des utilisateurs ainsi que les contraintes de puissances de transmission maximales des antennes,
- Nous proposons deux approches différentes pour une allocation conjointe de puissance et de sous-porteuses pour les cas OMA et NOMA. L'une des approches est plus efficace pour les conditions de simulation difficiles (en termes de débits cibles et de puissances limites), alors que l'autre est plus performante pour les conditions moins contraignantes de débits cibles et de puissances de transmission.

Exemples de résultats

Les techniques "OMA-HDAS" et "NOMA-HDAS" reposent sur une modification de la phase d'initialisation des algorithmes antérieurs de réduction de puissance (tant pour l'OMA que le NOMA), de sorte que l'on s'assure que chaque utilisateur est servi par une antenne non contrainte et sur une de ses sous-porteuses tout au moins. Ainsi, la satisfaction des contraintes de puissances d'antennes et de débits cibles est possible. Les techniques "OMA-HDAS-Realloc" et "NOMA-HDAS-Realloc" opèrent en deux phases. Dans un premier temps, les algorithmes antérieurs de minimisation de puissance sont appliqués en ne considérant que les antennes non contraintes en puissance. Dans un second temps, la possibilité de réaffecter certaines sous-porteuses par les antennes contraintes en puissance est étudiée pour réduire davantage la puissance du système. Finalement, une correction optimale de puissance est appliquée à ces deux familles de méthodes, si nécessaire.

Pour situer les performances des approches proposées, nous les comparons dans la Fig. 8 au cas le plus favorable où aucune contrainte en puissance n'est considérée (cas des méthodes "DAS"), et au cas le plus défavorable où l'antenne contrainte est éteinte (cas des méthodes "SOFF"). Notons tout d'abord l'important palier de puissance qui existe entre les méthodes orthogonales et non orthogonales, pour lesquelles l'algorithme le moins performant requiert un minimum de 40 W de moins que n'importe lequel des autres méthodes OMA pour une puissance limite de 20 W et un débit de 5 Mbps par utilisateur. Ceci montre encore une fois le potentiel important du NOMA en mutual SIC pour minimiser la puissance des systèmes de communications. L'intérêt d'observer la performance de nos algorithmes pour des contraintes de puissances limites de transmission élevées est de donner une idée du minimum de puissance atteignable par chacune de nos deux approches. Ainsi, il est clair que la technique OMA-HDAS présente un meilleur pouvoir de réduction de puissance que la technique OMA-HDAS-Realloc. Cependant, OMA-HDAS ne se rapproche de son potentiel que pour des valeurs de puissances limites relativement élevées. De plus, la puissance délivrée par OMA-HDAS augmente considérablement lorsque la puissance limite décroît ; jusqu'au point où elle finit par dépasser



(a) (b)

Figure 8 – Puissance totale consommée en fonction de la contrainte en puissance de l'antenne pour les cas OMA (a) et NOMA (b) pour $K = 38$ utilisateurs avec un débit cible de $R_{req} = 5$ Mbps.

la puissance totale requise par la technique OMA-SOFF. Or pour les mêmes conditions de débits cibles et de puissances limites, l'évolution de la puissance OMA-HDAS-Realloc est bien plus maîtrisée, ce qui lui permet de continuer à délivrer des résultats qui sont sensiblement meilleurs que la solution triviale OMA-SOFF. Le même constat peut être fait pour les méthodes NOMA dans la Fig. 8b puisqu'elles souffrent/profitent des mêmes avantages/inconvénients que les méthodes OMA.

En guise de conclusion, OMA-HDAS-Realloc est de loin la meilleure méthode pour les conditions d'opération les plus contraignantes (en termes de puissances limites, nombre d'utilisateurs et de débits cible par utilisateurs), alors que OMA-HDAS est la mieux adaptée pour des conditions plus favorables.

Chapitre 4 : Utilisation la technique NOMA mutual SIC pour augmenter l'efficacité spectrale des systèmes CoMP

La méthode mutual SIC trouve son origine dans l'application des principes du NOMA au contexte DAS, où les signaux multiplexés sont envoyés par différentes RRH. Dans ce chapitre, nous visons une généralisation du concept de mutual SIC pour couvrir le cas d'un nombre arbitraire d'utilisateurs (≥ 3). Ce faisant, nous développons un nouveau formalisme du mutual SIC qui peut être directement appliqué au DAS, C-RAN ainsi qu'à toute autre architecture de réseaux (HetNets, small cells, etc.) sous la condition d'existence de protocoles de signalisation permettant la coopération entre les points de transmission. Le cadre du CoMP est sélectionné pour conduire l'étude puisqu'il permet de couvrir les cas de cellules uniques (*single cell*) ou multiples (*multi-cell*) tout en considérant les transmissions conjointes des signaux par plusieurs points de transmissions ou par un seul point de transmission. Nous présentons dans ce chapitre le modèle du système et nous posons le problème de maximisation de débit sous contrainte de puissances limites.

Les conditions fondamentales de débits et de PMC permettant un mutual SIC généralisé sont développées pour le JT-CoMP et le DPS-CoMP, et les cas particuliers de deux et de trois utilisateurs sont évalués. Les contributions majeures de ce chapitre peuvent être résumées de la manière suivante :

- Nous proposons d'améliorer le débit des utilisateurs au bords de la cellule ainsi que le débit global du système en introduisant le service en mode JT non seulement pour les utilisateurs périphériques, mais aussi pour les utilisateurs centraux de la cellule ;
- Nous développons les conditions permettant une annulation d'interférence en NOMA pour les contextes DPS et JT et nous montrons que, contrairement à la croyance générale des travaux antérieurs de la littérature, l'annulation successive de l'interférence des signaux d'utilisateurs centraux à la cellule est possible au niveau des utilisateurs périphériques ;
- Nous définissions rigoureusement les conditions permettant la faisabilité du mutual SIC pour un nombre arbitraire d'utilisateurs et nous l'appliquons pour le cas de groupes NOMA de trois utilisateurs ;
- Nous montrons que le JT est plus favorable à une opération d'annulation d'interférence que le DPS, sans être pour autant une condition nécessaire pour implémenter le mutual SIC ;
- Nous remettons en question l'idée d'associer systématiquement l'utilisateur à son antenne la plus proche (en termes de puissance de signal reçu (RSS)) pour maximiser la capacité du système. Par là même, nous proposons une technique d'association des utilisateurs aux antennes qui garantit l'application du mutual SIC.

Exemples de résultats

Dans la méthode que nous proposons pour le cas de trois utilisateurs, tous les utilisateurs sont servis en mode JT et un mutual SIC est appliqué au niveau de toutes les paires d'utilisateurs prises deux par deux, d'où le nom de "FullJT-TMSIC". Elle est comparée à la méthode "CellEdgeJT-CellCenterSIC" de [30] où le JT n'est utilisé que pour l'utilisateur périphérique et les utilisateurs centraux n'appliquent le SIC que pour le signal de l'utilisateur périphérique. Une autre variante est aussi proposée sous le nom de "CellEdgeJT-TMSIC", qui tente d'appliquer le mutual SIC au niveau des trois utilisateurs tout en ne servant que l'utilisateur périphérique en transmission conjointe (JT-CoMP).

La comparaison des résultats entre CellEdgeJT-CellCenterSIC et CellEdgeJT-TMSIC de la Fig. 9 montrent les améliorations apportées par la simple adoption du TMSIC sans changement du moyen de transmission des signaux pour les utilisateurs centraux, qui sont servis dans les deux cas par sélection dynamique du point d'accès (DPS-CoMP). Il en résulte une augmentation sensible de l'ES au pic des deux courbes avec 18.2 bps/Hz pour le CellEdgeJT-CellCenterSIC contre 27.8 bps/Hz pour le CellEdgeJT-TMSIC. D'autre part, la comparaison entre le FullJT-TMSIC et le CellEdgeJT-TMSIC montre l'intérêt de l'utilisation du JT pour servir tous les utilisateurs. Les gains ainsi obtenus, qui sont amplifiés par l'application du *triple mutual SIC* (TMSIC), démontrent bien la supériorité du JT par rapport au DPS avec une augmentation de 66% de l'ES atteinte.

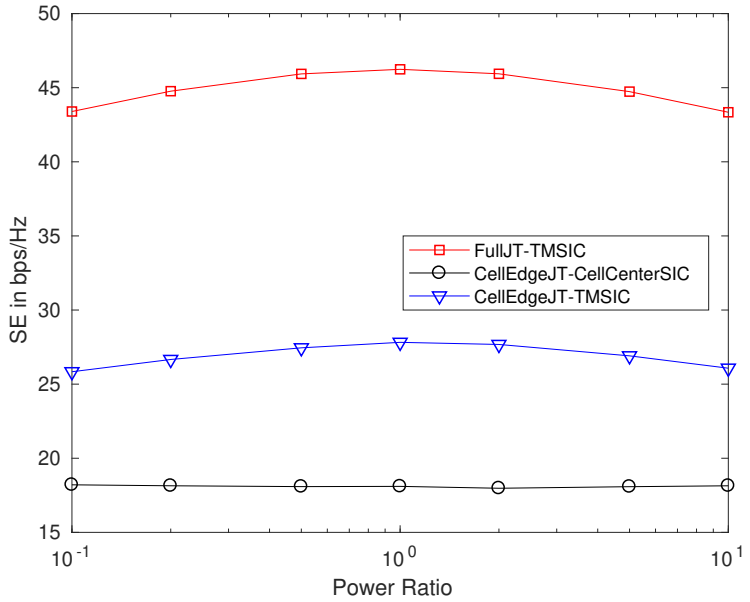


Figure 9 – Comparaison des procédures de maximisation du débit pour un système à trois utilisateurs et deux antennes avec $P_1 + P_2 = 4$ W.

Il est également intéressant d'observer le niveau d'équité atteint par ces méthodes pour fournir leurs ES respectives. L'indice d'équité de Jain [31] est affiché dans la table 1 pour une puissance totale des antennes $P_1 = P_2 = 2$ W. Cet indice prend la valeur 1 dans le cas d'une équité parfaite entre les utilisateurs (tous les utilisateurs atteignent la même ES), et de 1/3 pour le pire des cas (toute l'ES est atteinte par un seul utilisateur). Nous remarquons que le FullJT-TMSIC aboutit à une mesure d'équité qui est très proche de 1 (0.97) et qui est bien meilleure que celle obtenue pour le CellEdgeJT-CellCenterSIC (0.40). Le CellEdgeJT-TMSIC, quant à lui, délivre un niveau d'équité situé entre les deux. Ceci montre bien que, non seulement le FullJT-TMSIC est la meilleure stratégie au regard de l'ES obtenue, mais aussi qu'il délivre le plus haut niveau d'équité. En fait, c'est bien grâce à sa plus grande équité dans l'allocation de débit aux utilisateurs que le débit total du FullJT-TMSIC que nous proposons est supérieur aux autres.

Table 1 – Indice d'équité de Jain pour les systèmes à trois utilisateurs avec $P_{L_1}/P_{L_2} = 1$

	Jain fairness
FullJT-TMSIC	0.97
CellEdgeJT-CellCenterSIC	0.40
CellEdgeJT-TMSIC	0.62

Le gains considérables apportés par l'application du TMSIC suggèrent l'élaboration de techniques d'allocation de ressources en termes d'association d'utilisateurs, d'antennes et de sous-porteuses qui cherchent à favoriser la faisabilité du TMSIC avant tout. C'est par cette porte d'entrée que nous attaquons la problématique de positionnement d'antennes mobiles dans la cellule dans le chapitre suivant.

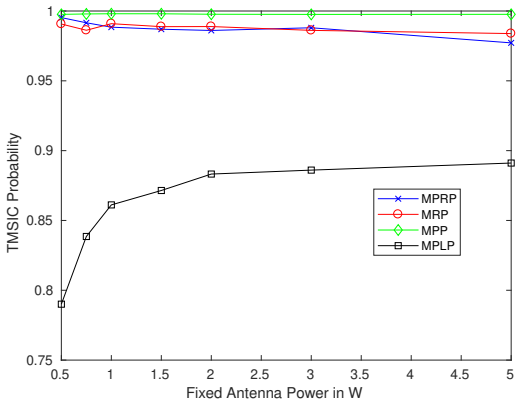
Chapitre 5 : Analyse des stratégies de placement de drones pour une annulation complète de l'interférence dans un système CoMP à deux cellules

L'utilisation de drones comme stations de base volantes se développe rapidement dans le domaine des communications sans fil afin d'apporter un support provisoire à des cellules encombrées. Ce chapitre considère un système à deux cellules où l'une des cellules est saturée, c'est-à-dire qu'elle ne peut plus servir ses utilisateurs, et est supportée par un véhicule aérien sans pilote (UAV) ou drone. Des procédures de positionnement du drone sont proposées pour alléger au mieux la charge de la cellule encombrée, avec une attention particulière portée à l'augmentation de l'ES du système par un service plus équitable des utilisateurs en bordure de cellule ainsi que des utilisateurs centraux de chacune des deux cellules adjacentes considérées. Dans le chapitre précédent, l'obtention d'un groupe d'utilisateurs sans interférence grâce à l'application du TMSIC a permis d'améliorer l'équité et l'ES du système. Par conséquent, l'idée maîtresse du placement des drones dans ce chapitre est de permettre le TMSIC tout en tenant compte des caractéristiques des liaisons air-sol (A2G) en termes de réalisations aléatoires de communications LoS et NLoS entre les utilisateurs et le drone. Les contributions majeures de ce chapitre peuvent être résumées comme suit:

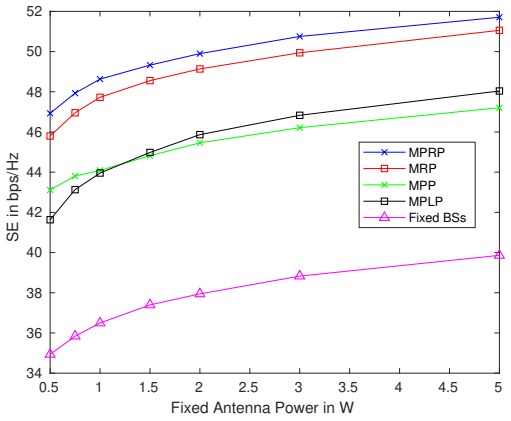
- Nous étudions le problème de positionnement des drones tout en prenant en compte les particularités du canal de propagation A2G en LoS/NLoS entre les utilisateurs et l'UAV au lieu de recourir au modèle de canal à évanouissement moyen qui est utilisé dans la littérature ;
- Nous introduisons un cadre d'étude probabiliste pour permettre le calcul de la probabilité de TMSIC associée à une position donnée du drone. Ceci permet la formulation du problème de positionnement du drone permettant de maximiser les chances d'application du TMSIC entre les utilisateurs ;
- Nous étudions plusieurs techniques de positionnement basées sur ce cadre probabiliste avec différents critères d'optimisation et nous les comparons aux techniques de positionnement basées sur la considération traditionnelle du *mean path loss*. Nous mettons également en évidence les compromis existant entre la capacité du système, l'équité et la complexité de calcul des approches étudiées.

Exemples de résultats

De par la méthodologie que nous avons élaborée pour placer les drones, les méthodes proposées de MPP, MRP et MPRP cherchent toutes à permettre l'application du TMSIC mais avec des objectifs différents pour chaque méthode. Le MPP vise une maximisation de la probabilité de TMSIC, le MRP vise une maximisation du débit atteignable par TMSIC, et le MPRP vise une maximisation du produit de la probabilité de TMSIC et du débit qui lui est associé. La méthode de MPLP, quant à elle, adopte le modèle de canal moyen (sans faire la distinction entre LoS et NLoS) pour effectuer le placement du drone. Nous commençons par remarquer dans la Fig. 10a que les méthodes MPP, MRP et MPRP qui prennent en compte le modèle aléatoire de propagation en LoS/NLoS, délivrent de bien meilleures probabilités de TMSIC que le MPLP. Ceci est normal puisque le MPLP base



(a) Probabilité de TMSIC en fonction de la puissance de l'antenne fixe.



(b) Efficacité spectrale en fonction de la puissance de l'antenne fixe.

ses calculs sur un modèle moins réaliste du canal qui ne permet pas de rendre compte des fluctuations dues au liaisons LoS et NLoS, et qui ont un impact certain sur la faisabilité du TMSIC. Par suite, nous notons le peu de différence qui existe entre les méthodes MPP, MRP et MPRP bien que la méthode MPP délivre la plus grande probabilité comme on pouvait s'y attendre. Ce résultat *a priori* contre-intuitif s'explique par la formulation générique des problèmes de positionnement que nous avons proposée, dans laquelle les conditions de débit et de PMC du TMSIC sont posées comme contraintes du problème. Il en résulte donc une faible différence en termes de probabilité de TMSIC.

Dans la Fig. 10b, l'ES atteinte pour chaque algorithme est représentée en fonction de la puissance de transmission à la station de base fixe. L'ES atteinte lorsque nous disposons de deux antennes fixes est aussi représentée à des fins de comparaison. L'amélioration des performances due à la mobilité des drones par rapport aux stations de base fixes est clairement observée pour toutes les techniques de positionnement. De plus, la prise en compte de la combinaison LoS/NLoS augmente significativement l'ES de 3 à 5 bps/Hz pour le MRP et le MPRP par rapport au MPLP. Cependant, la performance moyenne du MPP est à la traîne, car elle ne dépasse le MPLP que pour les petites valeurs de puissance limite $P_{L_1} = 1$ W avant de passer en dessous pour les valeurs limites de puissance supérieures à 1,5 W. Cela suggère que l'évolution de la position du drone avec l'augmentation de la valeur de P_{L_1} affecte les liaisons A2G de telle sorte que le taux d'augmentation du débit MPP est inférieur à celui de MPLP. En effet, une analyse du positionnement de l'UAV dans MPP et de son évolution avec la limite de puissance montre que les valeurs élevées de P_{L_1} ont tendance à placer l'UAV aux bords de la région de recherche, ce qui entraîne de faibles gains de canal et explique le débit plus faible par rapport à MPLP à $P_{L_1} = 5$ W.

Nous pouvons résumer les résultats de la figure 10b en affirmant que le fait de se concentrer exclusivement sur la probabilité TMSIC peut induire en erreur le placement du drone dans des zones où les liaisons A2G et le débit réalisable sont faibles. L'introduction du débit dans la fonction objectif donne un avantage qualitatif à la MRP par rapport à la MPP, puisque le débit est pris en compte pendant le positionnement, alors que la différence de probabilité TMSIC entre les deux est négligeable (cf. figure 10a). Cela dit, la combinaison du débit et de la probabilité dans le PRPM donne des résultats encore meilleurs puisque les deux objectifs sont pris en compte dès le début du processus de positionnement. Cependant, le gain en performance du MPRP et du MRP se fait au prix

d'une complexité supplémentaire par rapport au MPLP, puisque 64 combinaisons doivent être vérifiées pour le MRP et le MPP par rapport aux 8 ordres de décodage évalués par le MPLP. Dans le manuscrit, nous explorons plus en profondeur cette grande diversité dans les résultats de performance au niveau de chaque utilisateur, ce qui offre un large choix de sélection en fonction des priorités du système. Si la performance de l'utilisateur du bord de la cellule est prioritaire par rapport au débit total du système, le choix du MRP est le plus approprié. D'autre part, si la performance de l'utilisateur du centre de la cellule est priorisée, alors le MPRP et le MPP peuvent être employés, tout en gardant à l'esprit que le MPRP fournit la meilleure performance de débit global. Enfin, le MPLP peut également être utilisé pour favoriser l'utilisateur du bord de la cellule, tout en maintenant un bon débit global et en réduisant la complexité de l'optimisation par rapport au MRP en raison du modèle plus simple de *mean path loss*. Ce large éventail de choix fournit également au planificateur de réseau une multitude de réponses pour faire face aux variations dans le temps des exigences de trafic, où les priorités des utilisateurs peuvent changer et la stratégie de positionnement du drone peut être modifiée en conséquence.

Chapitre 6 : Application de NOMA mutual SIC dans les systèmes de communication *inband* D2D sous-jacents à un réseau cellulaire

Le nombre de dispositifs connectés ne cessant d'augmenter, des changements de paradigmes doivent être entrepris pour répondre à cette demande explosive. La communication D2D est l'une de ces solutions, qui permet d'augmenter le nombre de connexions, de réduire la latence et de décharger le trafic des réseaux mobiles sans nécessiter d'infrastructures de réseau supplémentaires. C'est pourquoi elle a suscité un intérêt croissant de la part du monde universitaire et de l'industrie au cours des dernières années [32–36]. Dans ce chapitre, nous proposons d'étudier l'interaction du NOMA mutual SIC avec l'écosystème D2D pour améliorer les performances du système. En supposant un réseau cellulaire pré-établi, l'objectif sera d'opérer le couplage D2D-utilisateur cellulaire (CU) et le contrôle de la puissance de telle sorte que le débit total du système D2D sous-jacent soit maximisé sans affecter la qualité de service des CUs. Le problème conjoint d'attribution des canaux et de la puissance est formulé, et il est montré que ce problème peut être séparé en problèmes disjoints d'allocation de puissance (PA) et d'attribution des canaux . Pour le problème de PA mutual SIC en mode FD, les conditions de mutual SIC pour FD-D2D sont d'abord dérivées, la réduction des contraintes du problème est ensuite effectuée, puis une résolution géométrique est proposée, permettant une résolution efficace du problème. Les principales contributions de ce chapitre peuvent être résumées comme suit :

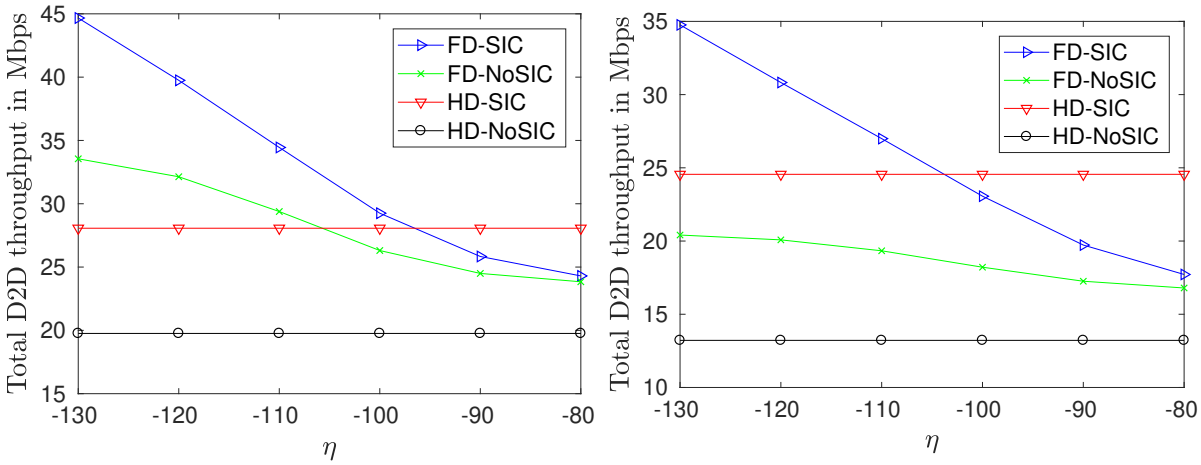
- Nous déterminons les conditions de SIC et de PMC permettant une annulation mutuelle de l'interférence entre D2D et CU ;
- Nous montrons que les PMC impliquent les conditions de SIC pour les modes de transmission HD et FD, ce qui permet une réduction considérable du problème de PA pour le cas du FD-SIC ;
- Nous résolvons analytiquement le problème de PA pour toutes les configurations, et particulièrement pour le FD-SIC où la méthodologie développée conduit à une

réduction drastique de la complexité ;

- La complémentarité entre le D2D et le NOMA mutual SIC est mise en évidence. La façon dont l'intégration du NOMA peut étendre l'applicabilité du D2D à des configurations d'utilisateurs et des scénarios de canaux plus larges est discutée.

Exemples de résultats

La Fig. 11 présente le débit total D2D en fonction du facteur d'annulation de la SI η, pour deux valeurs différentes de $R_{u,min}$ (débit requis des utilisateurs CU). On observe que les schémas d'allocation de ressources avec mutual SIC sont plus performants que leurs homologues sans SIC pour les scénarios de transmission HD et FD. En d'autres termes, les avantages de l'opération SIC en termes de SINR l'emportent sur la charge induite par les PMC supplémentaires sur la solution du problème de PA. En effet, une augmentation de 41 % du débit est observée sur la Fig.11a entre HD-SIC et HD-NoSIC (passant de 19,8 Mbps à 28,1 Mbps). Les augmentations de débit dues au mutual SIC pour le cas de la transmission FD varient de +2 % pour $\eta = -80$ dB à +33 % pour $\eta = -130$ dB. Les gains en performance du FD-SIC par rapport au FD-NoSIC augmentent avec les capacités d'annulation de la SI des dispositifs pour deux raisons : d'une part, la diminution de η relâche les contraintes d'applicabilité du mutual SIC, augmentant ainsi le nombre de paires D2D-CU qui bénéficient du FD-SIC (d'une moyenne de 0.36 paires D2D FD-SIC pour $\eta = -80$ dB à 1,92 paires pour $\eta = -130$ dB, avec $R_{u,min} = 1,5$ Mbps). D'autre part, la diminution de η réduit les termes d'interférence dans l'expression du débit D2D, ce qui se traduit par un débit atteint plus élevé.



(a) $R_{u,min} = 1.5$ Mbps (b) $R_{u,min} = 3$ Mbps

Figure 11 – Débit total du D2D en fonction du facteur d'annulation de la SI, η, pour $K = 20$ CUs, $D = 5$ paires de D2Ds, et $d_{max} = 100$ m.

Comme prévu, en comparant les performances pour différents débits d'UC requis entre les Figs. 11a et 11b, l'augmentation de R_u de $1,5$ Mbps à 3 Mbps diminue le débit D2D obtenu pour toutes les méthodes proposées. Cependant, le gain en pourcentage des performances des procédures SIC par rapport à NoSIC passe de 41 % à 86 % pour le cas HD, et de 33 % à 70 % pour le cas FD (pour $\eta = -130$ dB). La raison de cette augmentation de gain est que les algorithmes NoSIC sont fortement affectés par la valeur

de la puissance de CU (P_u) puisqu'ils souffrent de son interférence, ce qui n'est pas le cas des techniques SIC. En fait, même si le nombre total de paires D2D-CU pouvant aboutir à un FD-SIC diminue avec les contraintes plus sévères de mutual SIC dues à l'augmentation de $R_{u,min}$ (d'une moyenne de 1,6 paire pour $R_{u,min} = 1,5$ Mbps à 1,4 paire pour $R_{u,min} = 3$ Mbps, avec $\eta = -90$ dB), l'allocation de Munkres donne un nombre croissant de paires D2D-CU sélectionnées atteignant FD-SIC (ou HD-SIC) avec $R_{u,min}$ (d'une moyenne de 0,8 paire pour $R_{u,min} = 1,5$ Mbps à une moyenne de 1,24 paire pour $R_{u,min} = 3$ Mbps, avec $\eta = -90$ dB). Cela corrobore l'idée que la diminution du débit des techniques No-SIC avec $R_{u,min}$ est plus importante que celle des techniques SIC, à tel point que la contribution des techniques de mutual SIC dans la maximisation du débit est plus importante lorsque $R_{u,min}$ augmente. Ceci est vérifié en comparant le pourcentage de diminution du débit D2D pour chaque algorithme lorsqu'on passe de $R_u = 1.5$ Mbps à $R_u = 3$ Mbps : une diminution de 39 %, 33 %, 22 %, et 13 % est observée pour les algorithmes FD-NoSIC, HD-NoSIC, FD-SIC, HD-SIC respectivement. La plus grande réduction de performance de FD-NoSIC par rapport à HD-NoSIC justifie le déplacement du point d'intersection entre FD-SIC et HD-SIC vers la gauche lorsque $R_{u,min}$ augmente. En effet, FD-SIC et HD-SIC sont appliqués quand c'est possible, par dessus FD-NoSIC et HD-NoSIC respectivement. Si l'écart de performance entre FD-NoSIC et HD-NoSIC diminue, HD-SIC surpasse FD-SIC sur un plus large intervalle de valeurs de η avant que FD-SIC ne finisse par rattraper et dépasser HD-SIC pour des valeurs de η plus petites (c'est-à-dire pour de meilleures capacités d'annulation du SI des dispositifs).

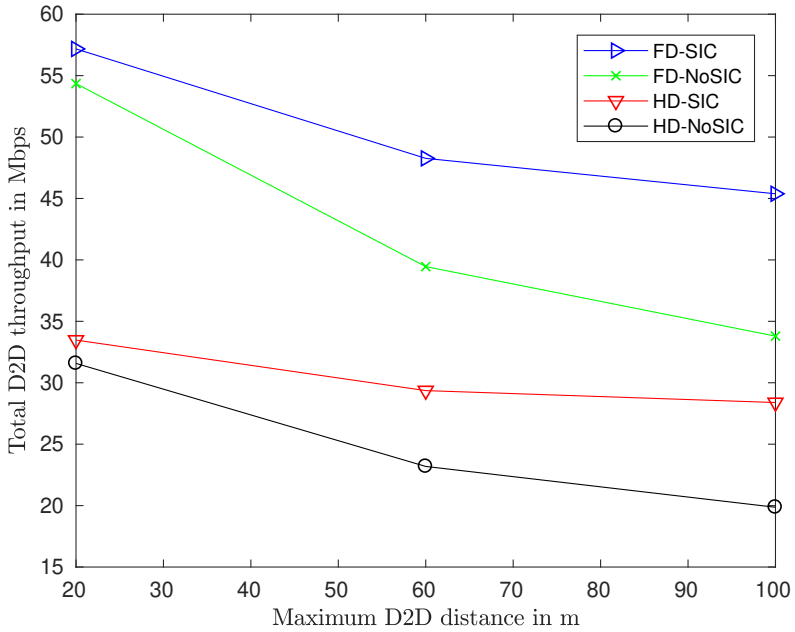


Figure 12 – Débit total du D2D en fonction de d_{max} pour un facteur d'annulation de la SI, η, de -130 dB.

Dans la Fig. 12, la variation du débit total D2D est présentée en fonction de la distance maximale de l'utilisateur D2D, d_{max}. L'augmentation de d_{max} conduit à une diminution significative des performances de toutes les méthodes proposées puisque h_d, le gain de canal de la liaison directe entre les utilisateurs d_1 et d_2 du couple D2D est réduit en moyenne. Cependant, cette augmentation de d_{max} s'accompagne d'une augmentation plus importante - en point de pourcentage - de la performance due au mutual SIC pour

les scénarios de transmission FD et HD, par rapport aux scénarios No-SIC. En effet, FD-SIC permet d'obtenir un débit D2D 128 % plus élevé que FD-NoSIC pour d_{max} = 100 m, qui est à comparer à l'augmentation de 81 % obtenue pour d_{max} = 20 m. Cela est dû au fait qu'il y a plus de paires D2D-CU compatibles avec le FD-SIC lorsque les utilisateurs D2D sont plus éloignés les uns des autres, puisqu'une moyenne de 1,96 paires appliquent FD-SIC avec succès pour d_{max} = 20 m contre 3,33 paires pour d_{max} = 100 m. La raison de cette augmentation est la diminution de h_d qui relaxe les conditions suffisantes de PMC, permettant ainsi plus de cas FD-SIC. Ceci met une fois de plus en évidence la complémentarité entre le D2D et le mutual SIC : bien que l'augmentation des distances D2D disqualifie généralement l'application classique du D2D, l'application du mutual SIC permet un regain d'intérêt pour la communication D2D.

Conclusions et perspectives

Dans cette thèse, nous avons étudié la combinaison de NOMA avec de multiples technologies de communication telles que D2D et FD, et des paradigmes de réseau comme DAS, CoMP, et UAVs afin de proposer des solutions nouvelles pour les réseaux de future génération reposant sur une gestion efficace des interférences.

Nous avons tout d'abord abordé le problème de la minimisation de la puissance de la liaison descendante dans une cellule DAS avec des exigences de taux d'utilisation. L'examen du concept de waterfilling pour l'allocation de puissance a permis de simplifier considérablement la complexité, ce qui a donné lieu à des schémas efficaces d'allocation conjointe de canal et de puissance pour le NOMA classique à une seule antenne. Ensuite, nous avons exploré les possibilités offertes par le DAS pour les signaux multiplexés en puissance provenant de différents RRHs. Cela a conduit à la définition du nouveau concept de mutual SIC qui a dévoilé les potentiels cachés de la diversité spatiale DAS et a permis une annulation complète des interférences entre utilisateurs. Les résultats obtenus ont montré la supériorité du NOMA mutual SIC par rapport à l'opération unique de SIC standard.

Pour aller plus loin, le cas pratique des antennes à puissance limitée a été exploré dans le contexte HDAS. La présence de contraintes de puissance sur les antennes de transmission pouvant potentiellement causer un échec dans la satisfaction des exigences de QoS des utilisateurs, les conditions d'allocation des canaux permettant de servir les utilisateurs avec succès ont alors été dérivées. La compréhension de ces contraintes a permis de façonner les stratégies d'allocation des ressources qui répondent aux demandes des utilisateurs pour diverses conditions de système. Deux approches distinctes ont été proposées pour tenir compte des limites de puissance de l'antenne pendant le processus de minimisation de la puissance : l'une donnant d'excellents résultats pour des conditions peu contraignantes, et l'autre présentant des performances robustes pour des conditions difficiles.

Par la suite, nous avons souhaité appliquer les principes de la procédure mutual SIC dans un cas plus général englobant les environnements multi-cellules dotés d'une coordination/coopération. Par conséquent, le concept de mutual SIC a été étendu pour prendre en compte la transmission JT-CoMP et un nombre arbitraire d'utilisateurs NOMA. Ensuite, les études de cas de double et triple mutual SIC ont été réalisées, montrant une amélioration considérable des performances par rapport aux techniques OMA JT-CoMP précédentes, ou aux techniques de SIC simples NOMA non coordonnées. En outre, un

résultat intéressant a été mis en évidence dans le cas deu double mutual SIC, où il a été démontré que favoriser l'interférence *annulable* par des choix non conventionnels d'association utilisateur-antenne peut être plus bénéfique que l'association traditionnelle antenne-utilisateur basée sur le RSS.

Les changements potentiels de paradigme dus au mutual SIC ont motivé la proposition de procédures de positionnement des réseaux assistés par drone qui permettent l'application du TMSIC et, par conséquent, héritent de tous ses avantages en termes d'équité et de débit. Un cadre probabiliste a été proposé pour tenir compte de la nature aléatoire des liaisons air-sol entre le drone et les utilisateurs, tout en visant une application TMSIC. Plusieurs métriques d'optimisation ont été proposées, fournissant un large panel de sélection pour le planificateur de réseau avec une multitude de réponses pour faire face aux variations dans le temps des besoins de trafic des utilisateurs.

Enfin, l'écosystème des communications D2D a été abordé en conjonction avec la communication FD et NOMA entre les CU et les dispositifs D2D. Les conditions de mutual SIC spécifiques à la communication FD-D2D ont été étudiées en détail et les conditions de canal nécessaires et suffisantes ont été identifiées. En outre, une représentation géométrique de l'espace de solution a permis une résolution optimale efficace du PA, permettant des affectations optimales ultérieures de D2D aux CU. De plus, l'application de la procédure de mutual SIC dans le contexte D2D s'est avérée particulièrement bénéfique à plusieurs égards. D'une part, des gains de performance significatifs ont été obtenus grâce à l'annulation des interférences, par rapport à la stratégie classique sans SIC entre les CU et les D2D. D'autre part, la mise en œuvre du mutual SIC a montré une grande complémentarité avec les applications D2D : lorsque le D2D classique ne parvient pas à apporter une augmentation de capacité supplémentaire à un système sans fil, en raison de distances D2D trop élevées, le mutual SIC peut être appliqué pour tirer parti de l'effet *near-far*.

Travaux futurs

Le travail présenté dans cette thèse a montré comment le concept clé de mutual SIC peut être adapté à divers scénarios de réseau et de cas d'usage tels que DAS, CoMP, réseaux assistés par drone et communications D2D. Ceci est bien normal puisque tout nouvel atout pour lutter contre les interférences est précieux pour les réseaux de demain qui seront sérieusement limités par les interférences. Pourtant, plusieurs aspects de ces études sont loin de dévoiler tout leur potentiel.

Tout d'abord, les schémas d'allocation de ressources dérivés supposent une connaissance parfaite de l'état du canal. Dans la pratique, cela est difficilement réalisable, et des recherches supplémentaires sont nécessaires pour déterminer le résultat des techniques de RA proposées dans le contexte d'une connaissance imparfaite de la statistique du CSI et/ou du CSI instantané. Par conséquent, une direction de travail possible pourrait être de concevoir des schémas RA robustes atténuant l'écart de performance entre le CSI parfait et le CSI bruité, où différents modèles de bruit CSI pourraient être supposés selon le contexte : [37–39].

Une suite directe de cette étude serait l'analyse de l'impact d'une mise en œuvre imparfaite de SIC sur les performances des procédures proposées. D'une part, le CSI erroné pourrait induire l'administrateur du réseau en erreur en l'amenant à appliquer le mutual SIC dans des scénarios inadéquats, ce qui pourrait se retourner contre lui en termes

d'interférence subie. D'autre part, des interférences résiduelles pourraient subsister à la suite d'une procédure SIC imparfaite en raison d'erreurs de quantification et d'estimation du canal résultant d'une égalisation imparfaite. La dégradation de performance induite nécessiterait des tests supplémentaires et éventuellement une atténuation par des schémas RA robustes prenant en compte l'imperfection mentionnée dans leur conception.

Bien que nous ayons proposé une procédure de mutual SIC généralisée dans le scénario CoMP, la complexité exponentielle des ordres de décodage qui l'accompagne limite la taille des groupes NOMA à un maximum de trois utilisateurs. Les travaux futurs pourraient consister à combiner des analyses expérimentales et théoriques afin de déterminer les ordres de décodage les plus probables pour obtenir un mutual SIC. Cela permettrait d'obtenir des gains de capacité linéaires pour chaque nouvel utilisateur ajouté, sans compromettre la complexité de l'ordonnancement. Un suivi direct de cette étude peut consister à concevoir des stratégies de regroupement des utilisateurs permettant un nombre maximal d'applications de mutual SIC. À cet égard, les techniques de pointe de regroupement centré sur l'utilisateur dans la CoMP peuvent être envisagées pour inclure plusieurs utilisateurs à la fois. En outre, l'étude peut être étendue pour explorer la mise en œuvre du mutual SIC dans les systèmes à entrées multiples et à sorties multiples.

Dans le dernier chapitre, la procédure géométrique proposée pourrait inspirer la résolution de problèmes PA de plus grande dimensionnalité où plus d'une seule CU accède à la même ressource que la paire D2D, ou inversement, plus de deux dispositifs sont en communication D2D. En outre, il pourrait être intéressant de dériver des modèles pour le couplage D2D-CU qui seraient purement basés sur la connaissance des conditions de canal, ou même plus loin, sur leur positionnement géographique relatif. Cela pourrait être réalisé à l'aide de divers outils (par exemple, des techniques d'apprentissage automatique), ce qui simplifierait l'étape d'attribution des canaux et faciliterait l'intégration des méthodologies proposées aux DAS.

Liste de Publications

Les résultats de ces travaux ont donné lieu à plusieurs publications :

Publication en conférence internationale

- A. Kilzi, J. Farah, C. Abdel Nour and C. Douillard, "Inband Full-Duplex D2D Communications Underlaying Uplink Networks with Mutual SIC NOMA," 2020 *IEEE 31st Annual Int. Symp. Pers., Indoor and Mobile Radio Commun. (PIMRC)*, London, United Kingdom, Sept. 2020.

Publications en articles de revues:

- A. Kilzi, J. Farah, C. A. Nour and C. Douillard,"Optimal Resource Allocation for Full-Duplex IoT Systems Underlaying Cellular Networks with Mutual SIC NOMA," in *IEEE Internet Things J.*, May 2021.

- A. Kilzi, J. Farah, C. Abdel Nour and C. Douillard, "Analysis of Drone Placement Strategies for Complete Interference Cancellation in Two-Cell NOMA CoMP Systems," in *IEEE Access*, vol. 8, pp. 179055-179069, Sept. 2020.

- A. Kilzi, J. Farah, C. Abdel Nour and C. Douillard,"Mutual Successive Interference Cancellation Strategies in NOMA for Enhancing the Spectral Efficiency of CoMP Systems," in *IEEE Trans. Commun.*, vol. 68, no. 2, pp. 1213-1226, Feb. 2020.
- A. Kilzi, J. Farah, C. Abdel Nour and C. Douillard, "New Power Minimization Techniques in Hybrid Distributed Antenna Systems With Orthogonal and Non-Orthogonal Multiple Access," in *IEEE Trans. Green Commun. Netw.*, vol. 3, no. 3, pp. 679-690, Sept. 2019.
- 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," in *IEEE Trans. Veh. Technol.*, vol. 67, no. 12, pp. 11873-11885, Dec. 2018.

Introduction

Nowadays, the place of mobile communications in the modern society spans much beyond the Information and Communications Technology (ICT) community, as it has contributed to shape today's society in unprecedented ways. For instance, the interplay between offer and demand follows Say's law where the offer enables applications through which greater demand surges, requiring thereby further supply [1]. The technical advancements, providing greater ease of use and wider services, have penetrated the consumer's day to day lives, substantially reshaping human activities. This effect, coupled with the ferocious competition for higher market shares, pushes Mobile Network Operators (MNOs) for promoting more hype through the advertisement industry and, as a result, conditions the society to always expect for more. The greedy side being triggered, every new technological breakthrough (e.g. birth of the iPhone in 2007) enables a new range of applications, which penetrates into societies' habits and morphs into actual needs, justifying thereby further demand to which the offer has to cope with. This self-reinforcing feedback resulted in an ever increasing demand for higher data rates, further data volumes, more connected devices, lower latency requirements for cheaper data plans [2]. Meanwhile, the emergence of the Internet of Things (IoT), machine-to-machine and vehicle-to-vehicle communications and other technologies greatly complexifies the traffic profiles, imposing the constraint of a greater flexibility from MNOs to meet the diversified demands of current and future generation networks. The International Telecommunication Union (ITU) has defined the requirements for International Mobile Telecommunications (IMT) for 2020 and beyond, its three main pillars being enhanced Mobile BroadBand (eMBB) (4K video streaming, virtual and augmented reality, viewpoint video, etc.), Ultra-Reliable Low-Latency Communications (URLLC) (e.g. remote medical surgery, transportation safety) and massive Machine-Type Communications (mMTC) (e.g. smart metering, network sensing).

To meet these stringent requirements, MNOs will have to resort to every tool available at their disposal and carefully aggregate them in order to provide the expected leaps in performance. In an effort to provide solutions for network growth, lessons could be learned from the past. According to the analysis of Cooper reported in [40] by Prof. Webb on the main enabling techniques for higher system capacity, the lion's share goes to the densification of mobile network deployment, allowing for a confined serving and enabling higher spectrum reuse. This is the driving idea behind the proliferation of small cells [41–43], Distributed Antenna Systems (DAS) and Cloud Radio Access Network (C-RAN) [44–46], which leads to the complex Heterogeneous Networks (HetNet) [47–49]. This being said, network densification cannot be indefinitely exploited as it meets its limits in the increasing inter-cell interference it generates. To unleash its full potentials, efficient interference management techniques need to be introduced, hence the adoption of the Coordinated Multipoint (CoMP) framework for inter-cell management, since 3GPPP release 11 [9], and its further enhancements in releases 14 and 15. Moreover, the surge of newly connected

devices will require breaking the orthogonality of previous generations multiple access schemes to accommodate more User Equipments (UEs) and provide a higher spectral efficiency using the same resources. As a result, Non-Orthogonal Multiple Access (NOMA) was selected as a study item in Long Term Evolution (LTE) release 13 - termed Multiuser Superposition Transmission (MUST) - and was adopted in the LTE standard since release 14 as an efficient component to tackle the capacity crisis. The key enabler is the adoption of complex receivers capable of canceling the interference of undesired signals. Thus, the interference management problem is the cornerstone for enabling further advancements of mobile communications. The interest for Device to Device (D2D) communications coupled to full-duplex (FD) transmission attests of this observation [33, 50–52]. On the one hand, interference avoidance is favored by allowing end-devices to bypass the network infrastructure and exchange information on a proximity-based trigger for a minimum interference footprint. On the other hand, the advancements of self-interference cancellation of FD receivers enable a virtual doubling of the capacity without requiring any additional network resources.

The general theme of this thesis revolves around the interference management problem for various mobile communications scenarios. Our aim is to efficiently combine the mentioned technologies hereinabove, to assess their combined gains and explore the key specific properties arising from such combinations. In a first part of the thesis (Chapters 2 and 3), we will be investigating NOMA signaling in DASs to meet the users Quality of Service (QoS) demands with minimal power consumption. The powering of multiplexed signals from different antennas paves the way for a complete intra-cell Successive Interference Cancellation (SIC) that we called "mutual SIC". This newly unveiled tool for interference management shows great potentials for the green communication scenarios of Chapters 2 and 3, as well as the rate craving scenarios studied in the second part of the thesis, in Chapters 4 to 6. Therefore, the mutual SIC concept is further investigated in the general framework of CoMP in Chapter 4, where its interference cancellation properties can be efficiently applied to combat inter-cell interference and enhance cell-edge user experience. Afterwards, an application of mutual SIC is proposed in Chapter 5 for the context of Unmanned Aerial Vehicules (UAV) placement in UAV-assisted mobile networks. Finally, the field of D2D communications is approached where the interference resulting from competing CUs and D2D devices is managed through the mixing of interference avoidance schemes with interference cancellation schemes in the mutual SIC procedure.

Thesis outline

Chapter 1 presents a general overview of the main enabling techniques for future generation networks which are tackled in this thesis. First, the principles of downlink power domain NOMA are presented. Then, the motivation for network densification is discussed, and the evolution of the network architectures going from Centralized Antenna Systems (CAS) to DAS, then to C-RANs, is reviewed in general, and more specifically from the perspective of Resource Allocation (RA). Afterwards, the concept of CoMP is explored for its potentials to efficiently manage the problems of inter and intra-cell interference resulting from the densified network topologies. Finally, we present the framework of D2D communications as another means for boosting the network performance and meeting the diversified demands, and we present the necessary technical background of FD

communications to be used in conjunction with D2D.

In Chapter 2, NOMA signaling is combined with the DAS setup to address the problem of downlink power minimization under strict user rate requirements. The chapter first presents a review of the state-of-the-art research on downlink power minimization using NOMA. Then, after presenting the system model and formulating the optimal RA problem, we discuss the need for suboptimal RA schemes and separating the NOMA layer from the Orthogonal Multiple Access (OMA) layer. Thus, optimal Power Allocation (PA) for fixed channel assignment is revisited for OMA signaling, and the design of an iterative joint channel and power allocation for OMA is proposed. Afterwards, a new PA scheme for the NOMA pairing step is proposed. In the second part of the chapter, we lay the theoretical background for the application of NOMA multiplexing in DAS, where the mutual SIC concept is introduced for the case of two multiplexed users per subband. Finally, several PA techniques are proposed for the application of mutual SIC in NOMA DAS.

In the third chapter, we adapt the power minimization solutions of Chapter 2 to the practical scenario of antenna-specific power limit constraints referred to as hybrid DAS. The optimal RA problem is reformulated for the context of hybrid DAS and then the optimal PA solution is derived for the case of OMA. A simple criterion is developed to guarantee the existence of viable RA schemes. Afterwards, two joint RA schemes are proposed to solve the power minimization problem in OMA. The first one accounts for the power constraints at the end of the algorithm, while the other considers the antenna power limits throughout the user-antenna-subcarrier allocation process. Finally, the additional NOMA layer, including iterative user pairing and subcarrier power allocation, is remodeled to capture the characteristics of the hybrid DAS scenario.

Chapter 4 is devoted to the generalization of the mutual SIC concept to multiple antenna systems in a multi-cell CoMP framework. A new mathematical formalism is proposed to generalize the mutual SIC concept to arbitrary NOMA cluster sizes, while including multi-point transmission methods. We highlight the inter-relation between the decoding orders at the level of every user with the generalized mutual SIC conditions, and we derive the fundamental conditions of Power Multiplexing Constraints (PMCs) and rate constraints enabling the application of mutual SIC for two CoMP configurations: joint transmission and dynamic point selection. Afterwards, we present two case studies: Dual Mutual SIC (DMSIC) for two-user clusters and Triple Mutual SIC (TMSIC) for three-user clusters.

In Chapter 5, a UAV positioning favoring TMSIC is proposed to support a two-cell system with a saturated antenna. The UAV positioning problem is formulated to account for TMSIC application. Next, a mathematical framework for modeling the problem of TMSIC feasibility through UAV placement in probabilistic terms is introduced. Several UAV positioning strategies are proposed based on different network optimization metrics related to the probabilistic model. Then, the used methodology for adequate performance assessment is presented. Finally, the trade-offs between the proposed methods are highlighted, and the use case applications for every positioning technique are discussed according to the user serving priorities.

Chapter 6 investigates the application of NOMA signaling to D2D communications with half and full duplex modes. The first section presents the state-of-the-art studies on NOMA with D2D applications. Next, the system model is presented and the joint channel and power allocation problem is formulated. The mutual SIC conditions for half-duplex

D2D are derived next and the resolution of the optimal PA is presented. Afterwards, the case of FD D2D applying mutual SIC is investigated. The constraints of mutual SIC and PMCs are derived, and a constraint reduction procedure is conducted to simplify the problem without any incidence on the solution performance. Then, a geometrical representation of the PA problem is proposed, enabling an efficient low-complexity resolution method. Finally, the complete resource (channel and power) allocation problem is solved by using the Munkres algorithm.

The last chapter concludes this thesis and suggests future works that could be conducted in each of the approached research domains.

Contributions of the thesis

This work contains several original contributions that are proposed to cope with the increasing expectations from mobile networks. Below is a summary of the main contributions of this thesis.

In Chapters 2 and 3, the aim was to derive efficient resource allocation schemes for the minimization of the downlink system power under QoS constraints. The main contributions of the first two chapters are:

- proposing an iterative waterfilling scheme that greatly reduces the complexity of NOMA user pairing,
- proposing a new PA scheme for single antenna NOMA pairing which outperforms the standard fractional transmit PA scheme without incurring additional complexity,
- introducing the concept of mutual SIC where interference cancellation is conducted at the level of all paired NOMA users on a subcarrier, achieving an important reduction in the transmit power, compared to single-SIC NOMA,
- providing and analyzing the optimal PA scheme for power minimization in the context of OMA hybrid DAS, leading to the proposal of a systematic criterion for assessing the feasibility of the solution given a predefined subcarrier allocation,
- proposing two different approaches for the joint channel and power allocation in HDAS, one being robust against harsh system conditions in terms of high user rates and low power antennas, the other being particularly suited for mild system conditions.

From Chapters 4 to 6, the intent is to profit from the interference cancellation capabilities of mutual SIC to maximize the system spectral efficiency in the various scenarios of CoMP serving, UAV assisted networks, and D2D enabled networks employing FD. The main contributions of every chapter can summarized as follows.
In Chapter 4:

- the generalization of the mutual SIC principle is conducted, from the case of two users with a single transmission antenna per signal, to an arbitrary number of multiplexed users with joint transmission of signals through multiple antennas,
- the proposal of a simple user-antenna pairing scheme to always enable mutual SIC application, challenging the practice of user-antenna association based on the maximum received signal strength criterion,

- the extension of joint transmission serving to cell-center users is proposed as a means to enhance the overall system performance and the cell-edge performance through the use of mutual SIC.

In Chapter 5:

- a probabilistic framework is introduced to account for the specificity of line-of-sight/non line-of-sight propagation and enable the computation of the TMSIC probability associated to the UAV position,

- several approaches using different optimization metrics are proposed showcasing the trade-offs between system capacity, user fairness and computational complexity.

In Chapter 6:

- the conditions for applying mutual SIC in FD-D2D systems are defined, and the PMC conditions are proven to encompass the rate conditions of mutual SIC,

- a geometrical representation is introduced to reduce the search space of optimal PA for FD-D2D using NOMA, enabling optimal RA by successive application of optimal PA and optimal channel assignment.

These contributions led to the following list of publications:

Journal papers

- A. Kilzi, J. Farah, C. Abdel Nour and C. Douillard, "Optimal Resource Allocation for Full-Duplex IoT Systems Underlaying Cellular Networks with Mutual SIC NOMA," under revision in *IEEE Internet Things J.*

- A. Kilzi, J. Farah, C. Abdel Nour and C. Douillard, "Analysis of Drone Placement Strategies for Complete Interference Cancellation in Two-Cell NOMA CoMP Systems," in *IEEE Access*, vol. 8, pp. 179055-179069, Sept. 2020.

- A. Kilzi, J. Farah, C. Abdel Nour and C. Douillard,"Mutual Successive Interference Cancellation Strategies in NOMA for Enhancing the Spectral Efficiency of CoMP Systems," in *IEEE Trans. Commun.*, vol. 68, no. 2, pp. 1213-1226, Feb. 2020.

- A. Kilzi, J. Farah, C. Abdel Nour and C. Douillard, "New Power Minimization Techniques in Hybrid Distributed Antenna Systems With Orthogonal and Non-Orthogonal Multiple Access," in *IEEE Trans. Green Commun. Netw.*, vol. 3, no. 3, pp. 679-690, Sept. 2019.

- 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," in *IEEE Trans. Veh. Technol.*, vol. 67, no. 12, pp. 11873-11885, Dec. 2018.

Conference paper

- A. Kilzi, J. Farah, C. Abdel Nour and C. Douillard, "Inband Full-Duplex D2D Communications Underlaying Uplink Networks with Mutual SIC NOMA," 2020 *IEEE 31st Annual Int. Symp. Pers., Indoor and Mobile Radio Commun. (PIMRC)*, London, United Kingdom, Sept. 2020.

Chapter 1

Background

We present in this chapter an overview of the main multiple access schemes, network architectures, and communication techniques that we address throughout this dissertation. We first discuss, in section 1.1, how the increasing number of connected devices pushes towards the adoption of NOMA, and then we present the principles of power domain NOMA, showcasing its advantages and highlighting its theoretical and practical conditions of application. In section 1.2, we elaborate on the paradigm shifts when moving from centralized antenna systems to densified distributed architectures such as DAS and Cloud Radio Access Networks (C-RANs). The enabling techniques due to DAS are presented from the perspective of resource allocation. Afterwards, given that network densification is self limited by the inter-cell interference it generates, the principles of CoMP, the most advanced framework for inter-cell interference coordination, are presented in section 1.3. Finally, the context of D2D communications is described in section 1.4. Its potentials to meet the diversified demand as well as offload the data traffic from the network core to its front-end devices are explained. Moreover, the symbiotic relationship that D2D holds with FD communications is presented.

1.1 Principles of Downlink NOMA

Historically, multiple access schemes have characterized every new generation of cellular networks. They include Frequency Division Multiple Access (FDMA) in 1G systems, Time Division Multiple Access (TDMA) in 2G, Code Division Multiple Access (CDMA) in 3G, and Orthogonal Frequency Division Multiple Access (OFDMA) in 4G networks. The pursuit in all these multiple access schemes was to enable broader multiple access by exploiting the orthogonality in the different dimensions of the network system, i.e. users are allocated distinct frequency channels or time slots or signature codes or resource blocks so that their signals do not interfere with one another when they access the network. This common theme of "orthogonality" is rooted back to the idea of interference avoidance through resource partitioning. However, with the rapid growth of mobile networks, it has become more and more evident that the "orthogonality" feature of multiple access will be a serious limiter to the number of accommodated users. Therefore, NOMA has been resorted to in order to cope with the increasing demand for connected devices [53, 54]. NOMA comes in various forms and techniques such as Multi-User Shared Access (MUSA) [55], Low Density Spreading (LDS) [56], Sparse Code Multiple Access (SCMA) [57], Power Domain NOMA (PD-NOMA) [28], Pattern Division Multiple Access (PDMA) [58] or Bit

Division Multiplexing (BDM) [59]. In this thesis, we will be dealing with PD-NOMA which was mainly adopted for downlink transmissions. Consequently, from hereinafter, PD-NOMA is simply referred to as NOMA.

NOMA breaks the orthogonality by allocating the same time/frequency resource to multiple users at the expense of additional receiver complexity. At the transmitter side, signals of different users are allocated different power levels (hence the power domain nomenclature), and superposition coding is used to transmit the combined users signals. Let x_1 and x_2 be the multiplexed signals of UEs 1 and 2, with respective powers P_1 and P_2, and let h_1 and h_2 be their experienced channel gains with $|h_1| > |h_2|$. In the NOMA framework, UE1 is referred to as the *strong* user, while UE 2 is labeled as the *weak* user. A higher power level is allocated to the weak user ($P_2 > P_1$) to compensate for its weaker channel gain, provide user fairness, and allow the decoding of UE 2's signal at the level of UE 1. The super-imposed signal transmitted by the Base Station (BS) is given by $x = x_1 + x_2$, and the received signals y_1 and y_2 at the level of UE 1 and UE 2 are given respectively by: $y_1 = xh_1 + n_1$ and $y_2 = xh_2 + n_2$, where n_i represents the Gaussian noise received by UE i with average power σ^2. At the level of UE 1, the SIC receiver is applied to extract x_1 from the total received signal. It proceeds first by detecting, demodulating and decoding the dominant signal which is x_2, prior to subtracting it from the total received signal as shown in Fig. 1.1.

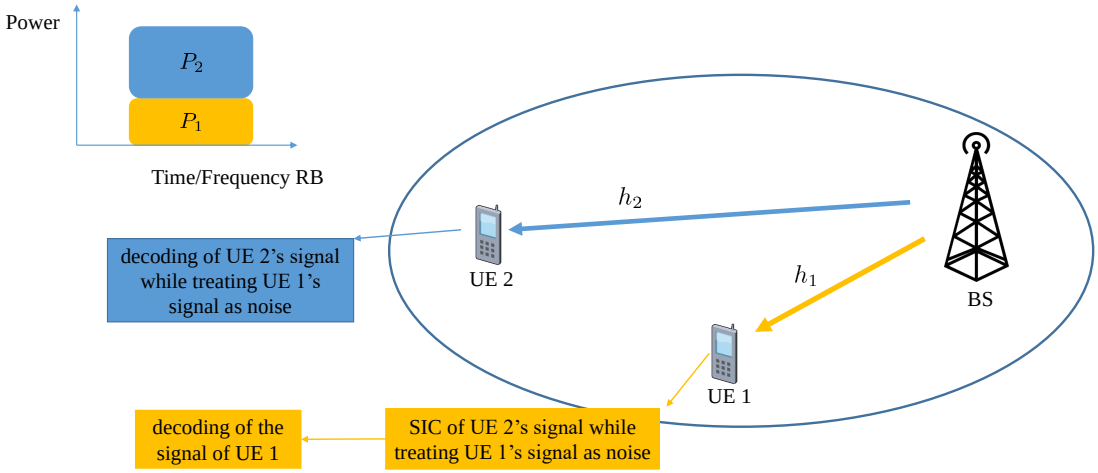


Figure 1.1 – Representation of a two-user NOMA system with UE 1 performing SIC before retrieving its signal.

Consequently, x_1 is decoded in an interference-free manner and its achievable rate according to the Shannon channel capacity theorem is given by:

$$R_1 = \log_2\left(1 + \frac{P_1|h_1|^2}{\sigma^2}\right).$$

At the level of the weak user, UE1'signal is treated as additional interference, and the achievable rate is given by:

$$R_2 = \log_2\left(1 + \frac{P_2|h_2|^2}{P_1|h_2|^2 + \sigma^2}\right).$$

For the general case of m multiplexed users with channel gains such $|h_1| > |h_2| > \ldots > |h_m|$, the power is allocated according to the descending order of channel gains, i.e. $P_m >$

$P_{m-1} > \ldots > P_1$ [60,61]. The i^{th} UE iteratively decodes and subtracts the signals of users with weaker channel gains - starting from x_m to x_{i+1} - before retrieving its own signal x_i while suffering from the interference of the remaining $i-1$ users. The rate for UEi is thus given by:

$$R_i = \log_2\left(1 + \frac{P_i|h_i|^2}{\sum_{j=1}^{i-1} P_j|h_i|^2 + \sigma^2}\right), \tag{1.1}$$

Note that for the case of three or more multiplexed users, enforcing $P_m > P_{m-1} > \ldots > P_1$ at the BS is not enough to guarantee that x_m will remain the dominant signal, since lower power signals may add up to a greater power level than P_m. This may induce SIC error propagation, threatening the throughput gains achieved by NOMA. To avoid running into that problem and to ensure SIC stability, the enforced Power Multiplexing Constraint (PMC) for the general case of m multiplexed subcarriers is done as follows:

$$\begin{aligned} P_m &> P_{m-1} + P_{m-2} + \cdots + P_1, \\ &\vdots \\ P_i &> P_{i-1} + P_{i-2} + \cdots + P_1, \\ &\vdots \\ P_2 &> P_1. \end{aligned} \tag{1.2}$$

This being said, in the literature, most papers considering downlink NOMA limit the number of multiplexed users to a maximum of three [28, 62, 63] since it was shown that the additional rate gains become marginal when m further increases [27], while the receiver complexity grows linearly with m. Note that the growing computational power of mobile devices enabled the implementation of interference cancellation as they have been incorporated in wireless standards under the name of Network-Assisted Interference Cancellation and Suppression (NAIC) in LTE since 3GPP release 12 [64].

1.2 Network Densification and Distributed Antenna Systems

The basic idea behind network densification is to bring network access nodes closer to the end users through the spreading of multiple Transmission Points (TPs) throughout the cell instead of having them grouped at the same location as for CAS. This enables a better cell coverage and enhances the cell capacity by improving the link quality due to reduced path loss and additional spatial diversity favoring Line-of-Sight (LoS) communication. Moreover, network densification increases the reuse per unit area of the available spectrum which significantly improves the network capacity.

1.2.1 Distributed and Centralized Densification

Network densification can be classified into distributed and centralized densification. Distributed densification corresponds to the geographical deployment of small cells, in areas where immense traffic is generated. Small cells, pico cells and femto cells are fully functioning BSs, capable of performing all the macro-cell functions (baseband and radio

processing) but with a lower power and smaller coverage areas. Each small cell having its own backhaul connection, coordination among them is not straightforward and distributed interference management protocols are required [3, 4]. On the other hand, when the baseband processing unit of a BS is decoupled from its radio units, centralized network densification can be achieved in DAS by deploying the Remote Radio Heads (RRHs) throughout the cell, while connecting them to a central processing unit referred to as Baseband Unit (BBU) through high-speed low-latency optical fibers. RRHs are responsible for the digital to analog conversion, analog to digital conversion, power amplification and filtering [65], while the BBU handles all the baseband processing, and higher level procedures such as user scheduling, medium access control, and Radio Resource Management (RRM). This star-like architecture achieves a full coordination between RRHs. The differences between DAS and small cells are depicted in Fig. 1.2.

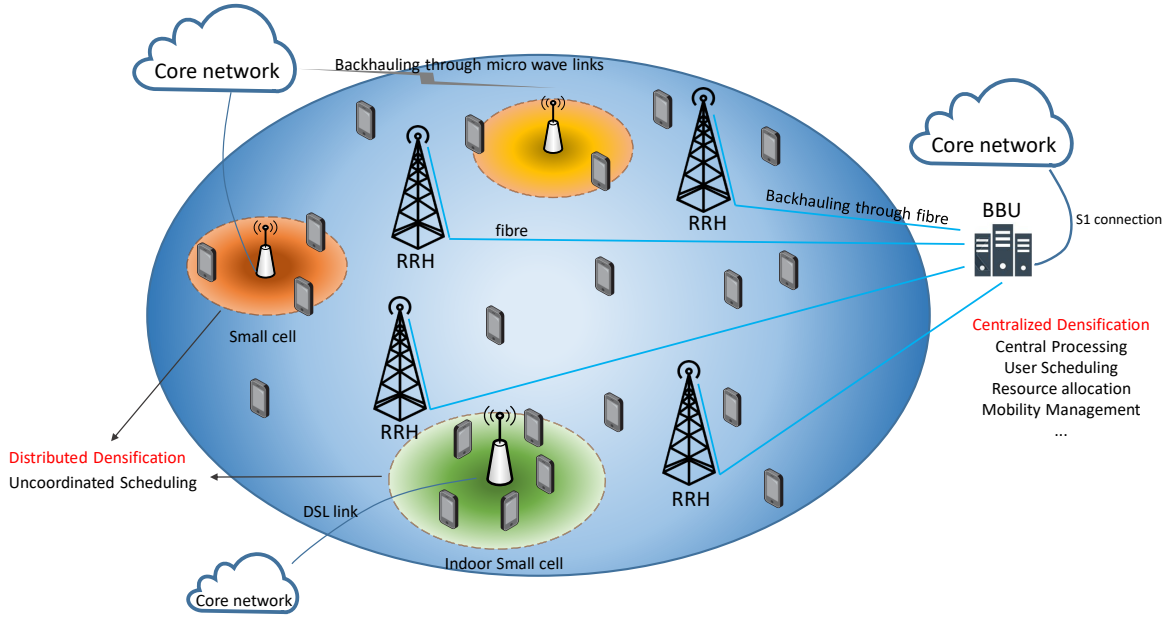


Figure 1.2 – Schematic of a densified heterogeneous network consisting of stand-alone small cells with individual backhaul connection, and distributed RRHs controlled by a single BBU entity.

Throughout the literature, a distinction has been made between deploying antennas for improving coverage as opposed to improving capacity. Small cell systems are typically seen as capacity boosters, capable of providing important capacity gains for small regions of high network activity by reusing the cell frequency. In this scenario, having a small coverage region enables a localized high capacity region that does not leak out excessive interference to the neighboring sites. On the other hand, coverage strengthening was the primary objective of early DAS deployment [44], where signals were simulcasted across all of the antennas to blanket the coverage region. While reasonable from a pure coverage perspective, this approach has the drawback of causing important out-of-cell interference compared to both small cells and CAS. Moreover, studies such as [5, 6] showed that more efficient user serving can be achieved through selection diversity, where one of the RRHs is selected to transmit the user signal. This approach is shown to provide greater capacity

and more power efficient user serving. Also, thanks to the centralized densification of DAS, BBU scheduling can operate such that some RRHs reuse all the spectrum while other RRHs dynamically share the cell frequency. For all these reasons, the potentials of DAS seem to us more appealing than those of small cells, especially from the perspective of resource allocation. That is why in this thesis, great importance has been given to the DAS setup with selection diversity in the proposed RA schemes.

1.2.2 More on Network Centralization

Network centralization can be taken one step further by grouping the BBUs of multiple cell sites in the same location, to form a shared BBU pool. This centralized network architecture is known as Cloud Radio Access Network (C-RAN), it was first proposed in [66] and detailed in [67]. C-RAN can be viewed as the natural extension of DAS centralization to a multi-cell scale, its advantages compared to DAS are manifold: Due to the high traffic variation in time and space, individual cell site BBUs are dimensioned according to the network busy hours which may be 10 times higher than off-the-peak hours [67]. When office BSs experience their peak load, residential BSs are at a their low, hence valuable BBU computational power is wasted. By virtualizing BBUs of diverse network areas and enabling dynamic reconfigurable mappings between RRHs and BBUs, the required baseband processing capacity of the pool is smaller than the sum of capacities of BSs taken individually.
The resulting hardware savings from adequate BBU dimensioning is called the statistical multiplexing gain of C-RANs. Although highly dependent on user distribution and traffic profiles, an average gain of 25% can be achieved [68, 69]. These hardware savings directly translate into a reduction of the CAPital EXpenditure (CAPEX) as well as the OPerating EXpenditure (OPEX) since important savings in cooling resources can be achieved, which represent 46% of cell site power consumption [67]. Moreover, even for negligible statistical multiplexing gains, grouping BBUs in the same location reduces site rental/acquisition costs, reducing thereby OPEX/CAPEX. In total, 15% CAPEX and 50% OPEX savings are envisioned in comparison to RAN with RRH [70].
Finally, one of the key features of C-RAN deployments is providing cooperative real-time RRM; therefore, resource allocation becomes possible on a multi-cell level. This provides a robust infrastructure to combat interference as it is the main limiter to network densification as discussed next.

1.2.3 On the Limits of Network Densification and the Cell Paradigm Shift

The limit to how far network densification can go is not necessarily bound to be less or equal to the user deployment density. Provided that idle mode capability is enabled [71, 72], many studies have pushed the ratio of deployed transmission nodes to UEs requiring network access beyond the intuitive unity limit [73–75]. The fundamental limit to network densification lies in the growing interference caused by the decreasing inter-site distance. It was demonstrated in [7] that when the density of small cells grows beyond a certain threshold, the experienced Signal to Interference and Noise Ratio (SINR) decreases as the interfering signals transition from non-LoS (NLoS) to LoS propagation, degrading the network performance. In fact, the problem of interference management is central to

all mobile communications systems. All the proposed Multiple Access (MA) schemes in every mobile generation can be summarized as a proposition to manage the problem of inter-user interference while sharing the same resources. The same is true at the level of traditional cellular architectures where frequency reuse schemes were resorted to for inter-cell interference mitigation. One could argue that inter-cell interference is basically inter-user interference taken on a network scale, but now that the network densification intensifies, the validity of such a distinction may be at question.

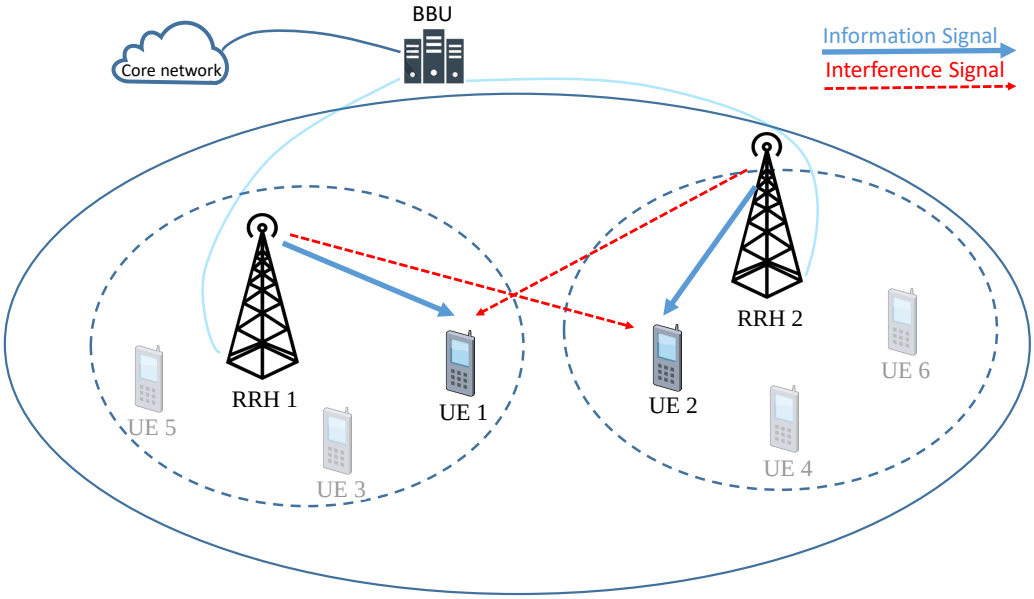


Figure 1.3 – Schematic of an inter-user/inter-cell interference scenario, in a two-antenna cell, illustrating the need for a new broader approach on handling inter-user/inter-cell interference in dense mobile networks.

To better illustrate the duality of the approach on handling interference, we consider the example of figure 1.3, where two users in a DAS cell are served from separate RRHs on the same time/frequency resource block. One possible strategy could be that RRHs 1 and 2 are reusing the cell spectrum. In that case, interference avoidance can be done through the selection of distant users from one another (e.g. UE 6 and UE 5). Another approach could be to consider the system as a two-user NOMA group where interference cancellation techniques could be attempted, i.e. taking advantage of the strong interfering signals for a better cancellation. It is not straightforward to determine whether inter-user interference is better attended to using traditional inter-cell mitigation techniques, or same-cell coordination techniques (which trace back to MA schemes). Each followed approach comes from a different background (inter-cell vs intra-cell interference management), and will lead to the adoption of different policies for the resolution of the same problem. In fact, the cell concept itself is at question as the network densifies, as it becomes more and more challenging to draw the line between neighboring cells. Indeed, the cell concept traces back to the geographic division of the space into hexagonal cells with a central BS serving the users in each region through a dedicated portion of the spectrum. Now that the cell architecture is split into multiple transmitting points with each of them having the potential to reuse the entire spectrum, the common understanding of cells needs to be revisited. As a consequence, the frontier between inter-cell and intra-cell interference management techniques should be revisited in a more holistic manner. In this

order of thought, we present next the concept of CoMP, as the most advanced framework for tackling the interference management problem.

1.3 Coordinated Multipoint

To mitigate the Inter-Cell Intereference (ICI), 3GPP proposed in release 9 [8], and then adopted in release 11 [9], the CoMP technique as the evolution of enhanced ICI Coordination (eICIC) to improve the performance of interference-prone users and enhance the overall network performance. The rationale is to apply coordination between adjacent cells, either in order to alleviate cell-edge interference without restricting the usage of network resources, or to intelligently take advantage of interference. The coordination can be done in a distributed or centralized fashion. In distributed coordination, cell sites are interconnected through the X2 interface to form a fully meshed network where Channel State Information (CSI) is exchanged. Thanks to this configuration, one of the coordinated cells can act as a master cell managing resource allocation and scheduling, while the others act as slaves. Such scenarios are typical for distributed CoMP between small cells [76–78]. In centralized coordination, a central unit processes the feedback information from cell sites to handle ICI and perform joint radio resource scheduling. For this purpose, CSI and user data must be made available at the level of the central unit, which implies high backhaul overhead with stringent latency requirements. For operators with free or cheap already available fiber resources, meeting these backhauling constraints is possible, hence star-like network architectures such as DASs and C-RANs are an appealing solution. In fact, the C-RAN architecture is considered as the main enabler for implementing CoMP technology, since BBU pools are directly interconnected in the same building, thus C-RAN deployment inherently provides the low-latency and high backhaul capacity required for CoMP. A schematic of centralized and decentralized CoMP architectures are provided in Fig. 1.4.

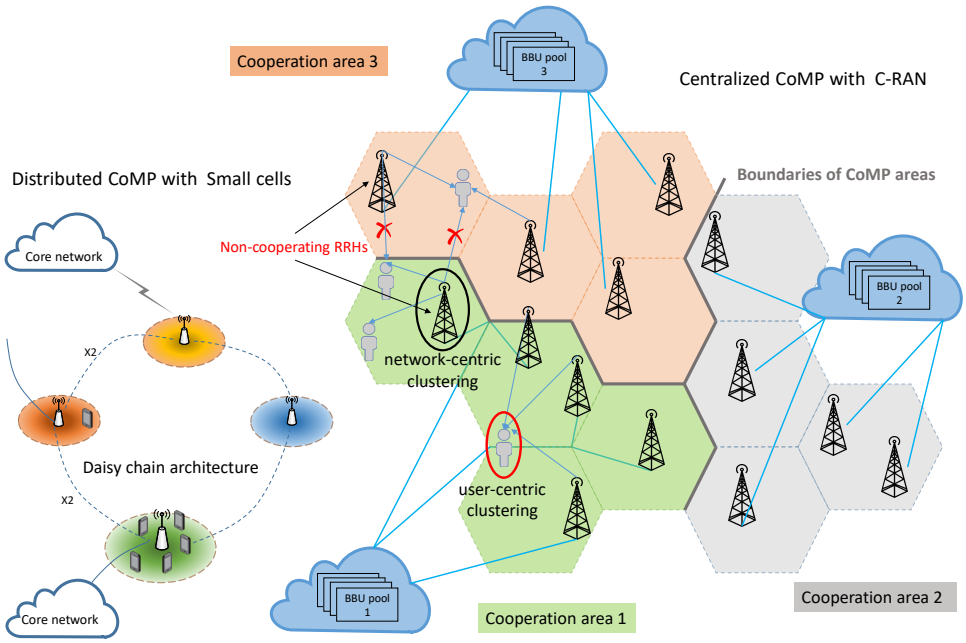


Figure 1.4 – An overview of CoMP implementation into different network architectures.

A common classification of CoMP techniques in the literature distinguishes between the schemes that involve the exchange of CSI only (sometimes referred to as coordination approaches) and those requiring the exchange of both CSI and user data (cooperative approaches). By virtue of this classification, Coordinated Scheduling (CS) and Coordinated Beamforming (CB) are presented first, in the next section, followed by Dynamic Point Selection (DPS) and Joint Transmission (JT) which are cooperative approaches as explained afterwards.

1.3.1 Coordinated Scheduling and Coordinated Beamforming

In CS, the cooperating nodes seek to avoid interference by allocating cell-edge users E1 and F1 different channels f_1 and f_2 (Fig. 1.5), while allocating other frequencies for inner users in the cell (e.g. user E2 in Fig. 1.5). This joint decision on the user-channel association is possible thanks to the sharing of user CSI between the corresponding nodes. Note, however, that the results of a CS coordination are not limited to the interference avoidance policy, but also take into account the potentially competing QoS requirements of both users, the available power at every RRH, the history of user serving, etc. That is to say the coordination results are parts of a whole in the ongoing resource allocation problem to best serve the two cells. These coordination results are applied every time scheduling is performed, which can be as short as 1 ms for LTE. Therefore, resources can be dynamically allocated even with instantaneous changes of UEs channel conditions.

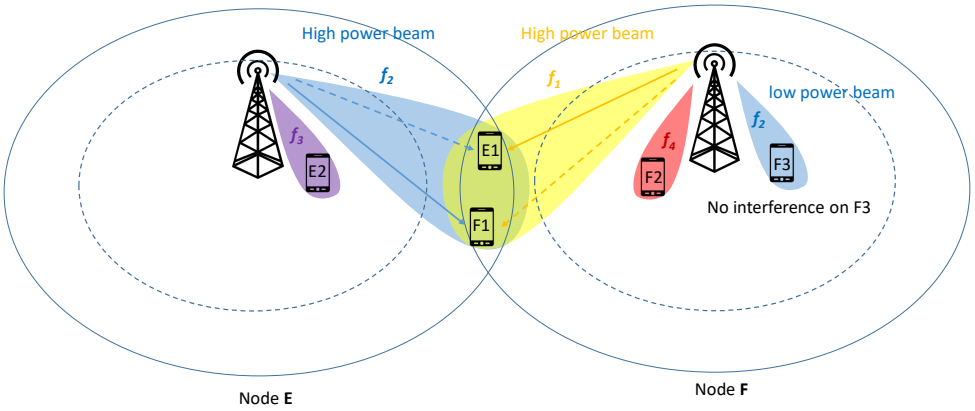


Figure 1.5 – CS, allocating cell edge users different frequency resources.

With CB (Fig. 1.6), users are served through the same time/frequency resource while being allocated different spatial resources, i.e. beam patterns. Thanks to the CSI sharing, which includes channel quality indicators and precoding matrix indicators, interference is prevented through each transmission node allocating the main beam to its user, and nullifying the beam to the other neighboring UE, as shown in Fig. 1.6.

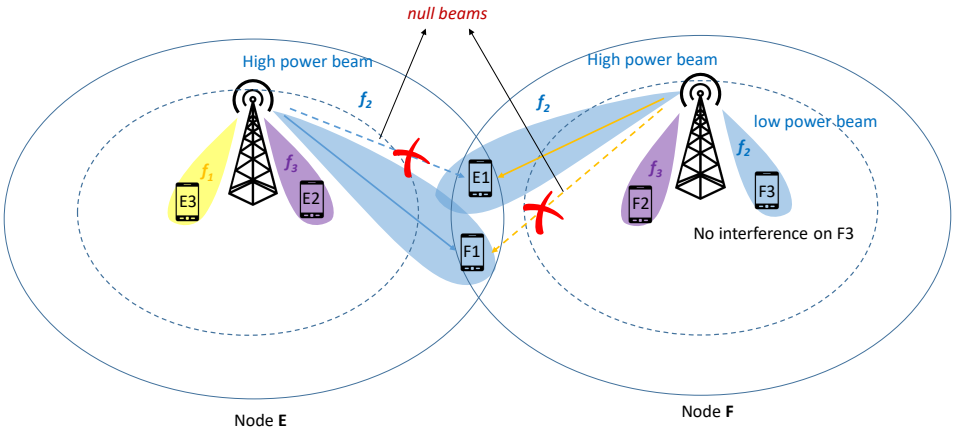


Figure 1.6 – CB, allocating cell edge users different beam patterns while using the same frequency.

Generally, CB is often used with CS, as shown in Fig. 1.7. On the one hand, CS can efficiently handle the interference, and on the other hand, better reception quality is ensured by CB.

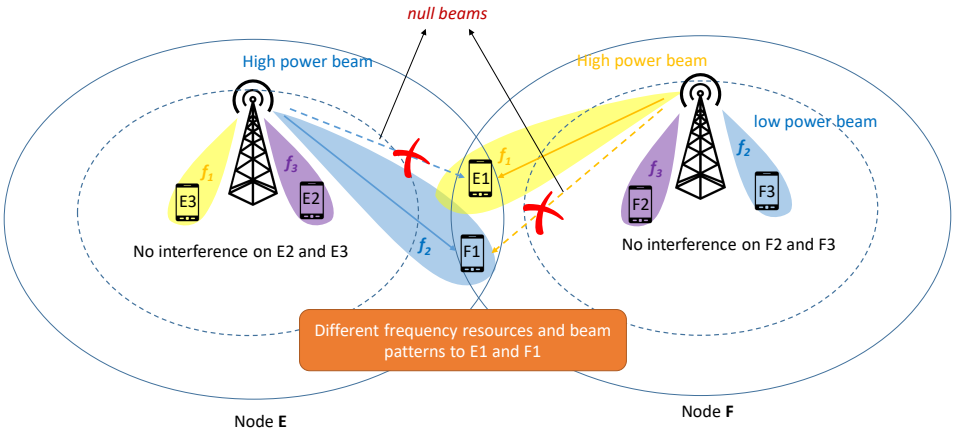


Figure 1.7 – Combining CS/CB schemes.

1.3.2 Dynamic Point Selection and Joint Transmission

In DPS, the data related to a UE is transmitted by a single transmitting node for a given time/frequency resource, as is done in CS. However, on top of CSI, the data should be available at all cooperating transmitters, which enables the selected point to change dynamically from one transmission time interval to another. Therefore, the RRH with the minimum path loss for the UE is always selected. The tighter latency in DPS, compared to CS, enables a cell/TP switching at the subframe level for a given UE.

With Joint Transmission CoMP (JT-CoMP), cooperating TPs transmit simultaneously the signal of the same user over the same time/frequency resource (c.f. Fig. 1.8). The joint processing of the data enables its precoding over the multiple transmitting nodes so that it is coherently combined at the level of each UE. JT-CoMP is the most promising coordination technique, but is also the most challenging one to implement, as discussed next.

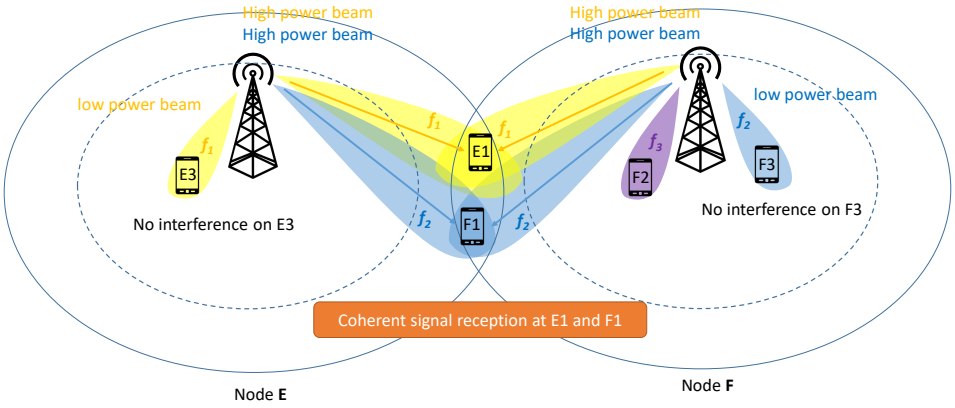


Figure 1.8 – JT Transmission from Nodes E and F to users E1 and F1.

On the evolution of the cell concept in C-RANs and CoMP

JT-CoMP enables the serving of UEs from multiple cell transmission nodes or RRHs, thus breaking to some extent the "cell" paradigm. The common understanding of cells would morph into the concept of CoMP-sets which are in essence the sets of "cell sites" that perform cooperation. The main challenge regarding CoMP is to come up with clustering techniques that would bridge the gap between theoretical expectations and practical performance gains of actual CoMP systems.

This gap is observed when moving from the ideal network wide cooperation area to e.g. two cooperation areas. In that case, the performance of simple clustering techniques can easily fall back to performance levels similar to uncoordinated networks. On the other hand, theoretical and practical results [79, 80] promise a linear performance gain with the increasing cooperation area, while assuming network-wide JT CoMP. Nonetheless, providing network-scale cooperation is simply not feasible - even within the C-RAN architecture - as cooperation will fatally span over geographically separated BBU pools (c.f. Fig. 1.4). Moreover, as the cooperation area increases, inter-cluster interference reduces by virtue of the greater distances, to the point where additionally canceled interferences are comparable to the randomized interference plus noise floor. Consequently, the gains of full network cooperation get asymptotically smaller [81] while the growth rate of the signaling burden due to CSI exchange (let alone user data) is unchanged if not increased. The challenge of CoMP clustering schemes is to strike the right balance between network performance and increasing signaling overhead and scheduling complexity. The physical limitations for providing the huge signaling exchange spanning over multiple BBU pools pushes towards a hybrid clustering scheme. On the one hand, the maximum cooperation

area is static and set to the coverage region of a single BBU pool. Note that thanks to the optical fiber advancements enabling the remote locating of RRHs 20 to 40 km away from the BBUs, considerable cooperation areas can be achieved (c.f. Fig. 1.4). On the other hand, within these fixed areas, dynamic user-centric clustering can be conducted, where users are allocated their own cluster of RRHs which can overlap with each other, instead of clustering RRHs in a network-centric manner and then serving users in the cluster from a subgroup of the RRHs [82, 83].
On a final note, there is an amount of inter-cluster interference which cannot be combated with CoMP; therefore, efficient network planning should consider the traffic profile, varying user densities (e.g. sub-urban vs. dense urban environments), the geographical topology, in order to: 1) optimize the RRH locations and determine the fixed RRH-BBU pools assignments such that the density of cluster-edge users suffering from inter-cluster interference is minimized; 2) elaborate smart CoMP clustering schemes capable of a fast adaption to the network spatial and temporal fluctuations within the available cooperation regions.

1.4 Device to Device Communication

To cope with the increasing demand for data, a completely opposite approach in network densification is to increase the number of wireless links per unit area, instead of deploying more access nodes (small cells or RRHs). The underlying idea is to enable the direct communication between close end-devices instead of having information transiting through BSs and the network core. D2D communication offloads uplink and downlink traffic from the network which can use the freed network capacity and power resources to serve other users. Also, by virtue of single-hop and proximity gains, an efficient D2D channel can be established, leading to high data rates with minimal transmit powers and very low latency. This enhances the system energy efficiency and localizes the interference footprint of devices, enabling densified local reuse of spectrum [10], [11]. Many services can benefit from D2D as depicted in Fig. 1.9, a non-exhaustive list includes: content sharing applications for the exchange of videos and photos between friends, multiplayer gaming, streaming services with enabled caching, mobile relaying for coverage extension, Vehicle-to-Vehicle (V2V) communication requiring strict latency constraints, context-aware applications which enable context-related mobile advertising, etc. In that regard, D2D communications are expected to open up new business opportunities to network operators and app developers to take advantage of this new market by providing proximity-based e-services, as forecast by the social networking service (SNC) research [84].

With D2D, near-by devices are authorized to communicate directly with one another with little to no information transiting through the cellular network. To establish a D2D link, a peer discovery process must be initiated before the communication phase can take place. When direct D2D discovery is used, the D2D communication is completely decentralized without requiring any intervention from the network (e.g. Bluetooth and WiFi-direct). However, direct discovery techniques use beaconing signals and scanning, making them time and energy consuming. Moreover, the uncontrollable interference in the unlicensed spectrum hinders the establishment of reliable QoS. Therefore, moving the D2D discovery process in the licensed band enables resorting to network assistance to mediate the discovery process [14, 32]. The UE initiates a D2D link request, prompting

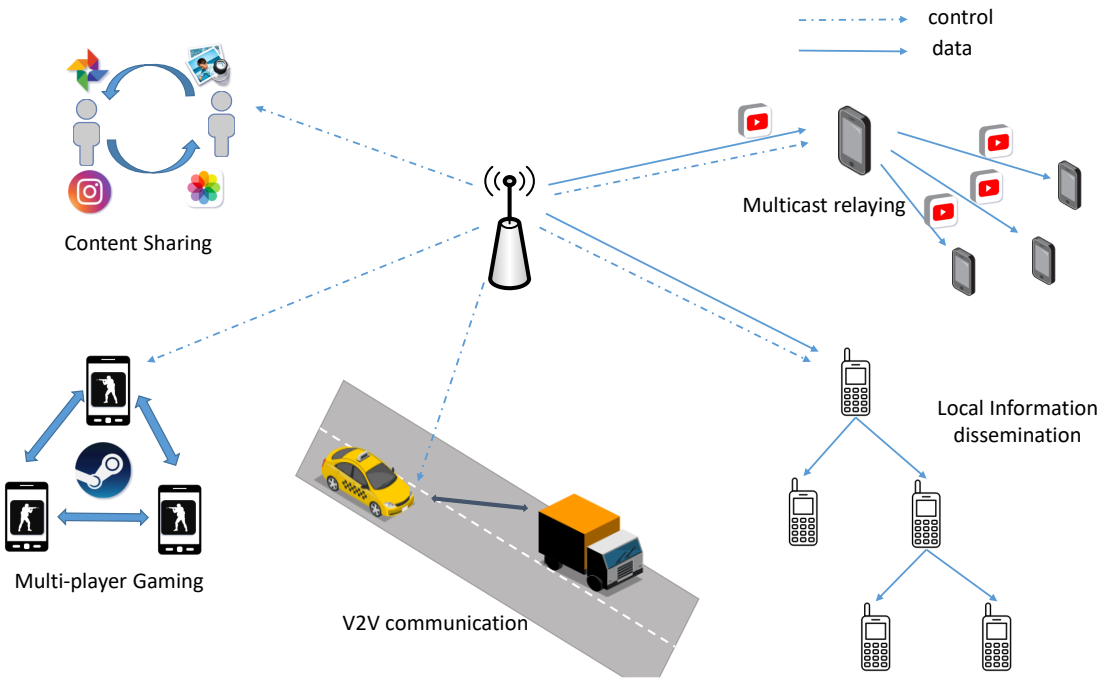


Figure 1.9 – A snapshot of possible D2D applications.

the BS to scan the network and recognize D2D candidates, coordinate the time/frequency resources, and provide back the identity information to the newly formed D2D pair.

In our work on D2D communications (Chapter 6), we are mainly interested in providing efficient power allocation and channel assignment of D2D communications, while assuming prior completion of the D2D discovery and pairing of devices.
As it was hinted out earlier, the idea of D2D is not new, as several services using direct communication already exist such as WiFi-direct and Bluetooth. The novelty in D2D is the utilization of the licensed bands of the cellular spectrum. From that distinction, the following classification can be made regarding D2D communications [12]:

- **Outband D2D communication:** D2D occurs in the unlicensed band without affecting the cellular network.
- **Inband D2D communication:** the D2D channel is allocated from the cellular band. Inband D2D can either be overlay or underlay.
  - **Overlay:** Dedicated communication links from the cellular spectrum are allocated to the D2D, preventing co-channel interference between the D2D system and the cellular network.
  - **Underlay:** In this case, the cellular spectrum is reused by D2D devices and the challenge resides in managing the interference between the D2D and the cellular network.

Due to the stochastic nature of the unlicensed band and to the challenges of coordinating the communication over two different bands (since outband communication requires a second radio interface and uses other wireless technologies such as WiFi Direct [13]), inband transmission has gained much attraction among the research community [14, 15].

Furthermore, due to the anticipated increase in the number of connected devices, dedicating cellular bands to D2D will not be a viable solution, thus most research focus on inband underlay D2D [16–19].

1.4.1 Full Duplex

A highly promising technology to be applied in conjunction with D2D is FD communications. FD enables the same UE (or any other equipment in the network) to transmit and receive information during the same time and using the same frequency [20]. Previous communication schemes either involved a simultaneous transmission and reception but using separate frequencies in the case of FDD (Frequency Division Duplex), or co-channel transmission and reception but using orthogonal time slots for TDD (Time Division Duplex). TDD schemes are also referred to as Half Duplex (HD) in the literature, as they are send-then-receive systems, and FDD can be found under the name of Out-of-band Full-Duplex (OBFD), whereas FD alone refers to In-Band Full-Duplex (IBDF). The achieved gains of FD can go up to a virtual two-fold increase in spectral efficiency compared to HD and OBDF systems. In return, a Self Interference (SI) is incurred due to the transmitted signal looping back into the receiver, thus limiting its appeal compared to HD. The challenge in designing FD equipment is in canceling the SI such that the Residual Self Interference (RSI) is comparable with the noise floor. SI cancellation techniques are grouped into three main categories: passive suppression, analog cancellation and digital cancellation, as depicted in Fig. 1.10.

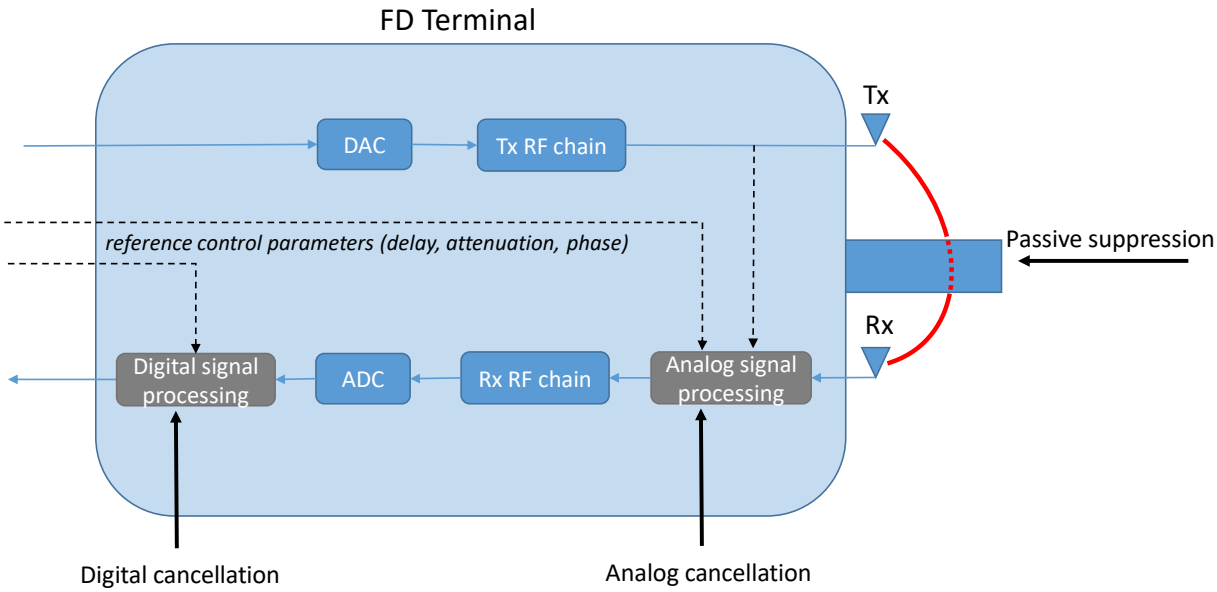


Figure 1.10 – Block diagram of the architecture of an FD transceiver implementing passive suppression, analog and digital self-interference cancellation (ADC = Analog to Digital Converter, DAC = Digital to Analog Converter, RF = Radio Frequency, Tx = Transmitter, Rx = Receiver).

Passive suppression occurs between the Tx and Rx antennas, it mainly consists in attenuating the received signal by separating the antennas the furthest apart on the

equipment, while also placing absorbing material in between them, and eventually applying appropriate polarization. Analog cancellation occurs at the level of the received RF signal before down-conversion. The sent signal is taken from the Tx chain, attenuated and delayed in order to mimic the Tx/Rx channel, and then subtracted from the received RF signal. Finally, digital cancellation occurs at the level of the digital baseband signal after the ADC block. Similarly to the analog cancellation, the necessary phase shifts and delay adjustments are applied to a reference signal from the transmitter in order to further reduce SI. Digital SI cancellation is limited by the dynamic range of ADCs, therefore it is essential to apply all three passive and active cancellation schemes at the FD receiver. Nowadays, the achieved improvement in antenna architecture and in transceiver circuitry allows a great reduction of the RSI [21–23], thereby advocating for the use of FD in future communication standards.

Most of high-level analyses on the capacity gains of FD [24–26] model the RSI as a complex Gaussian random variable with zero mean and variance ηP_{tx}, where η is the SI cancellation capability of the FD device and P_{tx} its transmission power. Thus, the power of the residual self-interference P_{RSI} is given by:

$$P_{RSI} = \eta P_{tx}. \tag{1.3}$$

The cancellation factor η can vary between 0 and 1, with $\eta = 0$ denoting perfect SI cancellation and $\eta = 1$ referring to the case where no cancellation is applied. In our thesis, actual values of η range from -80 dB to -130 dB. Therefore, the RSI is directly related to the transmit signal power, which makes FD most suited for low-power applications like in D2D networks. The surging interest for the combination of the FD communication with the D2D technology gave birth to new D2D applications and scenarios, as depicted in Fig. 1.11.

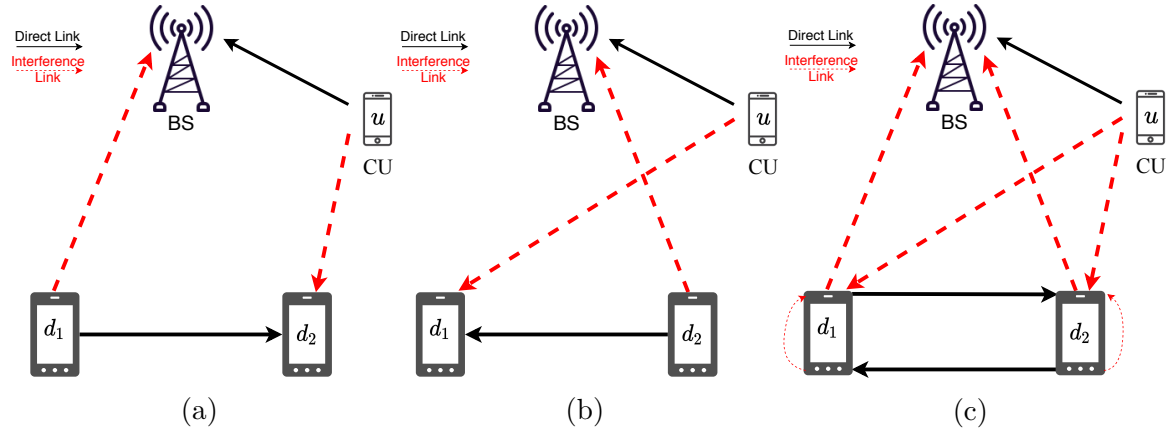


Figure 1.11 – D2D transmission underlaying a cellular system (a) HD transmission, first half time slot, d_1 transmits to d_2. (b) HD transmission, second half time slot, d_2 transmits to d_1. (c) FD transmission, d_1 and d_2 transmit to each other in the same time slot.

In this dissertation, we will be interested in the so-called bidirectional FD-D2D topology presented in Fig. 1.11c. In this use case, a D2D system is underlaying the cellular network. D2D devices are looking to exchange information, hence the bidirectional topology, while also benefiting from the FD technology at the level of both devices d_1 and d_2. In that case, the D2D devices will cause interference on the signal of the cellular

user at the level of the base station, and the cellular user's signal will interfere on both devices. The HD version of this topology is also presented in Fig. 1.11, where in Fig. 1.11a d_1 transmits information to d_2 while d_2 is receiving, and in Fig. 1.11b d_2 transmits information to d_1 while d_1 is receiving.

1.5 Summary

In this chapter, the necessary background on major candidate technologies to fulfill future generation network requirements is provided. First, the ability of power domain NOMA to increase system capacity and the number of connected devices is presented. Then, the key advantages, architectures and limits of network densification in the forms of DAS and C-RAN are elaborated. Afterwards, coordination among cell sites is discussed for the CoMP framework, and its main cooperation/coordination modes are described. Finally, the ecosystem of D2D communication is overviewed showing the gradual transition from unlicensed outband communication and signaling, to licensed bands underlay D2D, with enabled FD communication.

Chapter 2

NOMA Mutual SIC for Power Minimization in Distributed Antenna Systems

In this chapter, we study the problem of serving users by means of a downlink DAS using NOMA. The objective is to minimize the total cell power while guaranteeing the users their required rates. To that end, a power minimization strategy that operates successively in the orthogonal and non-orthogonal layers is proposed. After the presentation of the system model in section 2.2, the principles of the proposed waterfilling algorithm for PA are presented in details, guiding the elaboration of RA strategies for both OMA and NOMA (sections 2.4 and 2.5). Also, the combination of DAS and NOMA paves the way for a mutual SIC procedure, whose theoretical background is developed in section 2.5.2.1. Then, its incorporation to the global RA procedure is conducted for various PA schemes in sections 2.5.2.3 and 2.5.2.4. A complexity analysis of the proposed schemes is performed in section 2.6, and the performance of the proposed methods is evaluated in section 2.7. The conclusions are finally drawn in section 2.8.

The major contributions of this chapter are summarized as follows:

- We introduce several techniques that allow a significant complexity reduction of the waterfilling procedures used for PA in [85], for both orthogonal and non-orthogonal transmission, while adapting the allocation techniques to the DAS context.

- We propose a new NOMA PA scheme for user pairing that outperforms Fractional Transmit Power Allocation (FTPA) [27, 28], while taking into account the power multiplexing constraints.

- Unlike previous works, we investigate the use of different RRHs to power the multiplexed subcarriers in NOMA. This new setting gives rise to the concept of mutual SIC where paired users on a subcarrier can perform SIC at the same time, under well defined conditions.

- Finally, we propose new suboptimal algorithms to achieve joint subcarrier, RRH, and power allocation, in light of the newly uncovered potentials specific to the application of NOMA in the DAS context.

2.1 Related Works

Efficient RA is key in squeezing the achievable potentials out of DAS. For this purpose, the study in [5] explored the advantages of DAS and compared the achievable ergodic capacity for two different transmission scenarios: selection diversity and blanket transmission. In the first case, one of the RRHs is selected (based on a path-loss minimization criterion) for transmitting a given signal, whereas in the second, all antennas in the cell participate in each transmission, thus creating a macroscopic multiple antenna system. The results of [5] show that selection diversity achieves a better capacity in the DAS context, compared to blanket transmission. The same observations are made in [6]. In [86], RRH selection is also preconized as a mean to decrease the number of information streams that need to be assembled from or conveyed to the involved RRHs, as well as the signaling overhead.

2.1.1 Energy Efficiency Maximization in DAS

Several works target the optimization of system Energy Efficiency (EE) in DAS. In [87], two antenna selection techniques are proposed, either based on user path-loss information or on RRH energy consumption. Also, proportional fairness scheduling is considered for subband allocation with a utility function adapted to optimize the EE. In [88], Subcarrier Assignment (SA) and PA are done in two separate stages. In the first one, the number of subcarriers per RRH is determined, and subcarrier-RRH assignment is performed assuming initial equal power distribution. In the second stage, PA is performed by maximizing the EE under the constraints of the total transmit power per RRH, of the targeted bit error rate and of a proportionally-fair throughput distribution among active users. The optimization techniques proposed in [87,88] for DAS are designed for the orthogonal case. In other words, they allow the allocation of only one user per subcarrier.

2.1.2 NOMA in DAS and C-RAN

Applying power multiplexing on top of the orthogonal frequency division multiplexing (OFDM) layer has proven to significantly increase system throughput compared to orthogonal signaling, while also improving fairness and cell-edge user experience. A few previous works have studied the application of NOMA in the DAS context. An outage probability analysis for the case of two users in C-RAN is provided in [89] where all RRHs serve simultaneously both users. The results show the superiority of NOMA when compared to TDMA, in the context of C-RANs. In [90], the study investigates the application of distributed NOMA for the uplink of C-RANs. The partially centralized C-RAN architecture allows the use of joint processing by distributed antennas, in which RRHs can exchange correctly decoded messages from other RRHs in order to perform SIC. In [91], an efficient end-to-end uplink transmission scheme is proposed where the wireless link between users and RRHs on one side, and the fronthaul links between the RRHs and BBU on the other side are studied. User grouping on blocks of subcarriers is proposed to mitigate the computational complexity, and a fronthaul adaptation for every user group is performed in order to strike a tradeoff between throughput and fronthaul usage.

2.1.3 State of the Art of Power Minimization in the NOMA Context

Recent works tackle the downlink power minimization problem in the NOMA context. In [92], the proposed joint RA scheme consists in a deletion-based algorithm where the entire spectrum is first allocated to all users; then, optimal PA followed by the removal of users from subcarriers are iteratively conducted until the constraints of the maximum number of multiplexed users are satisfied. The algorithm presents near-optimal results, however, it proceeds with a high computational complexity as a numerical solver is required for solving the optimal PA in every iteration. Moreover, the PMCs are not taken into consideration. The PMCs state that the signal to be decoded first must have a higher power level than the other received signals, so that it is detectable at the receiver side. A similar deletion-based approach to [92] is followed in [29] but with consideration of PMCs. First the entire spectrum is allocated to all users and the optimal PA is obtained for a relaxed version of the minimization problem without PMCs. Then, the number of multiplexed users per subcarrier is reduced to a maximum of two (according to a simple criterion), before the iterative adjustment phase is conducted serially over all the users to meet the PMCs using bisection search. However, the proposed adjustment procedure does not take into account the rate coupling between multiplexed users. Thus, the obtained solution satisfies PMCs but without a guarantee of user rate satisfaction. Power minimization strategies are also proposed in [93] for Multiple-Input Multiple-Output NOMA (MIMO-NOMA), where PA and receive beamforming design are alternated in an iterative way. Constraints on the targeted Signal to Interference and Noise Ratio (SINR) are considered to guarantee successful SIC decoding. Provided results for a moderate number of users (4 or 6) show an important gain in performance with respect to OMA, however the subcarrier allocation problem is not included, only PA is considered. In [85], a set of techniques have been introduced, allowing the joint allocation of subcarriers and power, with the aim of minimizing the total power in NOMA-CAS. Particularly, it was shown that the most efficient method, from the power minimization perspective, consists of applying user pairing at a subsequent stage to single-user assignment, i.e. after applying OMA signaling at the first stage, instead of jointly assigning collocated users to subcarriers. The work in this chapter follows the same approach to perform power minimization.

2.2 System Model

The system consists in a total of R RRHs uniformly positioned over a cell where K mobile users are randomly deployed (Fig. 2.1). The RRHs are connected to the BBU through high capacity optical fibers. RRHs and users are assumed to be equipped with a single antenna. Users transmit their CSI to RRHs, and the BBU collects all the CSI from RRHs. Perfect CSI is assumed throughout the thesis (the influence of imperfect or outdated CSI is not the aim of our work). Alternatively, the BBU can benefit from channel reciprocity to perform the downlink channel estimation by exploiting the uplink transmissions. Based on these estimations, the BBU allocates subcarriers, powers, and RRHs to users in such a way to guarantee a transmission rate of $R_{k,req}$ [bps] for each user k. The system bandwidth B is equally divided into S subcarriers to form the set $\mathcal{S} = [\![1 \dots S]\!]$. Each user k is allocated a set $\mathcal{S}_k$ of subcarriers. From the set of K users, a maximum of $m(n)$ users $\{k_1(n), k_2(n), \dots, k_{m(n)}(n)\}$ are chosen to be collocated on the n^{th} subcarrier

$(1 \leq n \leq S)$, where $k_i(n)$ refers to the i^{th} user multiplexed on subcarrier n with $r_i(n)$ its powering antenna. Classical OMA signaling corresponds to the special case of $m(n) = 1$. Let $h_{k,n,r}$ be the squared channel gain between user k and RRH r over subcarrier n, and $\mathbb{H}$ the three-dimensional channel gain matrix with elements $h_{k,n,r}$, $1 \leq k \leq K$, $1 \leq n \leq S$, $1 \leq r \leq R$. As shown in Fig. 2.1, NOMA subcarriers can be served by the same RRH or by different RRHs. For instance, one can consider serving User 1 and User 2 on the same subcarrier SC 1 $(n = 1, k_1(1) = 1, k_2(1) = 2)$ by RRH 1 $(r_1(1) = 1, r_2(1) = 1)$, while User 2 and User 3 are paired on SC 2 $(n = 2, k_1(2) = 3, k_2(2) = 2)$, and served by RRH 1 and RRH 2 respectively $(r_1(2) = 2, r_2(2) = 1)$.

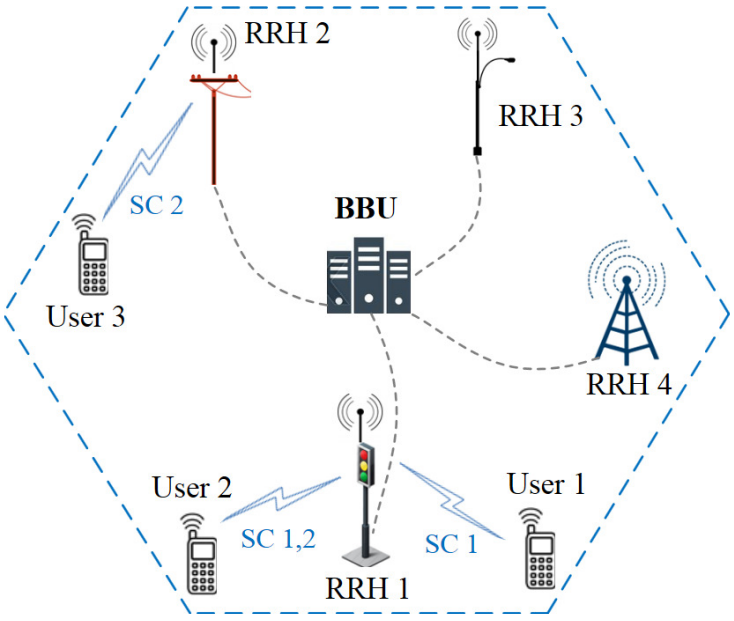


Figure 2.1 – Example of a downlink DAS setup with four RRHs and three NOMA-served users.

In the rest of the chapter, and without loss of generality, we will consider a maximum number of collocated users per subcarrier of 2, i.e. $m(n)$= 1 or 2. On the one hand, it has been shown that the gain in performance obtained with the collocation of 3 users per subcarrier, compared to 2, is minor in downlink NOMA [27]. On the other hand, limiting the number of multiplexed users per subcarrier limits the SIC complexity at the receiver terminals. We will denote by first (resp. second) user on a subcarrier n the user which has a higher (resp. lower) channel gain on n between the two paired users, when their subcarrier is powered by the same RRH. Let $P_{k_i(n),n,r_i(n)}$ be the power of the i^{th} user on subcarrier n transmitted by RRH $r_i(n)$. When the same antenna powers the signals of multiplexed users over a subcarrier n $(r_1(n) = r_2(n) = r)$, user $k_1(n)$ can remove the inter-user interference from any other user $k_2(n)$ if its channel gain verifies $h_{k_2(n),n,r} < h_{k_1(n),n,r}$ [27,63], and treats the received signals from other users as noise. The theoretical throughputs $R_{k_i(n),n,r}, 1 \leq i \leq 2$, on n are given by the Shannon capacity limit as follows:

$$R_{k_1(n),n,r} = \frac{B}{S} \log_2 \left(1 + \frac{P_{k_1(n),n,r} h_{k_1(n),n,r}}{\sigma^2}\right), \tag{2.1}$$

$$R_{k_2(n),n,r} = \frac{B}{S} \log_2 \left(1 + \frac{P_{k_2(n),n,r} h_{k_2(n),n,r}}{P_{k_1(n),n,r} h_{k_2(n),n,r} + \sigma^2}\right), \tag{2.2}$$

where N_0 and $\sigma^2 = N_0 B/S$ are respectively the power spectral density and the power level (over a subcarrier) of additive white Gaussian noise, including randomized inter-cell interference, and assumed to be constant over all subcarriers.

2.3 Problem Formulation

We first consider the case where the same RRH powers the signals of both paired users on each subcarrier; the case of two different RRHs powering the multiplexed signals is treated separately in section 2.5.2. Taking into account the PMCs specific to NOMA, the corresponding optimization problem can be formulated as:

$$\{\mathcal{S}_k, P_{k,n,r}\}^* = \underset{\mathcal{S}_k, P_{k,n,r}}{\arg\min} \sum_{k=1}^{K} \sum_{n \in \mathcal{S}_k} \sum_{\substack{i=1, \\ \text{s.t.} k_i(n)=k}}^{2} P_{k,n,r_i(n)}, \tag{2.3}$$

subject to:

$$\begin{cases} \displaystyle\sum_{n \in \mathcal{S}_k} \sum_{\substack{i=1, \\ \text{s.t.} k_i(n)=k}}^{2} R_{k,n,r_i(n)} = R_{k,req}, \forall k \in [\![1 \dots K]\!], & (2.4) \\ P_{k,n,r} \geq 0, \forall (k,n,r) \in [\![1 \dots K]\!] 55 \times [\![1 \dots S]\!] \times [\![1 \dots R]\!], & (2.5) \\ P_{k_2(n),n,r} \geq P_{k_1(n),n,r}, \forall n \in \mathcal{S}/m(n) = 2. & (2.6) \end{cases}$$

The problem consists in finding the optimal subcarrier-RRH-user allocation, as well as the optimal PA over the allocated subcarriers, so as to minimize the objective function, that is the total transmit power of the cell. This must be done under the rate constraints (2.4), positive power constraints (2.5), and PMCs (2.6). The first constraint imposes a minimum rate requirement $R_{k,req}$ for every user k, that must be achieved over the subcarriers $\mathcal{S}_k$ allocated to k. The second condition ensures that all power variables remain non-negative (a null power variable corresponds to an unallocated subcarrier). Finally, the last constraint accounts for the power multiplexing conditions where the power $P_{k_2(n),n,r}$ of the weak user must be greater than the power $P_{k_1(n),n,r}$ of the strong user. Solving this optimization problem resides in determining the optimal allocation set $\mathcal{S}_k$ for every user k, as well as finding the optimal power allocation over the allocated subcarriers. Therefore, the optimization problem at hand is mixed combinatorial and non-convex, which justifies the introduction of suboptimal solutions. However, instead of completely splitting the subcarrier assignment from the power allocation, we aim at designing a power minimization algorithm that iteratively performs user-subcarrier-RRH assignment based on the estimation of the system power for a given iteration. This approach proved its efficiency in [85] for the CAS context and will be used next. Also, the joint subcarrier assignment and power allocation strategy is operated in an initial OMA phase (section 2.4), followed by the additional NOMA layer for the case of the same RRH powering multiplexed signals in section 2.5, and different RRHs powering the multiplexed signals in section 2.5.2.

2.4 Power Minimization in OMA Signaling

The problem in (2.3) is NP-hard [94, 95] even for the OMA case, and its solution resides in finding the optimal subcarrier assignment which consists in a subcarrier-user-RRH allocation ($\mathcal{S}_k^* \; \forall k$), and the optimal PA ($P_{k,n,r}^*$) corresponding to that SA. This being said, for any fixed SA (including the optimal one $\mathcal{S}_k^*$), the optimal PA for power minimization in OMA is the well known waterfilling algorithm [96]. Therefore, we start by presenting the properties of the waterfilling algorithm in details, then the gained insights allow the design of an efficient joint channel and power allocation scheme.

2.4.1 Optimal PA: The Waterfilling Algorithm

In the orthogonal context, there is no inter-user interference hence the global downlink power minimization problem reduces to K separate power minimization problems, one for every user in the cell.

Consider a user k allocated the subcarrier set $\mathcal{S}_k$ of size N_k with the subcarrier n of k being transmitted by antenna $r(n)$; the waterfilling procedure for minimizing the total user power while meeting its required rate is the solution to the problem:

$$\{P_{k,n,r}\}^* = \arg\min_{P_{k,n,r}} \sum_{n\in\mathcal{S}_k} P_{k,n,r(n)}, \tag{2.7}$$

subject to:

$$\sum_{n\in\mathcal{S}_k} \frac{B}{S} \log_2\left(1 + \frac{P_{k,n,r(n)} h_{k,n,r(n)}}{\sigma^2}\right) = R_{k,req}. \tag{2.8}$$

This problem is efficiently solved by means of standard optimization techniques, its Lagrangian is given by:

$$L(P_{k,n,r}, \lambda) = -\sum_{n\in\mathcal{S}_k} P_{k,n,r(n)} + \lambda(R_{k,req} - \sum_{n\in\mathcal{S}_k} \frac{B}{S} \log_2\left(1 + \frac{P_{k,n,r(n)} h_{k,n,r(n)}}{\sigma^2}\right)).$$

The Karush-Kuhn-Tucker (KKT) condition for achieving optimality is given by:

$$\begin{aligned} &\partial L/\partial P_{k,n,r(n)} = 0, \forall n \in \mathcal{S}_k, \\ \Leftrightarrow\; &-1 + \frac{B}{S\ln(2)} \frac{\lambda h_{k,n,r(n)}}{\sigma^2 + P_{k,n,r(n)} h_{h,n,r(n)}} = 0, \forall n \in \mathcal{S}_k, \\ \Leftrightarrow\; &P_{k,n,r(n)} + \frac{\sigma^2}{h_{k,n,r(n)}} = \frac{\lambda B}{S\ln(2)} = \text{constant}, \forall n \in \mathcal{S}_k. \end{aligned} \tag{2.9}$$

The solution is called the waterfilling solution because one can construe the solution graphically by thinking of the curve of inverted channel signal-to-noise ratios as being filled with energy (water) to a constant line given by $W_k \triangleq \lambda B/S\ln(2)$, with more power being allocated to high gain subcarriers. The waterline is determined by replacing (2.9)

in (2.8) which yields:

$$\sum_{n\in\mathcal{S}_k} \frac{B}{S}\log_2\left(\frac{W_k h_{k,n,r(n)}}{\sigma^2}\right) = R_{k,req},$$

$$W_k = \sigma^2 \frac{2^{\frac{SR_{k,req}}{BN_k}}}{\left(\prod\limits_{n\in\mathcal{S}_k} h_{k,n,r(n)}\right)^{1/N_k}}. \tag{2.10}$$

The waterline is proportional to the background noise, inversely proportional to the geometrical mean of the channel gains ($\triangleq \overline{h}$), and grows exponentially with the required rate. Note the impact of the allocated bandwidth BN_k in reducing the waterline, hence the total power: even if a subcarrier addition does not enhance the geometrical mean, a waterlevel drop still occurs due to the increment of N_k. However, if the channel mean is affected by the change in $\mathcal{S}_k$, the waterline variation cannot be predicted beforehand. Therefore, the evolution of the waterline and user power with subcarrier addition and retraction is studied next to gain insights on SA for OMA.

2.4.1.1 Subcarrier Addition

A subcarrier is added to $\mathcal{S}_k$ if its addition decreases the user power through waterfilling. Since subcarrier addition presents an additional power burden, the only way this may lead to a power decrease is through a decrease of the waterline. We first evaluate the condition on h_{new} (the gain of the added subcarrier) for a waterline decrease, then we show that a waterline decrease does indeed translate into a power decrease before defining the way a subcarrier should be selected to maximize that power decrease.

From (2.10), an iterative relation is derived between the old waterline $W_k(N_k)$, the new waterline $W_k(N_k+1)$ and the channel gain of the added subcarrier h_{new}:

$$W_k(N_k+1)^{N_k+1} = \frac{W_k(N_k)^{N_k}}{h_{new}/\sigma^2}. \tag{2.11}$$

To compare the new waterlevel to the previous one, we compute their ratio:

$$\frac{W_k(N_k+1)}{W_k(N_k)} = \frac{W_k(N_k)^{\frac{N_k}{N_k+1}}}{(h_{new}/\sigma^2)^{\frac{1}{N_k+1}} W_k(N_k)},$$

$$= \left(\frac{\sigma^2/h_{new}}{W_k(N_k)}\right)^{\frac{1}{N_k+1}}.$$

The waterline decreases when the added subcarrier verifies:

$$h_{new} > \frac{\sigma^2}{W_k(N_k)} \Leftrightarrow h_{new} > \frac{\overline{h}}{2^{\frac{SR_{k,req}}{BN_k}}}. \tag{2.12}$$

This relation provides the precise condition on the link quality for subcarrier admission to $\mathcal{S}_k$. Indeed, not only is the previous observation confirmed regarding waterline decrease given that $\overline{h}$ is unchanged, but also it is shown that an added subcarrier can decrease the waterline even though the average channel gain is degraded. We prove next that adding a subcarrier verifying (2.12) leads to a power decrease which is maximal when h_{new} is at its highest.

Proof. The expression of the total user power before and after subcarrier addition is given respectively by:

$$P_{k,tot}(N_k) = N_k W_k(N_k) - \sum_{n\in\mathcal{S}_k} \frac{\sigma^2}{h_{k,n,r(n)}},$$
$$P_{k,tot}(N_k+1) = (N_k+1)W_k(N_k+1) - \sum_{n\in\mathcal{S}_k} \frac{\sigma^2}{h_{k,n,r(n)}} - \frac{\sigma^2}{h_{new}}.$$

Expressing the power variation in terms of $W_k(N_k), N_k$ and h_{new}, we get:

$$\Delta P = P_{k,tot}(N_k+1) - P_{k,tot}(N_k) = (N_k+1)\left(\frac{W_k(N_k)^{N_k}}{h_{new}/\sigma^2}\right)^{\frac{1}{N_k+1}} - N_k W_k(N_k) - \frac{\sigma^2}{h_{new}}. \quad (2.13)$$

Taking the derivative of ΔP with respect to h_{new}, we get:

$$\frac{\partial(\Delta P)}{\partial h_{new}} = -\frac{\sigma^{\frac{2}{N_k+1}}}{h_{new}^{\frac{1}{N_k+1}+1}} W_k(N_k)^{\frac{N_k}{N_k+1}} + \frac{\sigma^2}{h_{new}^2},$$
$$= \Big[-\left(\frac{\sigma^2}{h_{new}} W_k(N_k)^{N_k}\right)^{\frac{1}{N_k+1}} + \frac{\sigma^2}{h_{new}} \Big] \frac{1}{h_{new}}.$$

Studying the negativity of $\partial(\Delta P)/\partial h_{new}$ we get:

$$\frac{\sigma^2}{h_{new}} \leq (\frac{\sigma^2}{h_{new}})^{\frac{1}{N_k+1}} W_k(N_k)^{\frac{N_k}{N_k+1}},$$
$$\Leftrightarrow (\frac{\sigma^2}{h_{new}})^{1-\frac{1}{N_k+1}} \leq W_k(N_k)^{\frac{N_k}{N_k+1}},$$
$$\Leftrightarrow (\frac{\sigma^2}{h_{new}})^{\frac{N_k}{N_k+1}} \leq W_k(N_k)^{\frac{N_k}{N_k+1}},$$
$$\Leftrightarrow \frac{\sigma^2}{h_{new}} \leq W_k(N_k).$$

Which is the same condition as in (2.12). Therefore, since a prerequisite of the selected subcarrier is to verify (2.12), the derivative of ΔP is negative and the greater is h_{new}, the smaller is the power decrease (in algebraic value). Note that if (2.12) is met with equality, then from (2.11), $W_k(N_k+1) = W_k(N_k)$ and the power variation is null. Thus, for $h_{new} > \sigma^2/W_k(N_k)$, the subcarrier addition yields a power decrease. This concludes our proof. □

As a conclusion, the subcarrier addition verifying (2.12) leads to a waterline decrease and a power decrease. The decrease is maximized when the selected subcarrier is such that h_{new} is as high as possible.

2.4.1.2 Subcarrier Removal

So far, it has been implicitly assumed that all the subcarriers in $\mathcal{S}_k$ receive positive powers through waterfilling. Looking back at (2.9), this is not guaranteed since the waterlevel may be low enough so that negative powers are allocated to subcarriers. This occurs for

every subcarrier $n \in \mathcal{S}_k$ with $h_{k,n,r(n)} < \sigma^2/W_k(N_k)$, the opposite of condition (2.12) for adding a subcarrier. From hereinafter, a *useful* or *valid* subcarrier is one that verifies (2.12), otherwise it is *useless*. In the literature [97], such subcarriers are dealt with by invoking a subcarrier removal routine: the subcarriers are first sorted according to their channel gains, then the subcarrier receiving the most negative power (i.e. having the lowest gain subcarrier) is removed from $\mathcal{S}_k$. The waterline is updated and the search for negative powers is repeated until useful-only subcarriers remain in $\mathcal{S}_k$. Since adding a useful subcarrier decreases the waterline, it can be easily shown that removing a useless subcarrier also decreases the waterline. Therefore, we propose to slightly modify the routine by removing all the useless subcarriers at once instead of removing them one at a time.

On another hand, if a change of ΔR in the required rate of a user is observed, based on (2.10), the new waterline is obtained from the previous one through:

$$W_{k,new} = W_k 2^{\frac{S\Delta R}{BN_k}}. \tag{2.14}$$

This straightforward relation between the new and old waterlines provides a comprehensive complexity reduction compared to the dichotomy-based waterfilling approach used in [85]. Note that an increase in the required rate does not cause any negative power issues because the waterline increases, making (2.12) easier to satisfy, whereas a negative ΔR might cause problems. In such a case, the new waterline is obtained from (2.14), then the subcarrier removal routine explained previously is executed. If this does not lead to a subcarrier removal, the power decrease is given by:

$$\Delta P = N_k(W_{k,new} - W_k) = W_k(2^{\frac{S\Delta R}{BN_k}} - 1). \tag{2.15}$$

Having presented the behavior of the user waterline and total power for the addition and removal of a subcarrier in terms of the user subcarrier set $\mathcal{S}_k$ and the channel gain quality of the candidate subcarrier, we are now equipped to tackle the problem of joint SA and PA in the next section.

2.4.2 Joint Subcarrier Assignment and Power Allocation in OMA

The determining parameters for the total user power are its required rate, the quality of the mean channel gain on its allocated subcarriers ($\overline{h}$), and the number of allocated subcarriers. In the system power minimization problem, the user required rates are given, and the joint SA and PA is all about sharing the system bandwidth among users and conducting the adequate user-subcarrier association to minimize the global system power. Our proposed complete joint SA and PA procedure for OMA is referred to as OMA-DAS; it resides in an iterative user-subcarrier-RRH allocation with a power update after each allocation. The algorithm is composed of an initialization phase and an algorithm core, as shown next.

Worst-Best-H: WBH

Given the importance of the best subcarriers of a user in reducing its power, the initialization phase must make sure that users which can potentially consume the most power get there best subcarriers first. Considering $h_{k,max}$, the best channel gain of k over all

the subcarrier-RRH pairs in the system, the selected user is: $k^* = \arg\min_k h_{k,max}$. It is allocated its best subcarrier-RRH pair which is loaded in power to provide $R_{k^*,req}$. This initialization step is repeated until each user is assigned one subcarrier, after which the priority is changed, as will be explained next. These steps are shown in details in algorithm 2.1, where $\mathcal{S}_f$ is the set of allocated subcarriers, $\mathcal{S}_p$ the set of free subcarriers, and $\mathcal{U}_0$ the set of an uninitialized users.

Algorithm 2.1 WBH

Initialization: $\mathcal{S}_p = [1:S], \mathcal{U}_0 = [1:K], \mathcal{S}_f = \emptyset$
while $\mathcal{U}_0 \neq \emptyset$ **do**
$\forall k \in \mathcal{U}_0 : (n\_\max_k, r\_\max_k) = \arg\max_{n \in \mathcal{S}_p, r \in [\![1 \ldots R]\!]} (h_{k,n,r})$
$k^* = \arg\min_{k \in \mathcal{U}_0} h_{k,n\_\max_k, r\_\max_k}$
$n^* = n\_\max_{k^*}; r^* = r\_\max_{k^*}$
$P_{k^*,n^*,r^*} = \sigma^2 (2^{R_{k^*,req} S/B} - 1)/h_{k^*,n^*,r^*}$
$P_{k^*,tot} = P_{k^*,n^*,r^*}$,
$\mathcal{S}_{k^*} = \mathcal{S}_{k^*} \cup \{n^*\}$
$\mathcal{S}_f = \mathcal{S}_f \cup \{n^*\}$
$\mathcal{S}_p = \mathcal{S}_p \cap \{n^*\}^c$
$\mathcal{U}_0 = \mathcal{U}_0 \cap \{k^*\}^c$
end while

Orthogonal multiplexing

After the WBH phase, the system power consumption is at its highest. The core of OMA-DAS resides in an iterative subcarrier allocation phase where the system power is reduced after each subcarrier allocation. To efficiently allocate the bandwidth among users and thus minimize the system power, the most power consuming users should be prioritized, i.e. the users that request the highest total transmit power from the antennas. Therefore, the subcarrier allocation phase resides in selecting the most power consuming user which is then allocated the best subcarrier-RRH pair available as it reduces the most its power consumption (section 2.4.1.1). Following that allocation, the power of the selected user is updated through (2.13), updating thereby the user priority for subsequent subcarrier allocations. The process is repeated until the system power decrease becomes negligible or until the allocation of all the subcarriers. Note that after the WBH phase, the system power decreases with every subcarrier allocation by at least ρ. The threshold ρ is chosen in such a way to strike a balance between the power efficiency and the spectral efficiency of the system, since unused subcarriers are released for use by other users or systems. The complete OMA-DAS RA scheme is presented in algorithm 2.2.

Remark. *For each user, subcarriers are allocated in the descending order of channel gains. Since an allocated subcarrier is guaranteed to be useful (c.f.* (2.12)*), previously allocated subcarriers with higher channel gains than* $h_{new} \triangleq h_{k^*,n^*,r^*}$ *are also valid after updating the power subsequently to the allocation of* h_{new}*. Therefore, no negative powers arise from the resulting waterline decrease.*

Next, the NOMA user pairing phase is considered; it is applied on top of the OMA layer, that is after the OMA-DAS algorithm.

Algorithm 2.2 OMA-DAS

Phase 1: WBH
Phase 2: Orthogonal multiplexing
$\mathcal{U}_p = [1:K]$// set of active users
while $\mathcal{U}_p \neq \emptyset \wedge \mathcal{S}_p \neq \emptyset$
  $k^* = \arg\max_k P_{k,tot}$// identify the most power-consuming user
  $(n^*, r^*) = \arg\max_{(n,r)} h_{k^*,n,r}$ // identify its most favorable subcarrier-RRH pair
  if $h_{k^*,n^*,r^*} > \sigma^2 / W_{k^*}(N_{k^*})$
    Calculate $W_{k^*}(N_{k^*}+1), \Delta P(k^*, n^*, r^*)$ using (2.11) and (2.13)
    if $\Delta P(k^*, n^*, r^*) < -\rho$ // (n^*, r^*) allows a significant power decrease
      Attribute (n^*, r^*) to k^*,
      Remove n^* from $\mathcal{S}_p$,
      Add n^* to $\mathcal{S}_f$,
      Update $P_{k^*,tot}$
    else
      remove k^* from $\mathcal{U}_p$ // k^*'s power can no longer be decreased significantly in OMA
    end if
  else
      remove k^* from $\mathcal{U}_p$ // k^*'s power can no longer be decreased at all in OMA
  end if
end while

2.5 Power Minimization in NOMA Signaling

The NOMA layer consists in an iterative user pairing phase to further decrease the system power after OMA-DAS. The general idea behind user pairing in the NOMA RA schemes we develop is to select the most power consuming users and pair them onto the subcarriers that reduce their total power the most. The followed PA strategy for user pairing and the reasons and mechanisms for a power decrease subsequent to a user pairing are described next in details.

Given a selected user k for pairing, the multiplexing over a candidate subcarrier is considered only for subcarriers in $\mathcal{S}_f$ but not in $\mathcal{S}_k$. Thus the pairing is seen by the user as a further bandwidth allocation as in the OMA phase, hence the possibility of a power decrease, with the exception that additional interference is present on the candidate channel due to the power of the first user that was initially allocated this candidate in the OMA phase. Therefore, user k is allocated the subcarrier as a second user on subcarrier n ($k = k_2(n) = k_2$). By doing so, the rate achieved by the first user already allocated on n is not jeopardized. The resulting power from that bandwidth allocation is handled as follows:
When allocating a subcarrier n to user k_2, the additional rate brought to the user ($R_{k_2,n,r}$ in (2.2)) must be compensated for by lowering the rate on the sole subcarriers of k_2 (denoted as $\mathcal{S}_{k_2}^{sole}$) to prevent any rate excess. The sole subcarriers of a user are the ones that did not get paired in a prior pairing phase, neither as first, nor as second users. The waterline is decreased only over these subcarriers in order to avoid long chains of power modifications that would arise by changing the power of previously multiplexed subcarriers. Such changes would in turn induce power modifications on users paired

with k_2 on those subcarriers in an earlier phase, thus leading to power modifications on their own subcarriers, especially their multiplexed ones, and so on. To avoid such an inconvenient behavior, the powers of multiplexed subcarriers get fixed once the pairing is performed and the task of rate compensation is carried out over the user's sole subcarriers. Initially, all the subcarriers of a user are sole subcarriers ($\mathcal{S}_k^{sole} = \mathcal{S}_k, \forall k$), and the initial waterlevel of every user in the NOMA phase is the final waterline obtained in the OMA phase. To determine the power variation subsequent to a user pairing, we let $N_{k_2}^{sole}$ be the cardinal of $\mathcal{S}_{k_2}^{sole}$. From equation (2.10), the waterline over the sole subcarriers of k_2 is given by:

$$W_{k_2}(N_{k_2}^{sole}) = \sigma^2 \left(2^{SR_{k_2}^{sole}/B} / \prod_{m \in \mathcal{S}_{k_2}^{sole}} h_{k_2,m,r(m)} \right)^{1/N_{k_2}^{sole}},$$

with $R_{k_2}^{sole}$ the total rate of sole subcarriers of k_2. Since k_2 is allocated as a second user on the new subcarrier n, its sole subcarrier set is unchanged. Also, the additional rate $R_{k_2,n,r}$ due to the allocation of n corresponds to the rate decrease $\Delta R_{k_2,n,r}$ that should be compensated for on the sole subcarriers of k_2, so as to ensure the global rate constraint $R_{k_2,req}$. In other words, the variation that rate $R_{k_2}^{sole}$ undergoes is opposite to the rate addition that comes along the new subcarrier assignment. We can write the new rate that must be achieved on $\mathcal{S}_{k_2}^{sole}$ as $R_{k_2}^{sole'} = R_{k_2}^{sole} + \Delta R_{k_2,n,r}$ where the rate decrease $\Delta R_{k_2,n,r}$ is negative and equal to $-R_{k_2,n,r}$. Recall from (2.14) that the waterline expression after a target rate variation ΔR is given by:

$$W_{k_2,new}(N_{k_2}^{sole}) = W_{k_2}(N_{k_2}^{sole}) 2^{S\Delta R/BN_{k_2}^{sole}}. \tag{2.16}$$

Thus, the power variation of k_2 due to pairing is the sum of two terms: a power increase relative to the newly allocated subcarrier $P_{k_2,n,r}$, and a power decrease over $\mathcal{S}_{k_2}^{sole}$ due to the rate compensation and given by (2.15), which leads to the expression in (2.17):

$$\Delta P = N_{k_2}^{sole} W_{k_2}(N_{k_2}^{sole}) (2^{\frac{-SR_{k_2,n,r}}{BN_{k_2}^{sole}}} - 1) + P_{k_2,n,r}. \tag{2.17}$$

We present next the power control mechanism for determining the multiplexed subcarrier's power by distinguishing two serving cases, one when the antennas of the two multiplexed users are the same, and another when they are different.

2.5.1 Same Serving RRH

The allocated power for user k_2 selected as a second user over subcarrier n has to verify the PMC condition $P_{k_2,n,r} > P_{k_1,n,r}$. A PA rule commonly used in the literature [27,28] is FTPA where $P_{k_2,n,r}$ is given by: $P_{k_2,n,r} = P_{k_2,n,r}(h_{k_2,n,r}/h_{k_1,n,r})^{-\zeta}$, with ζ a decay factor, set to 0.5 in this thesis. The idea behind this design is to exploit the gap between the channel gains of the multiplexed users. The greater the gap in channel coefficients, the greater the allocated power to the second user on the studied subcarrier. This method will be referred to as "SRRH". While this approach guarantees SIC stability in an efficient manner, it is not optimized for the context of power minimization which is ours. Therefore, we introduce next the SRRH-LPO algorithm where $P_{k_2(n),n,r}$ is set such that power variation is minimized. The algorithm relies on the Local Power Optimization (LPO) PA scheme developed below.

Local Power Optimization

The power decrease incurred by a candidate subcarrier n in the SRRH technique is greatly influenced by the amount of power $P_{k_2,n,r}$ allocated to user k_2 on n using FTPA. Indeed, the addition of a new subcarrier translates into an increase of the power level allocated to the user on the one hand, and conversely into a power decrease for the same user due to the subsequent waterline reduction on its sole subcarriers on the other hand. Therefore, we propose to optimize the value of $P_{k_2,n,r}$ in such a way that the resulting user power reduction is minimized:

$$\min_{P_{k_2,n,r}} \Delta P_{k_2},$$

subject to:

$$P_{k_2,n,r} \geq P_{k_1,n,r}.$$

By expressing $R_{k_2,n,r}$ in terms of $P_{k_2,n,r}$ using (2.2), we can formulate the Lagrangian of this optimization problem as:

$$L(P_{k_2,n,r},\lambda) = P_{k_2,n,r} + N_{k_2}^{sole} W_{k_2}(N_{k_2}^{sole}) \left(\left(1 + \frac{P_{k_2,n,r} h_{k_2,n,r}}{P_{k_1,n,r} h_{k_2,n,r} + \sigma^2} \right)^{-\frac{1}{N_{k_2}^{sole}}} - 1 \right) + \lambda (P_{k_2,n,r} - P_{k_1,n,r}),$$

where λ is the Lagrange multiplier. The corresponding KKT conditions are:

$$\begin{cases} 1 + \lambda - \dfrac{W_{k_2}(N_{k_2}^{sole}) h_{k_2,n,r}}{P_{k_1,n,r} h_{k_2,n,r} + \sigma^2} \left(1 + \dfrac{P^*_{k_2,n,r} h_{k_2,n,r}}{P_{k_1,n,r} h_{k_2,n,r} + \sigma^2} \right)^{\frac{-N_{k_2}^{sole}-1}{N_{k_2}^{sole}}} = 0, \\ \lambda (P^*_{k_2,n,r} - P_{k_1,n,r}) = 0. \end{cases}$$

We can check that the second derivative of the Lagrangian is always positive, and therefore the corresponding solution is the global minimum. For $\lambda = 0$, this optimum is:

$$P^*_{k_2,n,r} = \left(\left(\frac{W_{k_2}(N_{k_2}^{sole}) h_{k_2,n,r}}{P_{k_1,n,r} h_{k_2,n,r} + \sigma^2} \right)^{\frac{N_{k_2}^{sole}}{N_{k_2}^{sole}+1}} - 1 \right) \left(P_{k_1,n,r} + \frac{\sigma^2}{h_{k_2,n,r}} \right). \tag{2.18}$$

For $\lambda \neq 0$, $P^*_{k_2,n,r} = P_{k_1,n,r}$. However in such cases, with no power difference between the two paired users successful SIC decoding is jeopardized at the receiver side for the first user. To overcome this problem, we take:

$$P^*_{k_2,n,r} = P_{k_1,n,r}(1 + \mu), \tag{2.19}$$

with μ a positive safety power margin that depends on practical SIC implementation. In other terms, if the obtained $P^*_{k_2,n,r}$ in (2.18) verifies the power constraint inequality, it is retained as the optimal solution, otherwise, it is taken as in (2.19).
On another hand, considering that the selected user suffers from the interference of the strong user, and that $P_{k_1,n,r}$ is constant through the PA over the subcarriers of k_2, then the addition of n can be seen as a simple subcarrier addition in the waterfilling process (section2.4.1.1), but with a noise power of $\sigma^2 + P_{k_1,n,r} h_{k_2,n,r}$ instead of σ^2. Indeed, a closer

look at (2.18) reveals that the optimal power expression can be written as the difference of a waterline and an inverse channel gain:

$$P^*_{k_2,n,r} = \left(\frac{W_{k_2}(N^{sole}_{k_2})h_{k_2,n,r}}{P_{k_1,n,r}h_{k_2,n,r}+\sigma^2}\right)^{\frac{N^{sole}_{k_2}}{N^{sole}_{k_2}+1}} \times \left(\frac{P_{k_1,n,r}h_{k_2,n,r}+\sigma^2}{h_{k_2,n,r}}\right) - \frac{P_{k_1,n,r}h_{k_2,n,r}+\sigma^2}{h_{k_2,n,r}},$$

The first term of $P^*_{k_2,n,r}$ can be rearranged to yield (2.20) which is the application of (2.11) for the iterative waterline update after subcarrier addition, where h_{new} is set to $h_{k_2,n,r}$, and the background noise σ^2 is set to the background plus interference noise $\sigma^2 + P_{k_1,n,r}h_{k_2,n,r}$.

$$W'_{k_2} = \left(\frac{W_{k_2}(N^{sole}_{k_2})^{N^{sole}_{k_2}}}{h_{k_2,n,r}/(P_{k_1,n,r}h_{k_2,n,r}+\sigma^2)}\right)^{\frac{1}{N^{sole}_{k_2}+1}}. \tag{2.20}$$

Therefore, the optimal allocated power in (2.18) is also given by:

$$P^*_{k_2,n,r} = W'_{k_2} - P_{k_1,n,r} - \frac{\sigma^2}{h_{k_2,n,r}}. \tag{2.21}$$

The similarities with the waterfilling properties can also be expended to get the equivalent condition of (2.12) for subcarrier addition. This allows the rejection of candidate subcarriers whose allocation will necessarily increase the power of user k_2 if they do not satisfy:

$$h_{k_2,n,r} > \frac{\sigma^2}{W_{k_2}(N^{sole}_{k_2}) - P_{k_1,n,r}}. \tag{2.22}$$

This method, referred to as "SRRH-LPO", operates similarly to SRRH, except for the FTPA power allocation which is replaced by either (2.20) or (2.19). The SRRH and SRRH-LPO algorithms are presented below in algorithm 2.3.

Algorithm 2.3 SRRH, SRRH-LPO

Phase 1: OMA-DAS
Phase 2: NOMA pairing
$\mathcal{U}_p = [1:K]$ // reinitialize the set of active users
while $\mathcal{S}_f \neq \emptyset \wedge \mathcal{U}_p \neq \emptyset$
  $k_2 = \arg\max_k P_{k,tot}$
  for every $n \in \mathcal{S}_f \cap \{\mathcal{S}_{k_2}\}^c$ s.t. $h_{k_2,n,r} < h_{k_1,n,r}$ and (2.22) // $r_2(n) = r_1(n) = r$
    Calculate $P_{k_2,n,r}$ through FTPA for SRRH, LPO for SRRH-LPO
    Calculate $W'_{k_2}(N^{sole}_{k_2})$ using (2.16) for SRRH, (2.20) for SRRH-LPO
    Calculate $\Delta P_{k_2,n,r}$ using (2.17)
  end for
  $n^* = \arg\min_n \Delta P_{k_2,n,r}$ // end of the subcarrier search phase
  if $\Delta P_{k_2,n^*,r} < -\rho$ // subcarrier allocation
    Assign k_2 on n^* and remove n^* from $\mathcal{S}_f$
    Fix P_{k_1,n^*,r^*} and P_{k_2,n^*,r^*}, update $P_{k_2,n,r}, \forall n \in \mathcal{S}^{sole}_{k_2}$ // thus $P_{k_2,tot}$ is updated
  else remove k_2 from $\mathcal{U}_p$
end while

Note that after the WBH phase, an iteration in OMA and NOMA phases both result in either the allocation of a subcarrier-RRH pair, or the rejection of a user from the set $\mathcal{U}_p$ of active users (in case of a negligible power decrease). Either ways, the total number of available subcarriers or active users is decreased by one every iteration. In the best case (in terms of complexity), OMA and NOMA phases involve K iterations, that is when all the users get dismissed without getting allocated a subcarrier in the OMA phase, or paired over a subcarrier in NOMA. In the worst case, the algorithm ends with an empty set of active users and a single subcarrier in $\mathcal{S}_p$ or $\mathcal{S}_f$ (for OMA and NOMA respectively), or with a single user in $\mathcal{U}_p$ and empty sets in $\mathcal{S}_p$ and $\mathcal{S}_f$ (for OMA and NOMA respectively). Therefore, during the OMA phase, the sum $|\mathcal{S}_p| + |\mathcal{U}_p|$ goes from S at the beginning of orthogonal multiplexing, to 1 at the end, resulting in $S-1$ iterations. In the NOMA phase, the sum $|\mathcal{S}_f| + |\mathcal{U}_p|$ goes from $S+K$ at the beginning of the pairing, to 1 at the end of the pairing, resulting in $S+K-1$ iterations. These considerations are central to the complexity analysis in section 2.6 and they prove the stability of the proposed resource allocation approaches for OMA and NOMA.

2.5.2 Different serving RRHs

The rest of the chapter aims at designing specific NOMA RA schemes capturing the unique properties that arise in DAS when subcarrier multiplexing is done through different RRHs. We start by developing the theoretical foundation lying behind SIC implementation when different RRHs are used to power the multiplexed signals on a subcarrier. The results show that under some well defined conditions, both paired users can perform SIC on the subcarrier. Finally, we propose several RA schemes taking advantage of the capacity gains inherent to mutual SIC and combine them with single SIC techniques.

2.5.2.1 Theoretical Background

In the case where the same RRH powers both multiplexed users on a subcarrier, there always exists one strong user at a given time which is the user having the best subcarrier-RRH link. However, this isn't necessarily the case when different RRHs are chosen to power the subcarrier, since the concept of weak and strong users is only valid relatively to a specific transmitting antenna. Indeed, the greater diversity provided by powering multiplexed subcarriers by different RRHs involves four instead of two different user-RRH links and thus opens the possibility of having more than one "strong" user at a time. To simplify the notations, users $k_1(n)$ and $k_2(n)$ on a subcarrier n, and their transmitting antennas $r_1(n)$ and $r_2(n)$ are simply referred to as k_1, k_2, r_1, r_2.

Theorem 1. *Two users k_1 and k_2, paired on subcarrier n and powered by two different RRHs, respectively r_1 and r_2, can both perform SIC if:*

$$h_{k_1,n,r_2} \geq h_{k_2,n,r_2} \tag{2.23}$$

$$h_{k_2,n,r_1} \geq h_{k_1,n,r_1} \tag{2.24}$$

Proof. Let s_1 be the signal of user k_1 transmitted by RRH r_1 with power P_{k_1,n,r_1}, and let s_2 be the signal of user k_2 transmitted by RRH r_2 with power P_{k_2,n,r_2}. Therefore, the channel conditions experienced by every signal arriving at a given user are different: at the level of

k_1, the power levels of signals s_1 and s_2 are $P_{k_1,n,r_1}h_{k_1,n,r_1}$ and $P_{k_2,n,r_2}h_{k_1,n,r_2}$ respectively. Similarly, at the level of k_2, the power levels of signals s_1 and s_2 are $P_{k_1,n,r_1}h_{k_2,n,r_1}$ and $P_{k_2,n,r_2}h_{k_2,n,r_2}$ respectively. Depending on their respective signal quality, users k_1 and k_2 can decode signal s_2 at different rates. Let $R_{k_2}^{(k_1)}$ be the necessary rate at the level of user k_1 to decode the signal of user k_2 in the presence of the signal of user k_1. And let $R_{k_2}^{(k_2)}$ be the necessary rate to decode the signal of user k_2 at the level of k_2 in the presence of the signal of user k_1. The capacity that can be achieved by k_1 and k_2 over signal s_2 and in the presence of interfering signal s_1 is given by the Shannon limit:

$$R_{k_2}^{(k_1)} = \frac{B}{S}\log_2\left(1+\frac{P_{k_2,n,r_2}h_{k_1,n,r_2}}{P_{k_1,n,r_1}h_{k_1,n,r_1}+\sigma^2}\right) \tag{2.25}$$

$$R_{k_2}^{(k_2)} = \frac{B}{S}\log_2\left(1+\frac{P_{k_2,n,r_2}h_{k_2,n,r_2}}{P_{k_1,n,r_1}h_{k_2,n,r_1}+\sigma^2}\right) \tag{2.26}$$

For k_1 to be able to perform SIC, the rates should satisfy the following condition:

$$R_{k_2}^{(k_1)} \geq R_{k_2}^{(k_2)} \tag{2.27}$$

By writing: $R_{k_2}^{(k_1)} - R_{k_2}^{(k_2)} = \frac{B}{S}\log_2\left(\frac{X}{Y}\right)$, we can express $X-Y$ as:

$$X-Y = P_{k_1,n,r_1}P_{k_2,n,r_2}\left(h_{k_1,n,r_2}h_{k_2,n,r_1}-h_{k_2,n,r_2}h_{k_1,n,r_1}\right)+\sigma^2P_{k_2,n,r_2}\left(h_{k_1,n,r_2}-h_{k_2,n,r_2}\right) \tag{2.28}$$

Similarly, for user k_2, the rate condition that should be satisfied for the implementation of SIC at the level of k_2 is:

$$R_{k_1}^{(k_2)} \geq R_{k_1}^{(k_1)} \tag{2.29}$$

$R_{k_1}^{(k_2)}$ and $R_{k_1}^{(k_1)}$ can be obtained from (2.25) and (2.26) by interchanging indexes 1 and 2. Also, by writing: $R_{k_1}^{(k_2)} - R_{k_1}^{(k_1)} = \frac{B}{S}\log_2\left(\frac{Z}{T}\right)$, we get:

$$Z-T = P_{k_2,n,r_2}P_{k_1,n,r_1}\left(h_{k_2,n,r_1}h_{k_1,n,r_2}-h_{k_1,n,r_1}h_{k_2,n,r_2}\right)+\sigma^2P_{k_1,n,r_1}\left(h_{k_2,n,r_1}-h_{k_1,n,r_1}\right) \tag{2.30}$$

Note that for special case $r_1 = r_2 = r$, we get:

$$X-Y = \sigma^2P_{k_2,n,r}\left(h_{k_1,n,r}-h_{k_2,n,r}\right)$$
$$Z-T = -\sigma^2P_{k_1,n,r}\left(h_{k_1,n,r}-h_{k_2,n,r}\right)$$

Therefore, either (2.28) or (2.30) is positive, not both, which justifies why only the stronger user, the one with the higher channel gain, is able to perform SIC as it has been stated in all the literature on NOMA [27–29, 63, 92].

For both users to perform SIC, the rate conditions (2.27) and (2.29) must be verified at the same time. From (2.28) and (2.30), we infer that the following two conditions are sufficient to enable mutual SIC:

$$h_{k_1,n,r_2} \geq h_{k_2,n,r_2}$$
$$h_{k_2,n,r_1} \geq h_{k_1,n,r_1}$$

Indeed, these conditions ensure the positivity of each of the two terms in both $X-Y$ and $Z-T$. This concludes our proof. □

Regarding the PMCs, the key is to design the PA scheme in such a way that the received power of the first signal to be decoded is larger than the power of the other signal. The resulting power conditions for users k_1 and k_2 respectively become:

$$P_{k_1,n,r_1} h_{k_1,n,r_1} \leq P_{k_2,n,r_2} h_{k_1,n,r_2}$$
$$P_{k_2,n,r_2} h_{k_2,n,r_2} \leq P_{k_1,n,r_1} h_{k_2,n,r_1}$$

They can be combined into the following condition:

$$\frac{h_{k_1,n,r_1}}{h_{k_1,n,r_2}} \leq \frac{P_{k_2,n,r_2}}{P_{k_1,n,r_1}} \leq \frac{h_{k_2,n,r_1}}{h_{k_2,n,r_2}} \tag{2.31}$$

Remark. *If (2.23) and (2.24) are true, then* $\frac{h_{k_1,n,r_1}}{h_{k_1,n,r_2}} \leq \frac{h_{k_2,n,r_1}}{h_{k_2,n,r_2}}$. *In this case, a PA scheme can be found to allow a mutual SIC, i.e. there exist* P_{k_1,n,r_1} *and* P_{k_2,n,r_2} *such that (2.31) is fulfilled.*

Finally, conditions (2.23) and (2.24) are sufficient but not necessary for the application of mutual SIC. Actually, the conditions for the application of mutual SIC lie in the positivity of (2.28) and (2.30). If any of (2.23) or (2.24) is not valid, the power terms in (2.28) and (2.30) should be considered, since they affect the sign of both equations. However, a closer examination of (2.28) and (2.30) reveals that in practical systems, their numerical values are greatly dominated by their first common term, since in general $\sigma^2 << Ph_{k,n,r}$ [98,99]. In that regard, a simpler constraint on the channel gains is derived:

$$h_{k_1,n,r_1} h_{k_2,n,r_2} \leq h_{k_2,n,r_1} h_{k_2,n,r_1} \tag{2.32}$$

This constraint will be used instead of (2.23) and (2.24) in the sequel. Note that condition (2.32) also ensures the existence of a PA scheme that allows mutual SIC. When both users k_1 and k_2 perform SIC on a subcarrier n, their reachable rates on n are given by:

$$R_{k_1,n,r_1} = \frac{B}{S} \log_2 \left(1 + \frac{P_{k_1,n,r_1} h_{k_1,n,r_1}}{\sigma^2}\right) \tag{2.33}$$

$$R_{k_2,n,r_2} = \frac{B}{S} \log_2 \left(1 + \frac{P_{k_2 n,r_2} h_{k_2,n,r_2}}{\sigma^2}\right) \tag{2.34}$$

Following the introduction of mutual SIC, the RA strategy should be modified accordingly. Therefore, the next sections describe the development of novel RA techniques that can benefit from this new potential of the NOMA-DAS combination.

2.5.2.2 Mutual SIC UnConstrained (MutSIC-UC)

In addition to the selection of different antennas in the pairing phase of algorithm 2.3, the key modifications that must be accounted for when moving from single SIC to mutual SIC RA schemes involve:

- New subcarrier subset selection: only the subcarrier-RRH links satisfying (2.32) are considered for potential assignment in mutual SIC configurations.
- Power assignment: PMC (2.31) must be accounted for instead of (2.6).

To get a lower bound on the performance of mutual SIC-based RA, we solve a relaxed version of the problem without PMCs. This consideration reverts the optimal PA scheme in the pairing phase to the user-specific waterfilling solution in OMA. Therefore, the pairing phase in mutual SIC becomes a simple extension of the OMA resource allocation in algorithm 2.2. This method is referred to as MutSIC-UC.

To compensate for the disregarded constraints, subcarrier assignment should be followed by a power optimization step as shown in appendix 2.A. However, the set of possible power corrections grows exponentially with the number of multiplexed subcarriers. Therefore, alternative suboptimal strategies accounting for the power multiplexing constraints at every subcarrier assignment are investigated in the following sections.

2.5.2.3 Mutual SIC with Direct Power Adjustment (MutSIC-DPA)

From a power minimization perspective, the power distribution obtained through waterfilling is the best possible PA scheme. However, compliance with the PMCs is not guaranteed; therefore, a power adjustment might be resorted to for the multiplexed subcarriers.

When an adjustment is needed, the new value of P_{k_2,n,r_2} in (2.31) should fall between $P_{k_1,n,r_1}h_{k_1,n,r_1}/h_{k_1,n,r_2}$ and $P_{k_1,n,r_1}h_{k_2,n,r_1}/h_{k_2,n,r_2}$ (the value of P_{k_1,n,r_1} is fixed). However, since any deviation from the waterfilling procedure degrades the performance of the solution, this deviation must be minimal. Therefore, P_{k_2,n,r_2} is set at the nearest limit of the inequality (2.31), with some safety margin μ accounting for proper SIC decoding. This power adjustment is conducted at both the SA stage (to determine the best candidate couple (n, r_2) for user k_2) and the PA stage (following the selection of the subcarrier-RRH pair). After the subcarrier-RRH pairing, the powers on the multiplexed subcarrier of both users are kept unvaried, as in algorithm 2.3. This procedure will be referred to as MutSIC-DPA; its details are presented in algorithm 2.4.

Algorithm 2.4 MutSIC-DPA

Phase 1: OMA-DAS
Phase 2: NOMA pairing using mutual SIC
$k_2 = \arg\max_k P_{k,tot}$
$S_c = \{(n, r_2)$ s.t. (2.32) & (2.12) are verified$\}$
for every candidate couple $(n, r_2) \in S_c$
  Calculate $P^*_{k_2,n,r_2}$ and $\Delta P_{k_2,n,r_2}$ using (2.11) and (2.13)
  If $P^*_{k_2,n,r_2}$ verifies (2.31), set $P_{k_2,n,r_2} = P^*_{k_2,n,r_2}$
  If $\frac{P_{k_2,n,r_2}}{P_{k_1,n,r_1}} < \frac{h_{k_1,n,r_1}}{h_{k_1,n,r_2}} \Rightarrow$ set $P_{k_2,n,r_2} = (1+\mu)P_{k_1,n,r_1}\frac{h_{k_1,n,r_1}}{h_{k_1,n,r_2}}$, estimate $\Delta P_{k_2,n,r_2}$ using (2.14), (2.17)
  If $\frac{P_{k_2,n,r_2}}{P_{k_1,n,r_1}} > \frac{h_{k_2,n,r_1}}{h_{k_2,n,r_2}} \Rightarrow$ set $P_{k_2,n,r_2} = (1-\mu)P_{k_1,n,r_1}\frac{h_{k_2,n,r_1}}{h_{k_2,n,r_2}}$, estimate $\Delta P_{k_2,n,r_2}$ using (2.14), (2.17)
end for
$(n^*, r_2^*) = \arg\min_{(n,r_2)} \Delta P_{k_2,n,r_2}$.
Continue the assignment similarly to SRRH using DPA when needed

2.5.2.4 Mutual SIC with Sequential Optimization for Power Adjustment (MutSIC-OPAd, MutSIC-SOPAd, and Mut&SingSIC)

In order to improve on the MutSIC-DPA technique, we propose to replace the adjustment and power estimation steps by a sequential power optimization. Instead of optimizing the choice of P_{k_2,n,r_2} over the candidate couple (n, r_2), we look for a wider optimization in which powers of both first and second users on the considered subcarrier are adjusted, in a way that their global power variation $\Delta P_{k_1} + \Delta P_{k_2}$ is minimal:

$$\{P_{k_1,n,r_1}, P_{k_2,n,r_2}\}^* = \underset{P_{k_1,n,r_1}, P_{k_2,n,r_2}}{\arg\max} \; (-\Delta P_{k_1} - \Delta P_{k_2})$$

$$\text{subject to:} \quad \frac{h_{k_1,n,r_1}}{h_{k_1,n,r_2}} \leq \frac{P_{k_2,n,r_2}}{P_{k_1,n,r_1}}, \quad \frac{P_{k_2,n,r_2}}{P_{k_1,n,r_1}} \leq \frac{h_{k_2,n,r_1}}{h_{k_2,n,r_2}}$$

The power variations of users k_2 and k_1 are given by:

$$\Delta P_{k_2} = N_{k_2}^{sole} W_{I,k_2} (2^{\frac{\Delta R_{k_2} S}{B N_{k_2}^{sole}}} - 1) + P_{k_2,n,r_2}$$

$$\Delta P_{k_1} = (N_{k_1}^{sole} - 1) W_{I,k_1} \left(2^{\frac{\Delta R_{k_1} S}{(N_{k_1}^{sole}-1) B}} - 1\right) + P_{k_1,n,r_1} - P^I_{k_1,n,r_1}$$

where $P^I_{k_1,n,r_1}$ is the initial power allocated to k_1 on n, W_{I,k_1} and W_{I,k_2} are the initial waterlines of k_1 and k_2 before pairing, and ΔR_{k_1} and ΔR_{k_2} the rate variations over the remaining sole subcarriers of k_1 and k_2 (after pairing). They are given by:

$$\Delta R_{k_1} = -\frac{B}{S} \log_2 \left(\frac{\sigma^2 + P_{k_1,n,r_1} h_{k_1,n,r_1}}{\sigma^2 + P^I_{k_1,n,r_1} h_{k_1,n,r_1}} \right), \qquad \Delta R_{k_2} = -\frac{B}{S} \log_2 \left(1 + \frac{P_{k_2,n,r_2} h_{k_2,n,r_2}}{\sigma^2} \right)$$

The Lagrangian of this problem is:

$$\begin{aligned} L(P_{k_1,n,r_1}, P_{k_2,n,r_2}, \lambda_1, \lambda_2) = &- \lambda_1 \left(P_{k_1,n,r_1} \frac{h_{k_1,n,r_1}}{h_{k_1,n,r_2}} - P_{k_2,n,r_2} \right) - \lambda_2 \left(P_{k_2,n,r_2} - P_{k_1,n,r_1} \frac{h_{k_2,n,r_1}}{h_{k_2,n,r_2}} \right) \\ &- \Delta P_{k_1,n,r_1} - \Delta P_{k_2,n,r_2} \end{aligned}$$

The solution of this problem must verify the following conditions:

$$\begin{cases} \nabla L(P_{k_1,n,r_1}, P_{k_2,n,r_2}, \lambda_1, \lambda_2) = 0 \\ \lambda_1 \left(P_{k_1,n,r_1} h_{k_1,n,r_1} / h_{k_1,n,r_2} - P_{k_2,n,r_2} \right) = 0 \\ \lambda_2 \left(P_{k_2,n,r_2} - P_{k_1,n,r_1} h_{k_2,n,r_1} / h_{k_2,n,r_2} \right) = 0 \\ \lambda_1, \lambda_2 \geq 0 \end{cases}$$

Four cases are identified:

1. $\lambda_1 = 0, \lambda_2 = 0$
2. $\lambda_1 \neq 0, \lambda_2 = 0 \rightarrow P_{k_2,n,r_2} = P_{k_1,n,r_1} h_{k_1,n,r_1} / h_{k_1,n,r_2}$
3. $\lambda_1 = 0, \lambda_2 \neq 0 \rightarrow P_{k_2,n,r_2} = P_{k_1,n,r_1} h_{k_2,n,r_1} / h_{k_2,n,r_2}$
4. $\lambda_1 \neq 0, \lambda_2 \neq 0$

Case 1 corresponds to the unconstrained waterfilling solution applied separately to the two users. Case 4 is generally impossible, since the two boundaries of the inequality (2.31) would be equal. Considering case 2, by replacing P_{k_2,n,r_2} in terms of P_{k_1,n,r_1} in the Lagrangian and by taking the derivative with respect to P_{k_1,n,r_1}, we can verify that $P^*_{k_1,n,r_1}$ is the solution of the following nonlinear equation:

$$W_{I,k_2}\frac{h_{k_1,n,r_1}h_{k_2,n,r_2}}{h_{k_1,n,r_2}\sigma^2}\left(1+\frac{P_{k_1,n,r_1}h_{k_2,n,r_2}h_{k_1,n,r_1}}{\sigma^2}\right)^{-\frac{1}{N^{sole}_{k_2}}-1} + \frac{W_{I,k_1}h_{k_1,n,r_1}}{\sigma^2+P^I_{k_1,n,r_1}h_{k_1,n,r_1}}\left(\frac{\sigma^2+P_{k_1,n,r_1}h_{k_1,n,r_1}}{\sigma^2+P^I_{k_1,n,r_1}h_{k_1,n,r_1}}\right)^{-\frac{1}{N^{sole}_{k_1}-1}-1} - \frac{h_{k_1,n,r_1}}{h_{k_1,n,r_2}}-1=0 \quad (2.35)$$

Note that in practice, we also take into consideration the safety power margin μ in the calculation of P_{k_1,n,r_1}. Similar calculations are performed for case 3, and the solution that yields the lowest ΔP is retained. Also, if none of the cases provides positive power solutions, the current candidate couple (n, r_2) is discarded. This method of Optimal Power Adjustment (OPAd) is employed both at the subcarrier allocation stage (for the selection of the best candidate couple (n, r_2) for user k_2) and at the power allocation stage (following the selection of the subcarrier-RRH pair). It will be referred to as "MutSIC-OPAd".

In order to decrease the complexity of MutSIC-OPAd, inherent to the resolution of a nonlinear equation for every subcarrier-RRH candidate, we consider a "semi-optimal" variant of this technique, called "MutSIC-SOPAd": at the stage where candidate couples (n, r_2) are considered for potential assignment to user k_2, DPA is used for power adjustment to determine the best candidate in a cost-effective way. Then, the preceding OPAd solution is applied to allocate power levels to users k_1 and k_2 on the retained candidate.

At last, to further exploit the space diversity inherent to DAS and minimize the system transmit power, single SIC and mutual SIC algorithms are combined to take advantage of the full potential of NOMA techniques. Given the superiority of mutual SIC over single SIC schemes, we prioritize the allocation of subcarriers allowing mutual SIC by first applying MutSIC-SOPAd. Then, the remaining set of solely assigned subcarriers is further examined for potential allocation of a second user in the single SIC context, using the same RRH as that of the first assigned user. LPO is used for power allocation in this second phase. This method will be referred to as "Mut&SingSIC".

2.6 Complexity Analysis

In this section, we analyze the complexity of the proposed allocation techniques. Given that the algorithms consist in sequential blocks of OMA assignment and NOMA pairing, we analyze the complexity of each step independently and then derive the complexity of each algorithm by combining the corresponding steps. Also, to have a ground of comparison with the CAS scenario, the complexities of OMA-CAS, NOMA-CAS and OMA-DAS are presented. Throughout all the section, we assume that an average of S

iterations are needed for the completion of the OMA phase, and that S iterations are required on average for the NOMA phases.

In the OMA section, the core of the algorithm resides in searching for the most power consuming user, which presents a linear complexity with the number of users ($O(K)$), assigning him the best subcarrier-RRH pair ($O(RS)$), and iterating this process S times until all subcarriers are allocated. The resulting complexity is of $O(S(K+RS))$. Suppose, on the other hand, that we sort the $R\times K$ subcarrier vectors (of length S) in channel matrix $\mathbb{H}$ prior to subcarrier-RRH assignment. This means the channel matrix is rearranged in such a way that the subcarriers of each user are sorted in the decreasing order of channel gain, separately for each RRH. In this case, the assignment of the best available subcarrier-RRH pair to the selected user reduces to searching for the best antenna with a complexity linear with R. Sorting $\mathbb{H}$ reduces the complexity of the subcarrier-RRH allocation phase S times, while adding a sorting complexity of $KSR\log(S)$. Each allocation cycle consists then of user identification, followed by the search of the RRH providing the subcarrier with the highest channel gain. The resulting complexity of this new approach is therefore $O(KSR\log(S)+S(K+R))$. This approach is roughly $S/K\log(S)$ times less complex then the preceding one (without matrix reordering), hence it will be used hereinafter for all the algorithms.

Each allocation step in the pairing phase of NOMA using SRRH consists of the identification of the most power consuming user, followed by a search over the subcarrier space and a power update over the set of sole subcarriers for the user. Assuming an average number of S/K subcarriers per user, the total complexity of SRRH and SRRH-LPO is $O(KSR\log(S)+S(K+R)+S(K+S+S/K))$. In order to assess the efficiency of SRRH-LPO, we compare our solution to the optimal PA technique developed in [92]. More specifically, we apply SRRH-LPO to determine the user-subcarrier-RRH assignment; then we apply the optimal PA in [92] without PMCs. Only the simulations yielding solutions abiding by the PMCs are included for possible comparison. This technique will be referred to as SRRH-OPA; its complexity analysis and comparison with SRRH-LPO is provided in appendix 2.B. Note that OMA-CAS and NOMA-CAS complexities are derived from the DAS scenario through replacing R by 1.

Concerning MutSIC-UC, by following the same reasoning as for OMA-DAS, and accounting for the search of an eventual collocated user for S subcarriers, we get a total complexity $O(KSR\log(S)+S(K+R)+S(K+R-1))$.

As for MutSIC-DPA, the total complexity is $O(KSR\log(S)+S(K+R)+S(K+S(R-1)+S/K))$, where the $S(R-1)$ term stems from the fact that the search over the subcarrier space in the pairing phase is conducted over all combinations of subcarriers and RRHs, except for the RRH of the first user on the candidate subcarrier.

Regarding MutSIC-OPAd, let C be the complexity of solving the nonlinear equation (2.35). The total complexity is therefore $O(KSR\log(S)+S(K+R)+S(K+S(R-1)C+S/K))$. Given that MutSIC-SOPAd solves (2.35) only once per allocation step, its complexity is $O(KSR\log(S)+S(K+R)+S(K+S(R-1)+S/K+C))$. Consequently, the complexity of Mut&SingSIC is $O(KSR\log(S)+S(K+R)+S(K+S(R-1)+S/K+C)+S(K+S+S/K)$. The additional term corresponds to the single SIC phase which is similar to the pairing phase in SRRH.

Table 2.1 – Approximate complexity of the different allocation techniques.

RA Technique	**Complexity**	**RA Technique**	**Complexity**
OMA-CAS	$O(KS\log(S))$	MutSIC-UC	$O(KSR\log(S))$
NOMA-CAS	$O(S^2+KS\log(S))$	MutSIC-DPA	$O(RS^2+KSR\log(S))$
OMA-DAS	$O(KSR\log(S))$	MutSIC-OPAd	$O(RCS^2+KSR\log(S))$
SRRH	$O(S^2+KSR\log(S))$	MutSIC-SOPAd	$O(RS^2+SC+KSR\log(S))$
SRRH-LPO	$O(S^2+KSR\log(S))$	Mut&SingSIC	$O(RS^2+SC+KSR\log(S))$

To give an idea of the relative complexity orders, Table 2.1 summarizes the approximate complexity of the different techniques. In fact, the complexity of the methods employing a numerical solver depends on the resolution cost C that depends on the closeness of the initial guess to the actual solution. In that regard, MutSIC-SOPAd is roughly C times less complex than MutSIC-OPAd, and has a complexity comparable to MutSIC-DPA.

2.7 Performance Results

2.7.1 System Parameters

The performance of the different allocation techniques is assessed through simulations in the LTE/LTE-Advanced context [100]. The cell is hexagonal with an outer radius R_d of 500 m. For DAS, we consider four RRHs ($R = 4$), unless specified otherwise. One antenna is located at the cell center, while the others are uniformly positioned on a circle of radius $2R_d/3$ centered on the cell center. The number of users in the cell is $K = 15$, except for Fig. 2.5. The system bandwidth B is 10 MHz, divided into $S = 64$ subcarriers except for Fig. 2.5. The transmission medium is a frequency-selective Rayleigh fading channel with a root mean square delay spread of 500 ns. We consider distance-dependent path loss with a decay factor of 3.76 and lognormal shadowing with an 8 dB variance. The noise power spectral density N_0 is 4.10^{-18} mW/Hz. Perfect knowledge of the channel gain by the BBU is assumed throughout the thesis. For typical system parameters, the system performance in terms of transmit power is mainly invariant with ρ, thus ρ is set to 10^{-3} W. A detailed analysis of the system behavior in terms of ρ can be found in [101] for OMA systems. The safety power margin μ is set to 0.01. The performance results of OMA-CAS, NOMA-CAS and OMA-DAS are also shown for comparison.

2.7.2 Simulation Results

Fig. 2.2 presents the total transmit power in the cell as a function of the requested rate considering only SRRH schemes for NOMA-based techniques. The results show that the DAS configuration greatly outperforms CAS: a large leap in power with a factor around 16 is achieved with both OMA and NOMA signaling. At a target rate of 12 Mbps, the required total power using SRRH, SRRH-LPO and SRRH-OPA is respectively 17.6%, 24.5%, and 26.1% less than in OMA-DAS. This shows a clear advantage of NOMA over OMA in the DAS context. Besides, applying LPO allows a power reduction of 7.7% over FTPA, with a similar computational load. The penalty in performance of LPO with respect to optimal PA is only 2% at 12 Mbps, but with a greatly reduced complexity.

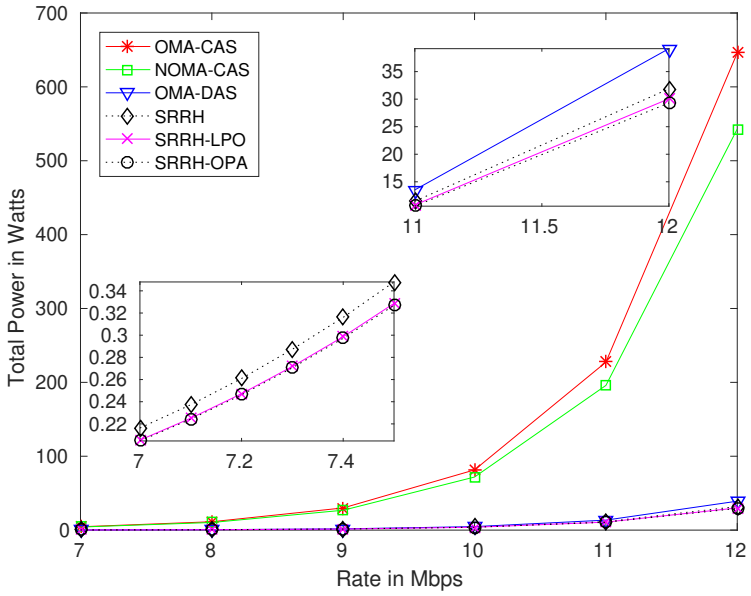


Figure 2.2 – Total power as a function of $R_{k,req}$ for DAS and CAS scenarios, with OMA and NOMA-SRRH schemes.

In Fig. 2.3, the results focus on the evaluation of mutual SIC and single SIC configurations. All three constrained configurations based on pure mutual SIC (MutSIC-DPA, MutSIC-SOPAd and MutSIC-OPAd) largely outperform SRRH-LPO. Their gain towards the latter is respectively 56.1%, 63.9% and 72.9%, at a requested rate of 13 Mbps. The significant gain of optimal power adjustment towards its suboptimal counterpart comes at the cost of a significant complexity increase, as shown in section 2.6. The most power-efficient mutual SIC implementation is obviously MutSIC-UC, since it is designed to solve a relaxed version of the power minimization problem by dropping all PMCs. Therefore, it only serves as a benchmark for assessing the other methods, because PMCs are essential for allowing correct signal decoding at the receiver side. Except for the OPAd solution, the best global strategy remains the combination of mutual and single SIC subcarriers, since it allows a power reduction of 15.2% and 15.6% at 12 and 13 Mbps respectively, when compared to MutSIC-SOPAd.

Fig. 2.4 shows the influence of increasing the number of RRHs on system performance. As expected, increasing the number of spread antennas greatly reduces the overall power, either with single SIC or combined mutual and single SIC configurations. A significant power reduction is observed when R is increased from 4 to 5, followed by a more moderate one when going from 5 to 7 antennas. The same behavior is expected for larger values of R. However, practical considerations like the overhead of CSI signaling exchange and the synchronization of the distributed RRHs, or geographical deployment constraints, would suggest limiting the number of deployed antennas in the cell.

In Fig. 2.5, we show the performance for a varying number of users, for the case of 4 RRHs and 128 subcarriers. Results confirm that the allocation strategies based on mutual SIC, or combined mutual and single SIC, scale much better to crowded areas, compared to single SIC solutions. The power reduction of Mut&SingSIC towards SRRH-LPO is 69.8% and 78.2% for 36 and 40 users respectively.

Table 2.2 shows the statistics of the number of non-multiplexed subcarriers, the num-

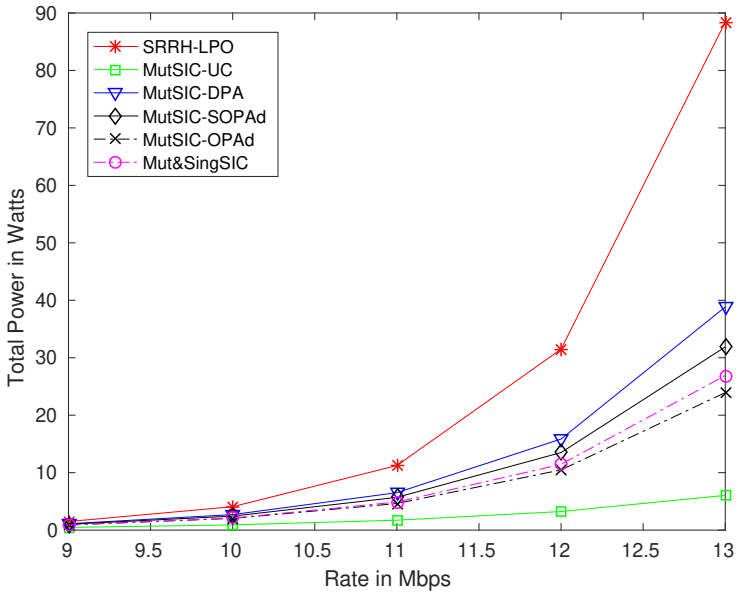


Figure 2.3 – Total power as a function of $R_{k,req}$ for the proposed NOMA-DAS schemes.

ber of subcarriers where a mutual SIC is performed, and the number of subcarriers where a single SIC is performed. On average, SRRH-LPO uses single SIC NOMA on 25% (resp. 32%) of the subcarriers for $R_{k,req} = 9$ Mbps (resp. 12 Mbps), while the rest of the subcarriers is solely allocated to users (a small proportion is not allocated at all, depending on the power threshold ρ). On the other hand, the proportions are respectively 17% and 23% with MutSIC-SOPAd. Therefore, in light of the results of Figs. 2.3 and 2.5, MutSIC-SOPAd not only outperforms SRRH-LPO from the requested transmit power perspective, but it also presents the advantage of yielding a reduced complexity at the UE level, by requiring a smaller amount of SIC procedures at the receiver side. This shows the efficiency of the mutual SIC strategy, combined with appropriate power adjustment, over classical single SIC configurations.

Table 2.2 – Statistics of subcarrier multiplexing, for K=15, S=64, and R=4.

RA technique	**Non Mux SC**	**SC MutSIC**	**SC SingSIC**
	$R_{k,req} = 9$ Mbps		
SRRH-LPO	48.1	-	15.9
MutSIC-SOPAd	53.4	10.6	-
Mut&SingSIC	39.2	10.6	14.2
	$R_{k,req} = 12$ Mbps		
SRRH-LPO	43.7	-	20.3
MutSIC-SOPAd	49.4	14.6	-
Mut&SingSIC	29	14.6	20.4

Note that in Mut&SingSIC, 17% (resp. 23%) of the subcarriers are powered from different antennas. This shows the importance of exploiting the additional spatial diversity, combined with NOMA, inherent to DAS.

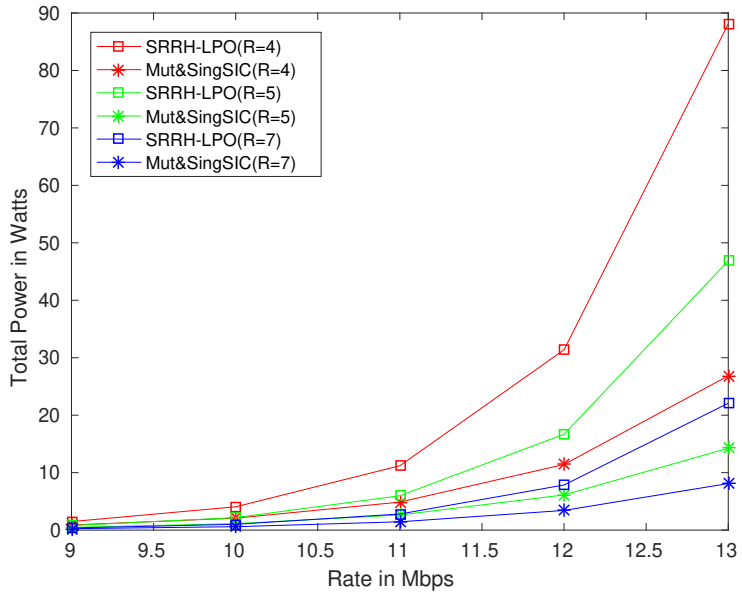


Figure 2.4 – Total power as a function of $R_{k,req}$ for NOMA-DAS schemes, with K=15, S=64, and R=4, 5 or 7.

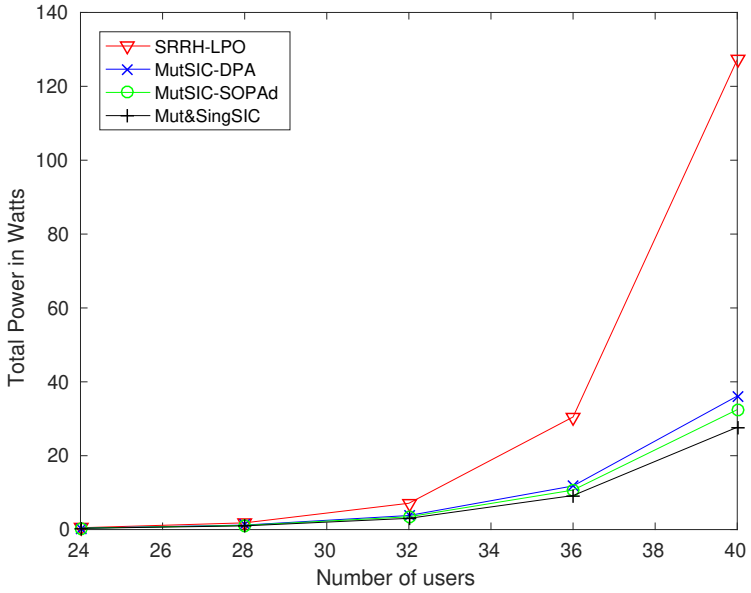


Figure 2.5 – Total power as a function of the number of users for the NOMA-DAS schemes, with $R_{k,req}$=5 Mbps, S=128, and R=4.

2.8 Conclusion

In this chapter, various RA techniques were presented for minimizing the total downlink transmit power in DAS for 5G and beyond networks. We first revisited the waterfilling principle prior to applying the acquired knowledge to designing efficient RAs in OMA and NOMA. Furthermore, we unveiled some of the hidden potentials of DAS for NOMA systems and developed new techniques to make the most out of these advantages, while

extracting their best characteristics and tradeoffs. Particularly, this study has enabled the design of NOMA with SIC decoding at both paired UE sides. Simulation results have shown the superiority of the proposed methods with respect to single SIC configurations. They also promoted mutual SIC with suboptimal power adjustment to the best tradeoff between transmit power and complexity at both the BBU and the UE levels. In order to address additional practical challenges of DAS, the next chapter focuses on transposing the solutions provided in this chapter to power minimization problems with power limited RRHs.

The contributions of this chapter led to the publication of the following journal paper:

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," in *IEEE Trans. Veh. Technol.*, vol. 67, no. 12, pp. 11873-11885, Dec. 2018.

Appendices

2.A Formulation of the Power Optimization Problem for the Constrained Case in Mutual SIC

For a predefined subcarrier-RRH-user assignment, the constrained power minimization problem for power assignment can be cast as the solution of the following optimization problem:

$$\max_{\{P_{k,n,r}\}} \left(-\sum_{k=1}^{K}\sum_{n=1}^{S}\sum_{r=1}^{R} P_{k,n,r} \right)$$

subject to:

$$\sum_{n\in\mathcal{S}_k} \log_2\left(1+\frac{P_{k,n,r}h_{k,n,r}}{\sigma^2}\right) = R_{k,req}, 1 \le k \le K$$

$$-\frac{P_{k_2,n,r_2}}{P_{k_1,n,r_1}} \le -\frac{h_{k_1,n,r_1}}{h_{k_1,n,r_2}}, \forall n \in \mathcal{S}_{mSIC}$$

$$\frac{P_{k_2,n,r_2}}{P_{k_1,n,r_1}} \le \frac{h_{k_2,n,r_1}}{h_{k_2,n,r_2}}, \forall n \in \mathcal{S}_{mSIC}$$

where $\mathcal{S}_{mSIC}$ is the set of subcarriers undergoing a mutual SIC. The corresponding Lagrangian with multipliers λ_k and $\beta_{i,n}$ is:

$$L(P,\lambda,\beta_1,\beta_2) = -\sum_{k=1}^{K}\sum_{n=1}^{S}\sum_{r=1}^{R} P_{k,n,r} + \sum_{n\in\mathcal{S}_{mSIC}} \beta_{1,n}\left(\frac{h_{k_2,n,r_1}}{h_{k_2,n,r_2}} - \frac{P_{k_2,n,r_2}}{P_{k_1,n,r_1}}\right)$$
$$+ \sum_{n\in\mathcal{S}_{mSIC}} \beta_{2,n}\left(\frac{h_{k_2,n,r_2}}{h_{k_1,n,r_1}} - \frac{P_{k_1,n,r_1}}{P_{k_1,n,r_2}}\right) + \sum_{k=1}^{K}\lambda_k\left(R_{k,req} - \sum_{n\in\mathcal{S}_k}\log_2\left(1+\frac{P_{k,n,r}h_{k,n,r}}{\sigma^2}\right)\right)$$

Writing the KKT conditions leads to a system of N_e non-linear equations with N_e variables, where $N_e = 3|S_{mSIC}| + K + S$ (taking into account the $S - |\mathcal{S}_{mSIC}|$ power variables on non-paired subcarriers). Knowing that $\beta_{1,n}$ and $\beta_{2,n}$ cannot be simultaneously non-zero, we have, for every subcarrier allocation scheme, a total of $3^{|\mathcal{S}_{mSIC}|}$ different possible combinations to solve, that is $3^{|\mathcal{S}_{mSIC}|}$ different variations of a square system of $2|\mathcal{S}_{mSIC}| + K + S$ equations (per subcarrier allocation).

2.B Complexity Analysis of SRRH-OPA and Comparison with SRRH-LPO

SRRH-OPA consists in successively applying SRRH-LPO to set the subcarrier-RRH assignment, and afterwards applying the optimal PA described in [29]. Therefore, the complexity of SRRH-OPA equals that of SSRH-LPO added to the complexity of optimal PA which is discussed next.

Following the optimal power formulation provided in [29], the relaxed version of the problem is as follows:
Let K_n be the set of multiplexed users on subcarrier n, $\mathcal{N}_M$ the set of multiplexed subcarriers, $\mathcal{S}^{sole}$ the set of sole subcarriers with $N^{sole} = |\mathcal{S}^{sole}|$, $k_1(n)$ the first user over the subcarrier n, where n is either a sole or a multiplexed subcarrier, $k_2(n)$ the second user over the subcarrier n, where n is a multiplexed subcarrier, $r(n)$ the RRH powering the signals on the subcarrier n, $R_{k,n,r}$ the rate achieved by user k on subcarrier n powered by RRH r. Using the same rate to power conversion procedure as in [29], the optimization problem can be expressed as follows:

$$\min_{R_{k,n,r}} \sum_{n \in \mathcal{N}_M \cup \mathcal{S}^{sole}} \frac{a(n)\sigma^2}{h_1(n)} + \frac{(b(n)-1)\sigma^2}{h_2(n)} \left[\frac{1}{h_2(n)} + \frac{a(n)-1}{h_1(n)} \right]$$

subject to:

$$\sum_{n \in S_k} R_{k,n,r(n)} = R_{k,req}, \forall k \in 1:K$$

Where $h_1(n) = h_{k_1(n),n,r(n)}$, $h_2(n) = h_{k_2(n),n,r(n)}$, $\sigma^2 = N_0 B/S$, $a(n) = 2^{R_{k_1(n),n,r(n)} S/B}$, and $b(n) = 2^{R_{k_2(n),n,r(n)} S/B}$. $R_{k_1(n),n,r(n)}$ is the rate achieved by the strong or sole user $k_1(n)$ on subcarrier n, and $R_{k_2(n),n,r(n)}$ is the rate delivered on the subcarrier n to the user $k_2(n)$. If n happens to be a sole subcarrier, then $R_{k_2(n),n,r(n)}$ is null. The Lagrangian of this problem is given by:

$$L(R_{k,n,r}, \lambda) = \sum_{n \in \mathcal{N}_M \cup \mathcal{S}^{sole}} (a(n)-1) \frac{\sigma^2}{h_1(n)} - \sum_{k=1}^{K} \lambda_k \left(\sum_{n=1}^{N} R_{k,n,r(n)} - R_{k,req} \right) + \left[\frac{(a(n)-1)\sigma^2}{h_1(n)} + \frac{\sigma^2}{h_2(n)} \right] \frac{b(n)-1}{h_2(n)}$$

After applying the KKT conditions, and including the K rate constraints, we obtain a system of $N_{sole} + 2card(\mathcal{N}_M) + K$ non-linear equations and unknowns ($N_{sole} + 2card(\mathcal{N}_M)$ rate variables and K Lagrangian multipliers). A numerical solver is used to determine

the solution, namely the *trust-region dogleg* method. Since finding an exact expression of this method's complexity is cumbersome, we propose to provide instead the average execution time ratio of SRRH-OPA with respect to SRRH-LPO, measured over a total of 1000 simulations at a rate of 12Mbps, for $K = 15$ users, $S = 64$ subcarriers and $R = 4$ RRHs. We observed that the execution time of SRRH-OPA is more than the double the one of SRRH-LPO, while the performance improvement is of only 2%. This showcases the efficiency of our LPO procedure, both in terms of its global optimal-like performance and in terms of its cost effective implementation.

Chapter 3

NOMA Mutual SIC for Power Minimization in Hybrid Distributed Antenna Systems

In the previous chapter, we studied the power minimization problem in the DAS context using NOMA, and we proposed several RA techniques that tackle the power minimization problem under user target rate requirements. In this chapter, we consider the power minimization problem in Hybrid DAS (HDAS) where antennas are supplied by various – low power and high power – energy sources. Antenna-specific power limits are taken into account and the problem is reformulated in this new HDAS scenario. After presenting the system model in section 3.2, the optimal PA for OMA with fixed subcarrier-RRH assignment is derived in section 3.3. This allows the design of adequate OMA RA techniques in section 3.4, and NOMA RA techniques in section 3.5. A comparative complexity analysis is conducted in section 3.6 for the proposed RA schemes, and their performance assessment is undergone in section 3.7. Finally, the conclusions are drawn in section 3.8. The major contributions of this chapter can be summarized as follows:

- We provide a deep thorough analysis of the optimal PA for the context of HDAS, showcasing the unique properties it exhibits with respect to classical unconstrained DAS, and highlighting the major differences in the obtained solutions.
- We derive the set of sufficient conditions on channel assignment and user-antenna pairing that guarantees the existence of a solution to satisfy the user rate requirements on the one hand, and the antenna power limits on the other hand.
- We provide two different approaches for joint channel and power allocation in both OMA and NOMA. One approach is more suited for harsh system conditions (in terms of required rates and total power limits), while the other is more effective for mild system conditions.

3.1 Related Works

The deployment of antennas throughout the cell in DAS allowed for greater coverage and enhanced signals strength by reducing the mean distance between users and their serving antennas. The new distributed cell architectures provide a robust framework to combat

inter-cell and intra-cell interference. These advantages have a green ecological impact as cells will be able to provide users with their requested services by utilizing the advantages of the network topology rather than resorting to an increase in the system transmission power.

However, the relative geographic proximity between the users and the antennas in DAS may give rise to more restrictive regulations on antenna power limits than before, in order to limit the electromagnetic field exposure, especially in sensitive locations in dense urban areas (e.g. around hospitals, police stations, etc.). Therefore, in such hybrid configurations, certain antennas in the cell may have restrictive transmit power constraints, e.g. due to their geographical position, their powering source or their small size, while others have access to a much higher amount of available power. The development of procedures which can deal with such different restrictions goes along with the philosophy of 5G and beyond communications in designing new smart networks that can dynamically adapt to various network demands and configurations. These procedures would also come in handy in situations where the operators use hybrid sources of energy to power the antennas deployed at different locations in the cell, including electric grids, local generators and various energy harvesting techniques. These different scenarios leading to power constrained antennas will be referred to from hereinafter as HDAS.

Several works in HDAS target the optimization of system EE with a power constraint on each RRH. In [102], the authors propose antenna selection as a means to maximize the EE of communication systems by successively activating antennas with a decreasing order in added efficiency. However, in a multi-carrier system where frequency selective channels are experienced by users, the possibility to use or not a particular antenna can be extended to each of the possible system subcarrier. In [88], SA and PA are done in two separate stages. In the first stage, the number of subcarriers per RRH is determined, and subcarrier/RRH assignment is performed assuming initial equal power distribution. In the second stage, optimal PA relying on the sub-gradient method is performed to maximize the EE under the constraints of the total transmit power per RRH, of the targeted bit error rate and of a proportionally-fair throughput distribution among active users. In [103], optimal PA is derived for EE maximization under antenna power constraints and proportionally fair user rates in a downlink MISO system. Differently from [88], a single-variable non-linear equation needs to be solved. However the resource allocation problem in its integrality is not addressed since the joint subchannel and power allocation is not studied. Moreover, no insights are inferred from the obtained PA towards the design of efficient user-channel assignment policies. The optimization techniques proposed in [88, 102, 103] for HDAS are designed for the case of OMA. In other words, they allow the allocation of only one user per subcarrier.

In NOMA, multiple users are enabled to access the same time-frequency block through multiplexing in the power domain. The power multiplexing scheme is coupled with SIC receivers to mitigate inter-user interference and enhance the system spectral efficiency. In CAS, the decoding order of downlink NOMA used to be determined according to the descending order of channel gains [27, 28, 104, 105]. When combining the study of NOMA with DAS, we showed in the previous chapter that under some specific subcarrier, user and powering antenna configurations, the two paired users on a subcarrier would be able to perform SIC. Based on this property, we developed techniques for joint subcarrier

and power allocation that aim at minimizing the total amount of power under user rate constraints in downlink NOMA. To the best of our knowledge, the problem of downlink power minimization in DAS networks with RRH power limits using NOMA has not been addressed yet. This problem is substantially different from the one in chapter 2 since heavily loaded antennas are generally the most important players in minimizing the system power. Thus, setting power limits on some of them will necessarily raise the system power consumption. To address this problem, we derive the optimal PA scheme for OMA (given a predefined subcarrier assignment) and explore thoroughly its properties, prior to introducing complete RA schemes that meet the system requirements based on the optimal PA, in both orthogonal and non-orthogonal scenarios.

3.2 System Description and Problem Formulation

The study is conducted on a downlink system consisting of a total of R RRHs uniformly positioned over a cell, where K single-antenna mobile users are randomly deployed. Each RRH is equipped with a single antenna, therefore, the terms "RRH" and "antenna" will be used interchangeably. Among these R RRHs, we consider a subset $\mathcal{RL} = \{RL_1, RL_2, \ldots, RL_F\}$ of $F < R$ power-limited (or constrained) antennas having each a respective power limit $P_{m_i}, i = 1, \ldots, F$, constituting the set $\mathcal{P} = \{P_{m_1}, P_{m_2}, \ldots, P_{m_F}\}$. The remaining $R - F$ RRHs have power limits much higher than those in $\mathcal{RL}$, that is why their power constraints will not be considered in the following. These antennas constitute the set $\mathcal{RU} = \{RU_1, RU_2, \ldots, RU_{R-F}\}$ of unconstrained antennas. All RRHs are connected to a single BBU through high capacity optical fibers and selection diversity [5] is assumed. The system bandwidth B is equally divided into a total of S subcarriers. Each user k is allocated a set $\mathcal{S}_k$ of subcarriers by the BBU in a way to achieve a requested rate $R_{k,req}$ [bps]. From the set of K users, a maximum of $m(n)$ users are chosen to be collocated on the n^{th} subcarrier ($1 \leq n \leq S$) using PD-NOMA [27, 106]. Classical OMA signaling corresponds to the special case of $m(n) = 1$. Also, in the sequel, we denote by DAS the system where $F = 0$ (i.e. none of its RRHs has a power limitation), and by HDAS the case where $F \neq 0$.

The hybrid distributed antenna system is illustrated in Fig. 3.1 where orthogonal signaling is used to serve User 2 on subcarrier SC 3, and non-orthogonal signaling is used to serve Users 1, 2 and 3 on subcarriers SC 1 and SC 2 from both $\mathcal{RL}$ and $\mathcal{RU}$ antennas. Let $P_{k,n,r}$ be the power of user k on subcarrier n, transmitted by RRH r, $\mathbb{H}$ the three-dimensional squared channel gain matrix with elements $h_{k,n,r}$, $1 \leq k \leq K$, $1 \leq n \leq S$, $1 \leq r \leq R$, $k_i(n)$ the i^{th} multiplexed user on subcarrier n, $r_i(n)$ the antenna powering the signal of the i^{th} user on subcarrier n, and $\mathcal{S}(RL_i)$ the set of subcarriers powered by the i^{th} antenna in $\mathcal{RL}$. At each receiver side, additive white Gaussian noise is assumed with a power spectral density N_0, leading to the same average noise power $\sigma^2 = N_0 B/S$ on each subcarrier. In this study, we limit the number of collocated users to a maximum of 2 per subcarrier, which limits the SIC complexity at the receiver side at the cost of a negligible performance drop, compared to 3 collocated users, as it was shown in [27].

When the same antenna is used to power the signals of collocated users on a subcarrier (e.g. User 1 and User 2 on subcarrier SC 1 in Fig. 3.1), the user with higher channel gain decodes, re-modulates and subtracts the signal of the weak user, whereas the weak user suffers from the interference caused by the signal of the strong user. Therefore, the rate expressions and PMCs of two collocated users k_1 and k_2 on subcarrier n with

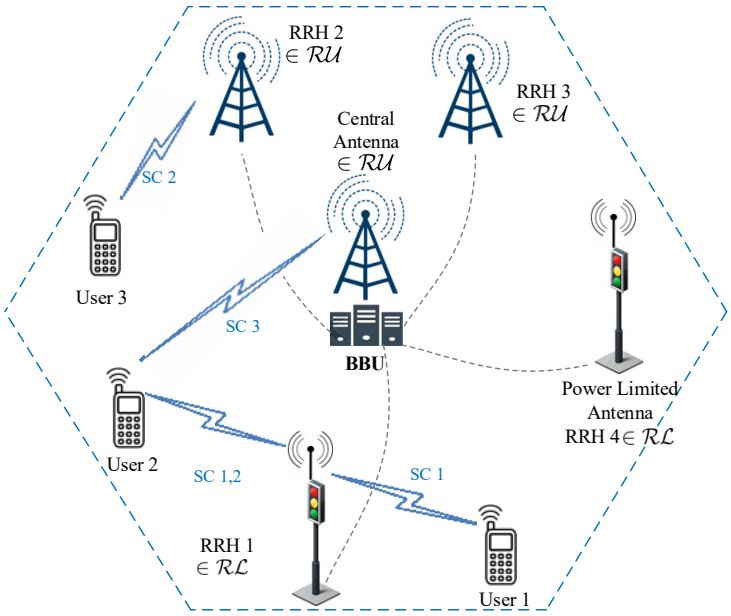


Figure 3.1 – HDAS cell with two power-limited RRHs (RRH 1 and RRH 4).

$h_{k_1,n,r} > h_{k_2,n,r}$ and $r = r_1(n) = r_2(n)$ for classical NOMA schemes are:

$$\begin{aligned} R_{k_1,n,r} &= \frac{B}{S}\log_2\left(1+\frac{P_{k_1,n,r}h_{k_1,n,r}}{\sigma^2}\right), \\ R_{k_2,n,r} &= \frac{B}{S}\log_2\left(1+\frac{P_{k_2,n,r}h_{k_2,n,r}}{P_{k_1,n,r}h_{k_2,n,r}+\sigma^2}\right), \\ & P_{k_2,n,r} > P_{k_1,n,r}. \end{aligned} \tag{3.1}$$

On the other hand, when the signal of the multiplexed users k_1 and k_2 on a subcarrier n and transmitted by two different RRHs $r_1(n)$ and $r_2(n)$ respectively (e.g. User 2 and User 3 on subcarrier SC 2 in Fig. 3.1), mutual SIC can be performed if the user channel gains verify:

$$\frac{h_{k_1,n,r_1(n)}}{h_{k_1,n,r_2(n)}} \leq \frac{h_{k_2,n,r_1(n)}}{h_{k_2,n,r_2(n)}}.$$

In such cases, their theoretical throughputs and power multiplexing constraints are given by:

$$\begin{aligned} R_{k_1,n,r_1(n)} &= \frac{B}{S}\log_2\left(1+\frac{P_{k_1,n,r_1(n)}h_{k_1,n,r_1(n)}}{\sigma^2}\right), \\ R_{k_2,n,r_2(n)} &= \frac{B}{S}\log_2\left(1+\frac{P_{k_2,n,r_2(n)}h_{k_2,n,r_2(n)}}{\sigma^2}\right), \\ \frac{h_{k_1,n,r_1(n)}}{h_{k_1,n,r_2(n)}} &\leq \frac{P_{k_2,n,r_2(n)}}{P_{k_1,n,r_1(n)}} \leq \frac{h_{k_2,n,r_1(n)}}{h_{k_2,n,r_2(n)}}. \end{aligned} \tag{3.2}$$

The aim in this chapter is to derive joint subcarrier and power allocation as well as user-pairing schemes that minimize the total transmit power while meeting the rate requirement of each user ($R_{k,req}$) and the power limit constraints on the $\mathcal{RL}$ antennas ($\mathcal{P}$). The introduction of power limit constraints on a subset of RRHs will lead to a more power

consuming solution than the one obtained previously in chapter 2, since the addition of any new constraint to an optimization problem may only result in the degradation of the solution's performance. Through comparison to the problem in chapter 2, having power constraints on some antennas results in a power transfer from the constrained antennas to the unconstrained ones, in such a way that the requested rates remain satisfied for each user. Hence, minimizing the total transmit power of the system under the user rate and antenna power limit constraints translates into searching for the best "power transfer" scheme that minimizes the excess in power compared to the unconstrained DAS solutions in chapter 2. Note that the number of constrained antennas F shall not reach R, that is to say that at least one antenna has to remain unconstrained in order to guarantee the satisfaction of the requested rate for all users. The global optimization problem of user-subcarrier-RRH assignment and PA, taking into account the rate requirements, power limits, and NOMA PMCs, can be formulated as:

$$\mathcal{OP}_1 : \{\mathcal{S}_k, P_{k,n,r}\}^* = \underset{\mathcal{S}_k, P_{k,n,r}}{\arg\min} \sum_{k=1}^{K} \sum_{n \in \mathcal{S}_k} \sum_{\substack{i=1, \\ \text{s.t.} k_i(n)=k}}^{2} P_{k,n,r_i(n)},$$

subject to :

$$\sum_{\substack{n \in \mathcal{S}_k \\ \text{s.t. } k_i(n)=k,\ i=\{1,2\}}} R_{k_i,n,r_i(n)} = R_{k,req}, \forall k, 1 \leq k \leq K, \tag{3.3}$$

$$\sum_{n \in \mathcal{S}(RL_j)} \sum_{\substack{i=1, \\ r_i(n)=RL_j}}^{2} P_{k_i(n),n,RL_j} \leq P_{m_j}, \forall j, 1 \leq j \leq F, \tag{3.4}$$

$$\forall n \in \{1, \ldots, S\}, \text{ s.t. } m(n) = 2 \begin{cases} (3.2), r_1(n) \neq r_2(n) & (3.5) \\ (3.1), r_1(n) = r_2(n). & (3.6) \end{cases}$$

As in the previous chapter, the problem at hand involves set selection as well as continuous variable optimization, hence its mixed-integer non-convex nature justifies the introduction of suboptimal schemes. Therefore, we follow the same approach for tackling the RA problem by first deriving the optimal PA for power minimization in the OMA context and for a fixed SA (section 3.3). Then the optimal PA properties lead to the elaboration of RA schemes for OMA (section 3.4) and NOMA (section 3.5).

3.3 Optimal Power Allocation for OMA HDAS

In the orthogonal scenario, every subcarrier n is allocated to one user and one antenna at most, referred to as $k(n)$ and $r(n)$ respectively. The optimal PA scheme for a predefined subcarrier allocation scheme is cast as the solution to the following problem:

$$\mathcal{OP}_2 : \{P_{k,n,r}\}^* = \min_{\{P_{k,n,r}\}} \sum_{k=1}^{K} \sum_{n \in \mathcal{S}_k} P_{k,n,r(n)},$$

subject to:

$$\sum_{n\in\mathcal{S}_k} \frac{B}{S}\log_2\left(1+\frac{P_{k,n,r(n)}h_{k,n,r(n)}}{\sigma^2}\right) = R_{k,req}, \forall k, 1\le k\le K, \tag{3.7}$$

$$\sum_{n\in\mathcal{S}(RL_i)} P_{k(n),n,RL_i} \le P_{m_i}, \forall i, 1\le i\le F. \tag{3.8}$$

The problem in hand can be solved by means of standard convex optimization techniques, its Lagrangian is given by:

$$\begin{aligned} L(P_{k,n,r},\lambda_k,\alpha_i) = &-\sum_{k=1}^{K}\sum_{n\in\mathcal{S}_k} P_{k,n,r(n)} + \sum_{k=1}^{K}\lambda_k\left(R_{k,req}-\sum_{n\in\mathcal{S}_k}\frac{B}{S}\log_2\left(1+\frac{P_{k,n,r(n)}h_{k,n,r(n)}}{\sigma^2}\right)\right) \\ &+\sum_{i=1}^{F}\alpha_i\left(P_{m_i}-\sum_{n\in\mathcal{S}(RL_i)} P_{k(n),n,RL_i}\right), \end{aligned} \tag{3.9}$$

where λ_k and α_i represent the Lagrangian multipliers relative to the rate and power constraints respectively. The corresponding KKT conditions are:

$$\begin{gathered} \nabla L(P^*_{k,n,r(n)},\lambda^*_k,\alpha^*_i) = 0, \\ \sum_{n\in\mathcal{S}_k}\frac{B}{S}\log_2\left(1+\frac{P_{k,n,r(n)}h_{k,n,r(n)}}{\sigma^2}\right) = R_{k,req}, \forall k, 1\le k\le K, \\ \sum_{n\in\mathcal{S}(RL_i)} P_{k(n),n,RL_i} \le P_{m_i}, \forall i, 1\le i\le F, \\ \alpha_i\left(\sum_{n\in\mathcal{S}(RL_i)} P_{k(n),n,r(n)} - P_{m_i}\right) = 0, \forall i, 1\le i\le F, \\ \alpha_i \ge 0, \forall i, 1\le i\le F. \end{gathered}$$

The expressions of the partial derivatives of L with respect to the power variables change according to whether the powering RRH is constrained or not. For subcarriers powered by a constrained antenna, we get:

$$\frac{\partial L}{\partial P_{k,n,r(n)}} = -1 - \frac{B\lambda_k}{S\ln(2)}\frac{h_{k,n,r(n)}/\sigma^2}{1+P_{k,n,r(n)}h_{k,n,r(n)}/\sigma^2} - \alpha_i = 0.$$

By setting $m_k = -B\lambda_k/S\ln(2)$, we have:

$$\frac{m_k}{\frac{\sigma^2}{h_{k,n,r(n)}}+P_{k,n,r(n)}} = 1+\alpha_i. \tag{3.10}$$

The subcarriers that are not powered by a constrained antenna do not feature an α_i term as in (3.10). Instead, their partial derivative yields:

$$\frac{m_k}{\frac{\sigma^2}{h_{k,n,r(n)}}+P_{k,n,r(n)}} = 1. \tag{3.11}$$

Equations (3.10) and (3.11) can be rearranged in the following form:

$$P_{k,n,r(n)} = \frac{m_k}{(1+\alpha_i)} - \frac{\sigma^2}{h_{k,n,r(n)}}, \forall n \in \mathcal{S}(RL_i), \tag{3.12}$$

$$P_{k,n,r(n)} = m_k - \frac{\sigma^2}{h_{k,n,r(n)}}, \forall n \notin \mathcal{S}(\mathcal{RL}), \tag{3.13}$$

where $\mathcal{S}(\mathcal{RL}) \triangleq \cup_{i=1}^{F} \mathcal{S}(RL_i)$ is the set of all subcarriers powered by a power constrained antenna. If $\mathcal{OP}_1$ was considered without the power constraints (3.8), the solution of the PA problem would revert to the classical case of power minimization with user-specific rate constraints, resulting in a user-specific waterfilling where the waterline m_k of user k would be the same for all of its subcarriers. However, in HDAS, the waterline becomes specific to the separate classes of subcarriers, grouped according to the transmitting antenna. More specifically, equations (3.12) and (3.13) show that, for every user k, a specific waterlevel $m_{ki} \triangleq m_k/(1+\alpha_i)$ is assigned for every subset of subcarriers allocated to k on the constrained antenna RL_i, whereas the remaining subcarriers of k that are not powered by a constrained RRH share a common waterlevel m_k. Furthermore, all the waterlevels of user k, corresponding to its powering RRHs, are related to m_k by the factors $(1+\alpha_i)$. Replacing the power variables by their expressions from (3.12) and (3.13) in the rate constraints (3.7), and applying some manipulations yield the following forms, for each user k:

$$\sum_{n \in \mathcal{S}_k \cap \overline{\mathcal{S}(\mathcal{RL})}} \log_2\left(m_k \frac{h_{k,n,r(n)}}{\sigma^2}\right) + \sum_{i=1}^{F} \sum_{n \in \mathcal{T}_{ki}} \log_2\left(\frac{m_k h_{k,n,r(n)}}{(1+\alpha_i)\sigma^2}\right) = \frac{R_{k,req}S}{B},$$

$$\sum_{n \in \mathcal{S}_k \cap \overline{\mathcal{S}(\mathcal{RL})}} \log_2\left(m_k \frac{h_{k,n,r(n)}}{\sigma^2}\right) + \sum_{n \in \cup_{i=1}^{F} \mathcal{T}_{ki}} \log_2\left(m_k \frac{h_{k,n,r(n)}}{\sigma^2}\right) - \sum_{i=1}^{F} |\mathcal{T}_{ki}| \log_2(1+\alpha_i) = \frac{R_{k,req}S}{B},$$

where $\mathcal{T}_{ki} \triangleq \mathcal{S}_k \cap \mathcal{S}(RL_i)$ is the set of subcarriers allocated to user k and powered by RL_i. Consequently, we obtain:

$$\sum_{n \in \mathcal{S}_k} \log_2\left(m_k h_{k,n,r(n)}/\sigma^2\right) - \sum_{i=1}^{F} |\mathcal{T}_{ki}| \log_2(1+\alpha_i) = \frac{R_{k,req}S}{B}.$$

Therefore, m_k can be written as:

$$m_k = \left(\frac{2^{\frac{R_{k,req}S}{B}}}{\prod_{n \in \mathcal{S}_k} h_{k,n,r(n)}/\sigma^2} \prod_{i=1}^{F} (1+\alpha_i)^{|\mathcal{T}_{ki}|}\right)^{1/|\mathcal{S}_k|},$$

$$m_k = W_k \prod_{i=1}^{F} (1+\alpha_i)^{\frac{|\mathcal{T}_{ki}|}{|\mathcal{S}_k|}}. \tag{3.14}$$

Recall from (2.10) that W_k is the common waterline that user k would have had in a "classical" waterfilling scheme, i.e. if the power constraints on $\mathcal{RL}$ were not taken into account ($m_k = W_k$ when $\alpha_i = 0, \forall i$). Therefore, α_i can be seen as the power correction factors that are applied to the unconstrained waterfilling solution to obtain the new power

pouring solution in (3.12) and (3.13). Note that if user k has all its subcarriers powered by non-constrained antennas ($\mathcal{T}_{ki} = \emptyset$, $1 \leq i \leq F$), then the allocated power to the user subcarriers is according to (3.13). The user will have a unique waterline m_k for all of its subcarriers, and $m_k = W_k$ since $|\mathcal{T}_{ki}| = 0, 1 \leq i \leq F$. Such a user is not affected by the power correction. Also, if a unique constrained RRH RL_i is exclusively serving a user ($\mathcal{T}_{ki} = \mathcal{S}_k$, and $\forall j \neq i, \mathcal{T}_{ki} = \emptyset$), the power of all of its subcarriers will be according to (3.12), meaning that a unique waterline is assigned to the user and it is given by $m_k/(1+\alpha_i)$. But we have from (3.14) $m_k = W_k(1+\alpha_i)$, therefore the waterline of the user is equal to W_k and it is not affected by the power correction. At last, another possibility for having unique waterlines (per user) resides in a system where the classical waterfilling solution abides by (3.8). Indeed, if all the α_i variables were null, the resulting Lagrangian would not account for the power constraints (3.8), hence the solution would be a simple user-based waterfilling: if $\alpha_i = 0$, $m_k = \frac{m_k}{(1+\alpha_i)} = W_k$ which results in a uniform waterlevel over all the subcarriers of the user k.
As a consequence, if an initial RA technique verifies the constraints on the antennas assuming a user-based waterfilling, this is indeed the best PA solution for the given SA. On the other hand, the proposed optimal PA technique can be applied in association with any SA scheme that constitutes a solution to problem $\mathcal{OP}_1$. Solving the power optimization problem reduces to determining the F Lagrangian variables α_i relative to the power constraints. By replacing (3.14) into (3.12), the power constraint (3.8), corresponding to $\alpha_i \neq 0$ in the KKT conditions, for the i^{th} antenna RL_i in $\mathcal{RL}$, can be written as:

$$\sum_{n \in S(RL_i)} \left(\frac{W_{k(n)} \prod_{j=1}^{F} (1+\alpha_j)^{\frac{|\mathcal{T}_{k(n)i}|}{|\mathcal{S}_{k(n)}|}}}{(1+\alpha_i)} - \frac{\sigma^2}{h_{k(n),n,r(n)}} \right) = P_{m_i}. \tag{3.15}$$

Equation (3.15) consists of F non-linear equations with unknowns α_i. In the sequel, the case of a single power-limited antenna ($F = 1$) is considered first in order to provide a clear analysis of the hybrid system behavior. Then, the generalized study for higher values of F is developed.

3.3.1 Single Power-Limited Antenna

For the special case of a single power-limited antenna, we simply denote by RL the considered RRH and α the Lagrangian variable relative to the corresponding power constraint. For each user, we can identify at most two sets of subcarriers and thus two waterlevels which are related by the factor $(1+\alpha)$. The waterlevel of the subcarrier set that is not powered by the constrained antenna RL is obtained from (3.14) as:

$$m_k = (1+\alpha)^{\frac{|\mathcal{T}_k|}{|\mathcal{S}_k|}} W_k. \tag{3.16}$$

This equation shows how the introduction of the constraint on one of the antennas affects the PA scheme, compared to the non-constrained case: since $|\mathcal{T}_k|/|\mathcal{S}_k| \leq 1$, and $\alpha > 0$, the waterline of the subcarriers in $\mathcal{T}_k$ decreases with respect to W_k (since $m_k/(1+\alpha) < W_k$), while that of the subcarriers in $\mathcal{S}_k \cap \overline{\mathcal{T}_k}$ increases (since $m_k > W_k$). This behavior is depicted in Fig. 3.2.

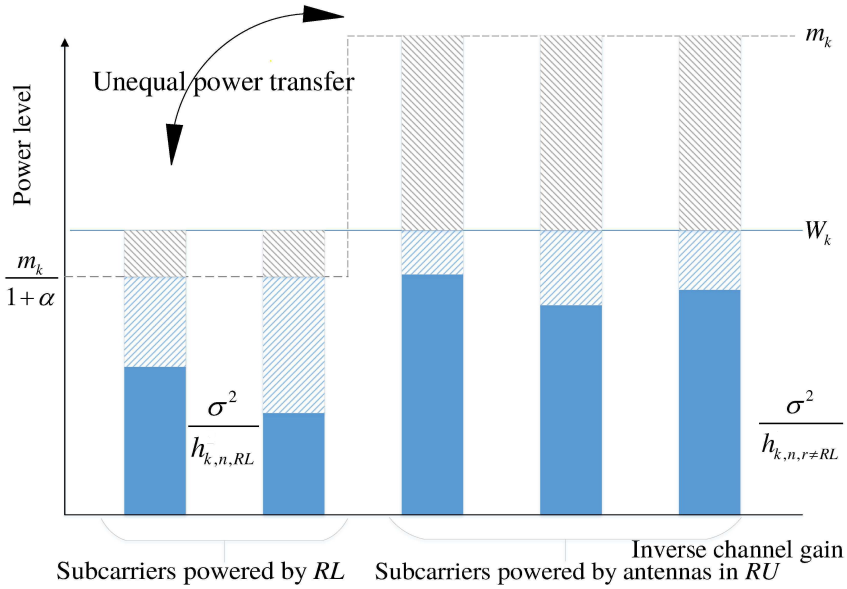


Figure 3.2 – Power pouring diagram for a user k after power correction.

When the PA solution of the unconstrained problem does not respect (3.8), then a power correction using (3.12) and (3.13) is necessary, and the rate transfer from the constrained antenna to the unconstrained ones translates into an unbalanced power transfer from antenna RL to the other antennas. This gives further justification to why α can be seen as the deviation factor from the unconstrained problem. A greater value of α translates into a greater deterioration of the performance of the solution towards that of the unconstrained problem, meaning a more important increase of the total power in HDAS compared to DAS.

For $F = 1$, the system of equations in (3.15) reduces to a single equation with a unique unknown α:

$$\sum_{n\in\mathcal{S}(RL)} \left(W_{k(n)}(1+\alpha)^{\frac{|\mathcal{T}_{k(n)}|}{|\mathcal{S}_{k(n)}|}-1} - \frac{\sigma^2}{h_{k(n),n,r(n)}} \right) = P_m. \tag{3.17}$$

There is no *a priori* guarantee for the existence of a solution to (3.17). An example of a situation with no solution is when every user served by RL is exclusively linked to RL. Keeping in mind that such users are not affected by the power correction, if their total power consumption is greater than P_m, then no PA could, at the same time, verify the antenna power constraint and provide the users the rates they are requesting. Therefore, it is of interest to assess the feasibility of a proposed SA before proceeding to the resolution of (3.17) through a numerical solver. By isolating the terms α from the others, (3.17) takes the following form:

$$\begin{gathered}\sum_{\substack{n\in\mathcal{S}(\mathcal{RL}),\\ \mathcal{T}_{k(n)}\neq\mathcal{S}_{k(n)}}} W_{k(n)}(1+\alpha)^{\frac{|\mathcal{T}_{k(n)}|}{|\mathcal{S}_{k(n)}|}-1} = P_m - \sum_{\substack{n\in\mathcal{S}(RL),\\ \mathcal{T}_{k(n)}=\mathcal{S}_{k(n)}}} W_{k(n)} + \sum_{n\in\mathcal{S}(RL)} \frac{\sigma^2}{h_{k(n),n,r(n)}}, \\ A(\alpha) = C.\end{gathered} \tag{3.18}$$

The first sum is a function of α, hence the notation $A(\alpha)$. It includes all the subcarriers in $\mathcal{S}(RL)$ belonging to users that are served by at least one non-constrained antenna.

C is constant (for a fixed subcarrier allocation) and accounts for: the power limit, the waterlines relative to the subcarriers of the users exclusively served by RL, and finally, the inverse channel gains of all subcarriers powered by RL. In order to have a solution, C needs to belong to the image of A when α spans the positive real axis. However, $A(\alpha)$ is a polynomial with negative, fractional exponents and positive coefficients. Thus, it is a strictly decreasing function of α, its co-domain is $]\lim_{\alpha\to\inf} A(\alpha) = 0; A(0)]$. Therefore, the condition that guarantees the existence of a solution is: $0 < C \leq A(0)$.

Proposition 3.1. *If the system requires a power correction, C will be necessarily smaller than $A(0)$.*

Proof. The left hand side of (3.17) is the power on RL for a given value of α. When $\alpha = 0$, and since the system requires correction, this power is the actual power of RL before any power correction takes place. This value is greater than P_m, that is $\sum_{n\in\mathcal{S}(RL)} \left(W_{k(n)} - \sigma^2/h_{k(n),n,r(n)}\right) \geq P_m$. By setting α to 0 in (3.18), we directly obtain $A(0) - C \geq 0$, i.e. $A(0) \geq C$ which concludes our proof. □

As a result, the existence of a solution is only conditioned by C being strictly positive. Finally, the uniqueness of the solution is an immediate result of the monotonic nature of function A.

3.4 Resource Allocation for HDAS using OMA

Having established the main properties and conditions of the optimal power allocation, we now seek efficient resource allocation schemes that meet the rate and power limit requirements while minimizing the total power. In the following, two different approaches are proposed to resolve $\mathcal{OP}_1$ in the OMA context: OMA-HDAS and OMA-HDAS-Realloc. They both aim at determining the subcarrier and PA schemes that minimize the overall power, while guaranteeing the power and rate allocation constraints. OMA-HDAS takes into consideration the antenna power constraints at the end of the algorithm, while OMA-HDAS-Realloc accounts for the loading of the constrained antennas throughout the algorithm.

3.4.1 The OMA-HDAS Approach

In the case of a single constrained antenna, a success-guaranteed RA scheme is one that ensures the positivity of C. The negativity of C refers to the scenario where satisfying the constraints of $\mathcal{OP}_1$ is impossible because the users that are solely served by RL require a higher power than P_m to reach their target rate. Since the power of such users is not affected by the power correction, it is easy to determine why $\mathcal{OP}_1$ is not feasible in this scenario. By extension to the general case ($F > 1$), the PA problem is not feasible when the requested power of users served exclusively by RL_i is greater than P_{m_i}, for any antenna in $\mathcal{RL}$. Therefore, one sufficient condition enabling the resolution of $\mathcal{OP}_1$ resides in removing the negative term from the right hand side of (3.18). This is achieved by imposing:

$$|\mathcal{T}_{ki}| < |\mathcal{S}_k|, \forall k, \forall i, 1 \leq i \leq F. \tag{3.19}$$

In other terms, a sufficient condition for a success-guaranteed RA scheme is to have every user served by $\mathcal{RL}$ allocated at least one subcarrier powered by a non-constrained

antenna. This is ensured by modifying the WBH phase as shown in algorithm 3.1 where every user is assigned a subcarrier-RRH pair from $\mathcal{RU}$ instead of $\mathcal{RL} \cup \mathcal{RU}$. Then, the power minimization strategy developed for OMA in algorithm 2.2 of section 2.4.2 is applied. Finally, the state of the $\mathcal{RL}$ antennas is checked: if a power level higher than the imposed limit is detected, the optimal PA described in section 3.3 is applied to perform power correction. The details of OMA-HDAS are presented in algorithm 3.1 where $\mathcal{S}_f$ represents the set of allocated subcarriers, $\mathcal{S}_p$ the set of free subcarriers, and $\mathcal{U}_0$ the set of active users in the modified WBH phase.

Algorithm 3.1 OMA-HDAS

Initialization: $\mathcal{S}_p = [1:S], \mathcal{U}_0 = [1:K], \mathcal{S}_f = \emptyset$
Phase 1: Modified Worst-Best-H
while $\mathcal{U}_0 \neq \emptyset$ **do**
$\forall k \in \mathcal{U}_0 : (n\_\max_k, r\_\max_k) = \arg\max_{n \in \mathcal{S}_p, r \in \mathcal{RU}} (h_{k,n,r})$
$k^* = \arg\min_{k \in \mathcal{U}_0} h_{k,n\_\max_k, r\_\max_k}$
$n = n\_\max_{k^*}; r = r\_\max_{k^*}$
$P_{k^*,n,r} = \sigma^2 (2^{R_{k^*,req} S/B} - 1)/h_{k^*,n_1,r_1}$, $P_{k^*,tot} = P_{k^*,n,r}$,
$\mathcal{S}_{k^*} = \mathcal{S}_{k^*} \cup \{n\}; \mathcal{S}_f = \mathcal{S}_f \cup \{n\}; \mathcal{S}_p = \mathcal{S}_p \cap \{n\}^c$,
$\mathcal{U}_0 = \mathcal{U}_0 \cap \{k^*\}^c$
end while
Phase 2: Orthogonal multiplexing // as in algorithm 2.2
Phase 3: Power correction
if $\exists i \in \{1, \ldots, F\}/P_{RL_i} > P_{m_i}$
Apply the power correction using (3.15), (3.12) and (3.13)
end if

The main advantages of OMA-HDAS are its relative simplicity and its guarantee for providing a solution to the system. However, separating the subcarrier-RRH assignment from the correction phase is far from optimal since a beneficial subcarrier-RRH allocation on $\mathcal{RL}$ in the first two phases of algorithm 3.1 may turn out to be penalizing after power correction. In fact, when no special care is given in the subcarrier-RRH allocation to account for the subsequent power correction, a great load may result on $\mathcal{RL}$, rendering the toll of the correction unacceptable. For instance, the power increase incurred by the power correction could be such that turning off the constrained antennas and applying the power minimization procedure of OMA-DAS (algorithm 2.2, section 2.4.2, chapter 2) over the unconstrained ones would be more profitable. This method is referred to as OMA-SOFF and will serve as a higher bound benchmark on the power consumption in DAS. To tackle the issues concerning OMA-HDAS, we propose in the next section a new approach for solving $\mathcal{OP}_1$.

3.4.2 The OMA-HDAS-Realloc Approach

To overcome the aforementioned drawbacks of OMA-HDAS, we seek an RA scheme that can systematically outperform the trivial solution of OMA-SOFF. For this purpose, the current algorithm undergoes two phases prior to the power correction. First, OMA-SOFF is applied: the constrained antennas are virtually shut off and the OMA-DAS power minimization technique is applied over $\mathcal{RU}$. In the second phase, the solution is

enhanced by gradually moving some subcarriers from $\mathcal{RU}$ to $\mathcal{RL}$, thus exploiting the better links available through the $\mathcal{RL}$ antennas. To do so, the most power demanding user k is selected and its subcarriers are considered for a potential reallocation.
To determine the subcarrier whose reallocation is the most profitable to user k, let us consider $r_{old} = r(n)$ the antenna powering the subcarrier n before reallocation, and r_{new} the candidate RRH considered for reallocation. To simplify the analysis, their corresponding channel gains are denoted by $h_{old} = h_{k,n,r_{old}}$ and $h_{new} = h_{k,n,r_{new}}$ respectively.

Proposition 3.2. *The subcarrier leading to the greatest power decrease for user k is the one having the highest ratio h_{new}/h_{old}, and the selected RRH is the one providing the largest channel gain on the selected subcarrier.*

Proof. The reallocation of a subcarrier can be decomposed into two consecutive steps: the removal of the subcarrier from the user, and its allocation to the user while being powered by a new antenna. If W_k is the waterline of user k prior to the reallocation, $W_k^{'}$ the intermediate waterline after the subcarrier removal, and $W_{k,new}$ the waterline after completing the reallocation, $W_{k,new}$ can be obtained from W_k based on the iterative waterline relation in (2.11) as follows:

$$\left.\begin{aligned} W_k^{'|\mathcal{S}_k|-1} &= \frac{W_k^{|\mathcal{S}_k|} h_{old}}{\sigma^2} \\ W_{k,new}^{|\mathcal{S}_k|} &= \frac{W_k^{'|\mathcal{S}_k|-1}}{h_{new}/\sigma^2} \end{aligned}\right\} \Rightarrow W_{k,new} = W_k \left(\frac{h_{old}}{h_{new}}\right)^{1/|\mathcal{S}_k|}.$$

The power variation of user k obtained from this potential reallocation is:

$$\Delta P = |\mathcal{S}_k| \left(W_{k,new} - W_k\right). \tag{3.20}$$

The subcarrier to be selected for reallocation must verify:

$$n^* = \arg\min_{n \in \mathcal{S}_k} \Delta P = \arg\min_{n \in \mathcal{S}_k} W_{k,new} = \arg\max_{n \in \mathcal{S}_k} \frac{h_{new}}{h_{old}}.$$

Then, it is straightforward that the selected RRH should be $r_{new} = \arg\max_{r \in \mathcal{RL}}(h_{k,n^*,r})$. This concludes our proof. □

When $h_{new}/h_{old} > 1$, the reallocation is applied and the total power and waterline level of the user are updated. This reallocation process is carried out until leading to an excess in power over every antenna in $\mathcal{RL}$. Note that if a reallocation would render a user without any sole subcarriers powered by $\mathcal{RU}$, then this reallocation is rejected in order to guarantee the existence of a solution to the system according to (3.19). The details of OMA-HDAS-Realloc are presented in algorithm 3.2 where $\mathcal{U}_p$ is the set of active users in the reallocation phase, $\mathcal{R}$ is the set of active antennas in $\mathcal{RL}$ for reallocation, and $\mathcal{S}_k^{\mathcal{RU}}$ is the set of subcarriers of k powered by antennas in $\mathcal{RU}$.

Algorithm 3.2 OMA-HDAS-Realloc

Phase 1: OMA-SOFF
Apply OMA-DAS (algorithm 2.2) but using only antennas in $\mathcal{RU}$
Phase 2: Subcarrier reallocation
Initialization: $\mathcal{U}_p = \{1, \ldots, K\}; \mathcal{R} = \mathcal{RL}; \mathcal{S}_k^{\mathcal{RU}} \triangleq \mathcal{S}_k \cap \overline{\mathcal{S}(\mathcal{RL})} = \mathcal{S}_k, \forall k$
while $\mathcal{U}_p \neq \emptyset$ & $\mathcal{R} \neq \emptyset$ **do**
$k^* = \arg\max_{k \in \mathcal{U}_p} (P_k)$
$(n^*, r^*) = \arg\max_{n \in \mathcal{S}_{k^*}^{\mathcal{RU}}, r \in \mathcal{R}} \left(\frac{h_{k^*,n,r}}{h_{k^*,n,r(n)}} \right)$ // with $r = r_{new}, r(n) = r_{old}$
Estimate ΔP according to (3.20)
if $\Delta P < -\rho$
$P_{k^*} = P_{k^*} + \Delta P$
$\mathcal{S}_{k^*}^{\mathcal{RU}} = \mathcal{S}_{k^*}^{\mathcal{RU}} \cap \{n^*\}^c$
Update $P_{\mathcal{RL}}$
if $|\mathcal{S}_{k^*}^{\mathcal{RU}}| == 1$
$\mathcal{U}_p = \mathcal{U}_p \cap \{k^*\}^c$ // remove k^* from the active user set to hold (3.19)
end if
if $P_{r^*} > P_m$
$\mathcal{R} = \mathcal{R} \cap \{r^*\}^c$ // remove r^* from the set of active antennas
end if
else
$\mathcal{U}_p = \mathcal{U}_p \cap \{k^*\}^c$ // user power can no longer be decreased
end if
end while
Phase 3: Power correction

3.5 Resource Allocation for HDAS using NOMA

To further reduce the system power, the NOMA layer is applied on top of OMA. For this purpose, the user pairing scheme of the Mut&SingSIC technique that was introduced in chapter 2, section 2.5.2.4, is adapted to account for the antenna power limits.

The allocation technique starts with an uncorrected version of the proposed solutions for OMA, then the algorithm tries to pair users in order to reduce system power prior to applying a power correction at a final stage. As previously discussed in chapter 2, each time pairing is performed on a subcarrier, the users powers on this subcarrier are kept unvaried for the subsequent allocation stages. In other words, they will not undergo any further modification in the succeeding PA steps, in order to avoid complex chains of modifications. Due to the power multiplexing constraints of mutual and single SIC subcarriers (3.5), (3.6), the optimal power correction described in section 3.3 for OMA can not be directly applied to the non-orthogonal context. Indeed, since the power allocated to multiplexed subcarriers is constant until the end of the pairing phase, the power correction has to be carried out on the sole subcarriers only (i.e. subcarriers occupied by a unique user). Moreover, the total amount of power on multiplexed subcarriers is deducted from the power limit on each constrained antenna. In other terms, the new power limit on the i^{th} power-constrained RRH is reduced to:

$$P'_{m_i} = P_{m_i} - \sum_{n \in S(RL_i) \text{ s.t } m(n)=2} P_{k,n,RL_i}. \tag{3.21}$$

Therefore, a necessary condition to allow the power correction of the system is to prevent any subcarrier pairing that would lead to a total power of multiplexed subcarriers greater than P_{m_i} for any antenna in $\mathcal{RL}$. To keep track of the total multiplexed power on $\mathcal{RL}$ antennas, we initialize the vector P_{RL_i} of $|\mathcal{RL}|$ elements to zero. For every subcarrier-RRH candidate, the powers $P_{k_2,n,r_2(n)}$ and $P_{k_1,n,r_1(n)}$ of the involved users k_1 and k_2 are added to their corresponding P_{RL_i} elements. If this addition results in an excess on an antenna from $\mathcal{RL}$, the current candidate pair is denied multiplexing. Meanwhile, the power limit on the multiplexed subcarriers per constrained antenna (i.e. the second term in the right-hand part of (3.21)) is set to a fraction β ($0 < \beta < 1$) of P_{m_i}, in order to leave room for power adjustment (correction).
Similarly to the orthogonal scenario, the subcarrier pairing must leave at least one sole subcarrier for each user powered by an RRH in $\mathcal{RU}$ in order to guarantee the existence of a solution to the PA problem. This pairing procedure can be coupled with either OMA-HDAS or OMA-HDAS-Realloc. Note that for the sake of simplicity, we restrict the choices of PA procedures for determining the power on multiplexed subcarriers to LPO (see section 2.5.1) for single SIC subcarriers, and DPA (see section 2.5.2.3) for mutual SIC subcarriers. Finally, the power correction is performed on the sole subcarriers with P'_{m_i} instead of P_{m_i} in (3.15). The details of the complete NOMA algorithms are presented in algorithm 3.3.

3.6 Complexity Analysis

In this section, we assess the complexity of the proposed resource allocation techniques. Just like for the previous chapter, the algorithms consist in sequential blocks of OMA assignment, OMA reallocation, and NOMA pairing, thus the complexity of each independent step is provided then the overall complexity of the proposed algorithms is deduced by combining the corresponding steps.

The complexity of OMA-HDAS is dominated by the matrix reordering of the channel gains for every $K \times R$ pair and the iterative subcarrier-RRH allocation. As shown in chapter 2, section 2.6, it amounts to a complexity of $O(KSR\log(S)+S(K+R))$. In OMA-HDAS-Realloc, the $\mathcal{RL}$ antennas are not used until the reallocation phase, thus $\mathbb{H}$ does not need to be sorted on the corresponding F RRHs. Therefore, the resulting complexity before the reallocation phase is $O(KS(R-F)\log(S)+(K+R-F)S)$.

During the reallocation phase, the most power consuming user is first selected ($O(K)$), then its subcarriers are checked for a potential emission from the $\mathcal{RL}$ antennas. The selected subcarrier satisfies Proposition 3.2 which requires determining the best antenna for every candidate subcarrier. Assuming an equal distribution of the number of subcarriers among users, the complexity of reallocating a single subcarrier is $O(K+FS/K)$. For the worst case of S reallocated subcarriers, the resulting complexity is upper-bounded by $O(S(K+SF/K))$.

Finally, the subcarrier pairing step consists of selecting the most power consuming user which costs $O(K)$, then searching for the subcarrier-RRH pair minimizing the total power ($O(SR)$). The process is repeated a maximum of S times leading to a complexity of $O(S(K+SR))$. At last, the power correction phase is carried when needed with a computational cost denoted by f, which depends on the numerical solver used to resolve the non-linear system in (3.15). Table 3.1 gives the upper bound to the complexity of

Algorithm 3.3 NOMA-HDAS & NOMA-HDAS-Realloc

Phase 1: OMA phase
OMA-HDAS or OMA-HDAS-Realloc but without power correction

Phase 2: User pairing
$\mathcal{U}_p = \{1, 2, \ldots, K\}$
while $\mathcal{U}_p \neq \emptyset$ **do**
  Select the most power consuming user k^*
  Select the couple (n^*, r^*) such that:
    ΔP is minimal
    The total power of multiplexed subcarriers over each antenna in $\mathcal{RL}$ is less than βP_{m_i}
  if $\Delta P < -\rho$
    Apply the user pairing
    Remove the selected subcarrier from $\mathcal{S}_{k^*}^{sole}$
    if Any of the multiplexed users has one remaining sole subcarrier
      Remove this user ($k^* = k_2(n^*)$ or $k_1(n^*)$) from $\mathcal{U}_p$
    end if
  else
    Remove k^* from $\mathcal{U}_p$
  end if
end while

Phase 3: Power correction
if $\exists i \in \{1, \ldots, F\} / P_{RL_i} > P_m$
  Apply the power correction on $\mathcal{S}_k^{sole}, \forall k$ using (3.12), (3.13), and using P'_{m_i} instead of P_{m_i} in (3.15).
end if

each technique.

Table 3.1 – Approximate complexity of the different allocation techniques.

Technique	Complexity
OMA-HDAS	$O(KSR\log(S) + S(K+R) + f)$
OMA-HDAS-Realloc	$O\big(KS(R-F)\log(S) + S(K+R-F) + S(K+SF/K) + f\big)$
NOMA-HDAS	$O(KSR\log(S) + S(K+R) + S(K+SR) + f)$
NOMA-HDAS-Realloc	$O\big(KS(R-F)\log(S) + S(K+R-F) + S(K+SF/K) + S(K+SR) + f\big)$

It can be seen that OMA-HDAS and OMA-HDAS-Realloc present similar complexities, since the computational cost of the reallocation phase is compensated by an initial sorting over a smaller user-antenna set of subcarrier vectors. The same applies when comparing the complexity of NOMA-HDAS to that of NOMA-HDAS-Realloc since their NOMA pairing phases are essentially the same. However, when comparing NOMA to OMA algorithms, a noticeable complexity increase can be observed. This is driven by the dominant factor S^2R as opposed to the S^2F/K term in the reallocation phase. Since the cost of power correction is the same for all techniques, we compare the relative complexities

of the algorithms before power correction. For the configuration of Fig. 3.3, that is $K = 38$ users, $R = 4$ RRHs, $S = 64$ subcarriers, $F = 1$ constrained RRH and $R_{req} = 5$ Mbps per user, OMA-HDAS-Realloc is 17.7% less complex than OMA-HDAS, while NOMA-HDAS is 1.7% less complex than NOMA-HDAS-Realloc. Both NOMA techniques are about 46% more complex than OMA-HDAS.

3.7 Performance Evaluation

The performance of the proposed resource allocation schemes iss assessed for various conditions of antenna power limits, user rate requirements and number of users in the cell. The reported simulation results are averaged over 10 000 iterations of various user distributions, for each simulation setup. Results are compared against the OMA-DAS and NOMA-DAS scenarios of the previous chapter. The system is simulated using a hexagonal cell model with an outer radius r_d of 500 m. The network topology consists of four RRHs distributed as follows: one central antenna and three antennas uniformly positioned on a circle of radius $2r_d/3$ centered on the cell center. In all the simulated scenarios, the central antenna is considered to have no power limitation. In all the figures, except for Fig. 3.5, a single antenna (randomly chosen) is power limited ($F = 1$), whereas in Fig. 3.5, one, two and three power limited antennas are considered. Users are randomly deployed in the cell. The transmission medium is a frequency-selective Rayleigh fading channel with a root mean square delay spread τ of 500 ns. Large scale fading is composed of path-loss with a decay factor of 3.76, and lognormal shadowing with an 8 dB variance. The system bandwidth B is 10 MHz, and is divided into $S = 64$ subcarriers. The noise power spectral density N_0 is -173 dBm/Hz and the power threshold ρ is set to 0.01 W.

The power margin β is an important parameter in NOMA algorithms. As explained in section 3.5, it is essential to ensure the success-guaranteed nature of the algorithms since the remaining power P'_{m_i} is the actual one being used to solve (3.15). The cost of the power correction does not really depend on how distant the actual power of each antenna i is from P'_{m_i} (i.e. $|P_{RL_i} - P'_{m_i}|$), but it greatly depends on the ratio of P_{RL_i} before correction to the effective power limit of the antenna P'_{m_i}. Therefore, a power excess of 1 W when $P'_{m_i} = 5$ W, incurs a much more graceful degradation compared to the case when $P'_{m_i} = 0.01$ W. Also, the high or low amount of power margin with respect to P'_{m_i} left by the pairing steps is entirely linked to the randomness of the channel realizations. To counteract this, the power margin factor β is used to ensure $P'_{m_i} > (1 - \beta)P_{m_i}$. The larger the β, the greater the risk of having a significant power correction toll. Conversely, the lower the β, the smaller the number of accepted subcarriers for multiplexing, and the smaller the power reduction observed between OMA and NOMA. The optimal value of β comes then as a tradeoff in order to minimize the total system power. This optimal value depends on the system parameters, e.g. the targeted rate, the number of users, etc. Nevertheless, practical tests show that any value of β between 0.7 and 0.8 always guarantees a near-optimal tradeoff by leaving enough room to P_{m_i} for power correction without hindering the pairing process. For this reason, the value $\beta = 0.75$ is selected.

Fig. 3.3 presents the total transmit power in the cell as a function of the power limit P_m. At first, we note the important gap between orthogonal and non-orthogonal RA schemes in which the worst performing NOMA algorithm requires at least 40 W less power than any other OMA scheme at a power limit of 20 W, to provide all 38 users with

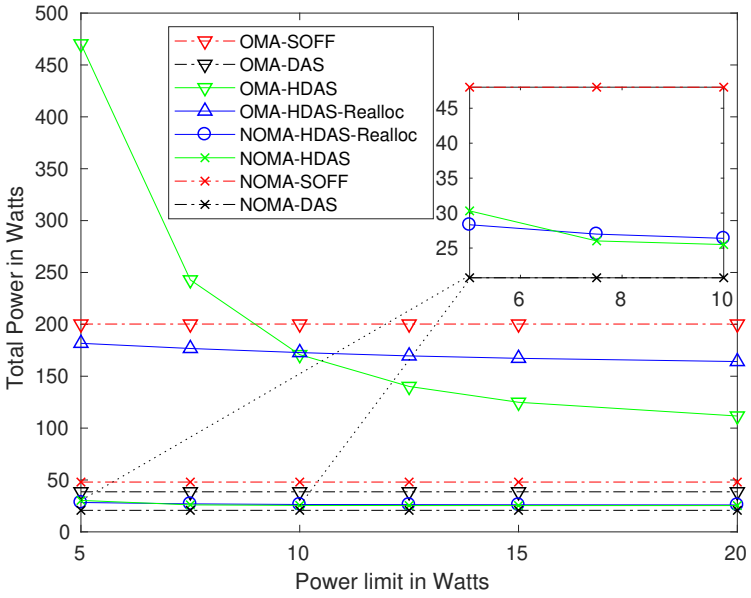


Figure 3.3 – Total power as a function of the antenna power limit for OMA and NOMA schemes, $R_{req} = 5$ Mbps, $K = 38$ users.

a requested rate of 5 Mbps. This amounts to a power decrease by more than a factor of two, which means that the complexity increase due to NOMA is largely overcome by the important power savings achieved over OMA. The performance of the algorithms under high power limit constraints gives an indication about the best performance that can be reached by each considered allocation technique. Therefore, in light of this remark and regarding orthogonal RA schemes, OMA-HDAS has a greater potential in limiting system power than OMA-HDAS-Realloc. However, OMA-HDAS only achieves this potential for relatively relaxed power constraints. Also, OMA-HDAS performance deteriorates drastically for more severe conditions: as the power limit decreases, OMA-HDAS leads to an increasingly more important power correction cost until it eventually gets worse than the trivial OMA-SOFF solution in which constrained antennas are simply shut off and the algorithm is run using the remaining antennas. On the other hand, OMA-HDAS-Realloc handles critical power conditions in a much more graceful way. Indeed, its total transmit power remains a reasonably better alternative than the trivial solution, while slightly increasing with the decreasing power limit. This is in accordance with the properties that were required from OMA-HDAS-Realloc in providing a solution that always outperforms the trivial solution.

As a conclusion, OMA-HDAS-Realloc performs better than OMA-HDAS by far for critical system conditions, whereas OMA-HDAS is better for the other extreme (i.e for loose system conditions of power limit, user rates and number of users). The same analysis can be drawn from the two competing NOMA algorithms as they suffer/benefit from the same advantages/drawbacks as shown in Figs. 3.3,3.4. The reason for this behavior being that each NOMA scheme is based on its orthogonal counterpart.

In Fig. 3.4, the performance of the OMA and NOMA schemes are presented as a function of the number of users. It can be observed that the behavior of the NOMA algorithms follows the lead of their OMA counterparts: starting from mild system conditions (i.e. for a relatively small number of users), NOMA-HDAS has barely an advantage over

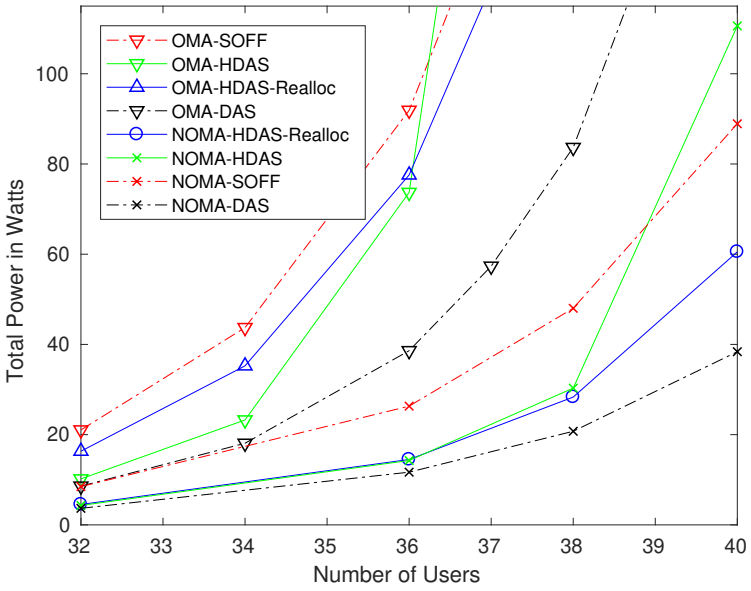


Figure 3.4 – Total power as a function of the number of users K for a requested rate of R_{req} = 5 Mbps with P_m = 5 W.

NOMA-HDAS-Realloc (14.3 W vs 14.5 W respectively for K = 32 users), till the point where the system conditions start to weigh too heavily on NOMA-HDAS, forcing important power corrections. The latter switches the balance in favor of NOMA-HDAS-Realloc which requires a transmit power of 60.5 W for a total of K = 40 users, incurring a 58% power increase with respect to NOMA-DAS against 188% inferred by NOMA-HDAS.

The percentage power increase of NOMA-HDAS-Realloc compared to NOMA-DAS increases with the number of users: 24% for K = 36 users, 37% for K = 38 users, and 58% for K = 40 users. This increase was expected since the total system power is increasing with the number of users while the imposed power-limit remains unchanged. Finally, we note the important reduction in the performance gaps when moving from OMA to NOMA, between HDAS-Realloc and HDAS algorithms, within the regions of mild system conditions. For example, for K = 32 users, a relative power difference of 60% is observed in the orthogonal context vs 5% of difference in the non-orthogonal one. This convergence of the algorithms in regions previously favorable to NOMA-HDAS promotes NOMA-HDAS-Realloc as a globally better candidate for resolving our RA problem.

Fig. 3.5 shows the evolution of the system power with the number of constrained antennas. As expected, the greater the number of constrained antennas, the more important the total power. Moreover, when comparing NOMA-HDAS and NOMA-HDAS-Realloc at 13 Mbps, we observe that the correction costs increase with the number of constrained antennas. At a rate of 13 Mbps, the difference between NOMA-HDAS and NOMA-HDAS-Realloc, is 0.27 dB (6.4%), 4.3 dB (169%) and 18.3 dB (6660%) for F = 1, 2 and 3 respectively. However, at lower values of the requested rate, the saved power of NOMA-HDAS with respect to NOMA-HDAS-Realloc is even larger for a larger number of constrained antennas (0.27 dB (6.4%), 0.77 dB (19.4%) and 2.5 dB (77.8%) for F = 1, 2 and 3 respectively at 12 Mbps). We conclude that the increase in the number of constrained antennas magnifies the differences between NOMA-HDAS-Realloc and NOMA-HDAS, both in critical and mild conditions.

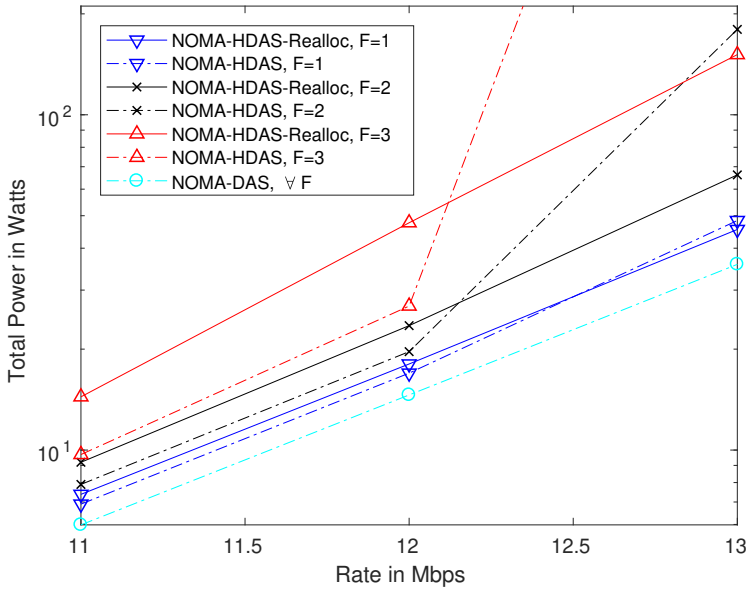


Figure 3.5 – Total power as a function of the target rate, for different values of the number of constrained antennas, with $K = 15$ users and $P_{m_i} = 15$ W.

3.8 Conclusion

In this chapter, we extended the proposed procedures developed in chapter 2 for downlink power minimization using mutual SIC NOMA to the context of HDAS while imposing power limitations on a subset of transmitting antennas. We first explored the characteristics of optimum PA in an orthogonal scenario, which enabled the design of RA schemes for both orthogonal and non-orthogonal contexts. The results suggest the use of different algorithms depending on system conditions: The NOMA-HDAS method is favored in the presence of low requested rates, high power limits and small numbers of served users and constrained antennas. On the other hand, the NOMA-HDAS-Realloc technique proves to be remarkably efficient in harsher system constraints, maintaining a significant advantage over the trivial solution of shutting down the constrained antennas. Thus, relying on a judicious antenna allocation in the first place is preferable over resorting systematically to the optimum power correction procedure.

The combination of NOMA with DAS gave birth to the mutual SIC concept leading to inter-user interference-free NOMA clusters. This complete interference cancellation proved its efficiency in the context of power minimization, as it was demonstrated in the last two chapters. In the second part of the thesis, starting from the next chapter, the potential of mutual SIC is explored for the dual problem, that is throughput maximization problems under power limit constraints.

The contributions of this chapter led to the publication of the following journal paper:

A. Kilzi, J. Farah, C. Abdel Nour and C. Douillard, “New Power Minimization Techniques in Hybrid Distributed Antenna Systems With Orthogonal and Non-Orthogonal Multiple Access,” in *IEEE Trans. Green Commun. Netw.*, vol. 3, no. 3, pp. 679-690, Sept. 2019.

Chapter 4

Enhancing the Spectral Efficiency of CoMP Systems using NOMA mutual SIC

The mutual SIC technique originated from the application of NOMA principles to the DAS setup, with multiplexed signals being sent by different RRHs. In this chapter, we seek to generalize the concept of mutual SIC in NOMA to cover the case of arbitrary numbers of multiplexed users. Meanwhile, we develop a new formalism for the mutual SIC procedure that can be directly applied to DAS and C-RAN as well as other network topologies (HetNets, small cells, etc.), provided that the signaling exchange enabling cooperation is available. The context of CoMP (cf. section 1.3) is therefore selected to conduct the study as it allows to encompass the cases of multi-cell and/or single cell scenarios while considering both joint transmission of signals by multiple TPs or single TP serving.

After providing an overview of previous works on NOMA in CoMP systems in the literature review of section 4.1, the system model is presented in section 4.2 where the system setup is described and the throughput maximization problem is clearly stated. Then in section 4.3, the fundamental conditions of PMC and rate conditions for a generalized mutual SIC are developed for JT-CoMP and DPS-CoMP. Afterwards, two case studies are conducted: in section 4.4, a two-user NOMA cluster is considered, and in section 4.5 three-user NOMA clusters are considered. The impact of mutual SIC on system throughput and fairness among users is presented in section 4.6, and the major conclusions of the chapter are drawn in section 4.7. The key contributions of this chapter can be summarized as follows:

- We propose to improve the cell-edge user rate and the overall system throughput by introducing JT not only for cell-edge but also for cell-centered users. In practice, JT is not necessarily applied to all users on all subbands, but may be restricted to users signals transmitted on subbands including at least one cell-edge user.

- We develop the conditions for allowing interference cancellation in NOMA for both DPS and JT scenarios, and show that, unlike previous CoMP techniques, SIC of the signals of inner users is possible at the level of the cell-edge user.

- We rigorously define the conditions allowing the feasibility of mutual SIC for any user and apply it to a three-user NOMA cluster.

- We show that JT is more favorable for enabling interference cancellation than DPS without being a necessary condition for achieving mutual SIC.

- We challenge the common practice of basing the user-antenna association on the Received Signal Strength (RSS) for achieving the highest system capacity, and we favor the associations allowing the much more profitable mutual SIC procedure.

4.1 Related Works

Several studies have proposed the combination of NOMA with CoMP techniques. In [107], the authors study a CoMP-NOMA system for downlink transmission and propose a suboptimal scheduling strategy that scales linearly with the number of users. In [108], the applicability of different NOMA-COMP scenarios is studied. The authors also argue that signals of users receiving CoMP transmissions must be decoded prior to those of non-CoMP users receiving single transmission. In [109], CoMP scenarios are studied in a HetNet system consisting of a macro BS, and multiple small BSs. The users requiring JT-CoMP transmission are first determined according to the RSS. Users with weak RSS – cell-edge users – are granted JT-CoMP transmission. The sub-optimal user-clustering for NOMA users developed in [110] is adopted, then a low complexity distributed power allocation for rate maximization is performed independently on every BS. In [30], a two-cell system made of one cell-edge user and two cell-center users (one in each cell) is studied. Alamouti code [111] is utilized with joint transmission to serve the cell-edge user with JT-CoMP in order to improve the performance of this user.

In all previous studies on NOMA-CoMP, only cell-edge users are considered as potential CoMP users. Also, cell-edge users are not considered able to decode and remove the signals of inner cell users. Finally, user-antenna association for non-CoMP users is based on the sole criterion of maximal RSS or channel gain. However, the concept of mutual SIC we introduced previously in chapter 2, and also applied in chapter 3, brings back into question the ideas of "strong" and "weak" users as they stand equally in front of interference cancellation. The configuration used in those chapters actually corresponds to an intra-site CoMP (using DPS), behaving as inter-site CoMP [112]. This new concept of mutual SIC relying on CoMP systems makes the combination of NOMA and CoMP much more interesting than their combination using the single SIC approach. Indeed, a complete interference cancellation (intra-cell and inter-cell) among users from the same NOMA cluster (whether they are cell-edge or cell-center users) becomes possible. Therefore, in this chapter we study the combination of NOMA and multicell-CoMP, establishing the conditions enabling a successful mutual SIC procedure at the level of all users, and assessing the performance by means of the system throughput metric.

4.2 System Model

We consider a two-cell downlink system where each cell has multiple RRHs deployed in a DAS setup such that the RRHs are connected to their BBUs through high capacity optical fibers (see Fig. 4.1). Similarly to previous setups, single-antenna RRHs are considered; hence, the terms "RRH" and "antenna" will be used interchangeably. Users transmit their CSI to RRHs, and the BBU collects all the CSI from RRHs and shares them with

other BBUs. Inter-BBU message exchange can be done through a direct X2 link between the BBUs of the two cells, in case of a fully meshed decentralized CoMP architecture, or both BBUs may be connected to a third party central unit in a star-like network, for a centralized CoMP architecture. In any case, whether the central unit coordinates the BBUs (which in turn control the RRHs), or whether BBUs exchange information in a decentralized manner, we assume that the information data relative to any user is made available at the level of both BBUs of each cell. Therefore, a Joint Processing (JP) CoMP scenario is considered with either a DPS scheme, where users are served by one antenna at a time, or a JT scheme in which a user may benefit from the transmission of the same signal over multiple antennas at the same time (cf. section 1.3.2). To focus on the cell-edge user scenario, we restrict the choice of serving antennas to the two located near the common frontier of the cells, one on each side. Let $\mathcal{K}$ be the set of users, with a maximum number of three users considered in the system, user 1 being a non-cell-edge user present in cell 1, user 2 a non-cell-edge user located in cell 2, and user 3 the cell-edge user. The serving RRH for user 1 in cell 1 is referred to as r_1 (or $r = 1$) and that of user 2 in cell 2 is referred to as r_2 (or $r = 2$). The system framework is presented in the schematic of Fig.4.1. Without loss of generality, three different geometric regions were defined, in which each user is randomly positioned.

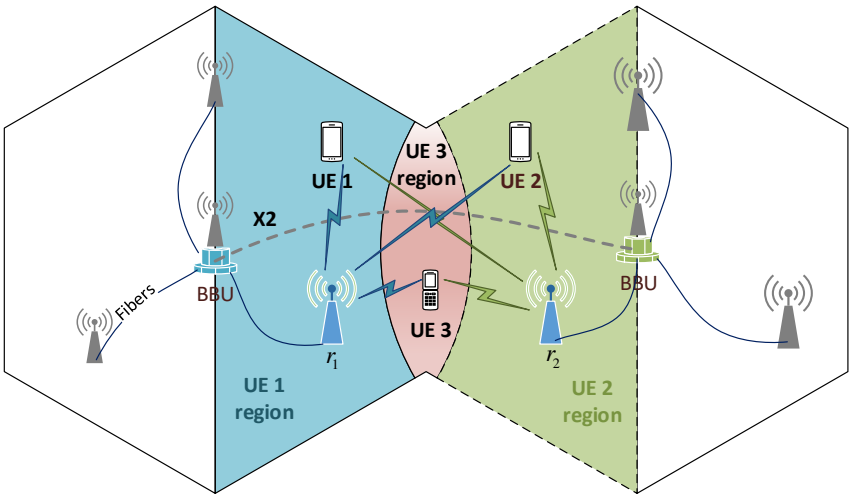


Figure 4.1 – Illustration of the two-cell DAS setup with the functional RRHs r_1 and r_2, and the three colored user regions (UE = user equipment).

The problem structure of this chapter is radically different from that of the two previous ones, since the purpose is to showcase the important advantages of combining mutual SIC with CoMP, rather than devising new resource allocation schemes. Indeed, in a practical implementation with a significant number of users in each cell, appropriate pairing or clustering methods must be incorporated in the resource allocation technique, so as to assign NOMA clusters of 2 or 3 users to subbands [110, 113–115]. This chapter is therefore focused on one of these particular clusters, with the main objective of the chapter being the study of the upper layer conditions enabling the combination of mutual SIC and CoMP (physical aspects of mutual SIC are out of the scope of the thesis). The resulting enhancements of the service quality of users in general, and cell-edge users in particular, are compared against classic NOMA scenarios [28, 104, 105], or previous CoMP scenarios [30]. To do so, the performance of different CoMP systems is analyzed from the system

capacity perspective. We aim to maximize the achievable total throughput of any given configuration, under the following sets of constraints:

- SIC constraints: the set of conditions that make the mutual SIC technique possible from the information theory perspective, i.e. the conditions on achievable rates at the respective users levels.
- PMCs: the set of conditions that make the mutual SIC technique feasible from a practical implementation perspective, i.e. the conditions on the received signals powers at the respective users levels. Let s_i be the signal of user i, $i \in \{1, 2, 3\}$. According to the SIC principle, if signals s_1, s_2 and s_3 are to be decoded in that order at the level of one of the users, then the signal power of s_1 at the level of that user must be greater than that of s_2 and s_3 combined, and the power level of s_2 must be greater than that of s_3. This guarantees SIC stability since every signal is ensured to be the dominant signal during its decoding [110, 116].
- Power limit constraints: the maximum total amount of transmit power available at the level of the RRHs.

As previously mentioned, the work in this chapter is conducted over a given NOMA cluster with known users and multiplexed subcarrier. Therefore, the subband index is dropped from the channel attenuation and power terms. Let $h_{k,r}$ be the channel attenuation experienced by a signal between antenna r and user k, and let $P_{k,r}$ be the power of signal s_k transmitted from antenna r to user k. This signal reaches k after experiencing the channel with attenuation factor $h_{k,r}$; the received signal power is therefore $P_{k,r}h_{k,r}$. In the case of JT, both antennas r_1 and r_2 are used for the transmission of the message to user k with transmit powers being respectively P_{k,r_1} and P_{k,r_2}. Hence, the received signal power is $P_{k,r_1}h_{k,r_1} + P_{k,r_2}h_{k,r_2}$. The system throughput is the sum of the rates achieved by all users in the system, its expression depending on whether DPS or JT is adopted and on the intra-cell and inter-cell interfering terms. When there is no interference (which is the case with a full mutual SIC between the three users), the rate expression for a user k is given by the Shannon capacity theorem:

$$R_k = \begin{cases} \dfrac{B}{S}\log_2\left(1 + \dfrac{P_{k,r}h_{k,r}}{\sigma^2}\right) \text{ for DPS,} & (4.1) \\ \dfrac{B}{S}\log_2\left(1 + \dfrac{P_{k,r_1}h_{k,r_1} + P_{k,r_2}h_{k,r_2}}{\sigma^2}\right) \text{for JT,} & (4.2) \end{cases}$$

σ^2 being the noise power over the subband bandwidth B/S normalized to 1. The problem formulation of sum-rate maximization over the transmit power variables $P_{k,r}$ takes the following generic form:

$$\max_{P_{k,r}} \sum_{i \in \mathcal{K}} R_i, \tag{4.3a}$$

such that:

$$\begin{cases} \text{Mutual SIC constraints are verified,} & (4.3b) \\ \text{PMCs are verified,} & (4.3c) \\ \text{Power limit constraints are verified.} & (4.3d) \end{cases}$$

In the following section, we derive the fundamental mutual SIC constraints for a general system of M users and two transmitting RRHs in a CoMP scenario. Then, attention is directed towards the application of the mutual SIC technique in a two-user and a three-user system in sections 4.4 and 4.5 respectively. The expressions of (4.3b), (4.3c) and (4.3d) will therefore be developed in each case.

4.3 Mutual SIC Conditions for CoMP Scenarios

In this section, we study the conditions, in terms of channel coefficients and transmit powers, that must be met to enable the mutual SIC procedure at the level of all users for any NOMA cluster size. To this end, a general framework for identifying the interfering user sets depending on the decoding order is introduced. The developed methodology is provided for JT transmission scenario which encompasses simpler DPS transmission schemes. In other words, the conditions concerning DPS-based mutual SIC schemes can be easily adapted from those shown in this section by canceling the transmitted power from one of the antennas to the user.

Let $\mathcal{M}$ be the NOMA cluster with dimension M, i.e. $\mathcal{M}$ is the set of users multiplexed over the same frequency resource. Given two users p and n randomly selected in that cluster, we seek to determine the conditions under which a successful mutual SIC can occur between the two users while in the presence of interfering signals from the remaining users in $\mathcal{M}$. In chapter 2, section 2.5.2.1, the mutual SIC conditions were developed for the special case of two users per cluster and a single-cell system. The rate conditions that must be verified to guarantee mutual SIC can be translated into SINR conditions: For n to successfully decode (and cancel) the signal of p denoted by s_p, the SINR of s_p at the level of n, denoted by $SINR_n^{s_p}$ must be greater than the SINR of s_p at the level of p itself ($SINR_p^{s_p}$). Therefore, the conditions of mutual SIC at the level of both users are:

$$\begin{cases} SINR_p^{s_n} > SINR_n^{s_n} & \text{SIC of } s_n \text{ at user } p, \quad (4.4) \\ SINR_n^{s_p} > SINR_p^{s_p} & \text{SIC of } s_p \text{ at user } n, \quad (4.5) \end{cases}$$

Determining the SINRs requires the knowledge of the interfering signals at the level of every user, at the time of decoding signals s_n and s_p. For example, if p managed to decode the signal of a third user m in the cluster while n did not, the SINR of p will not suffer from the interference caused by s_m, while decoding either s_n or s_p. The same cannot be said of user n in that case, which highlights the importance of the decoding order at every user. Indeed, the SINR terms vary according to this decoding order, which is instructed by the BBU to the RRH and then to the user via signaling. Therefore, mutual SIC conditions depend on each possible decoding order. Let $\mathcal{I}_p$ and $\mathcal{I}_n$ be the sets of interfering users on users p and n respectively. $\mathcal{I}_p$ and $\mathcal{I}_n$ can be each partitioned into two sets, a set of common interfering users between n and p named $\mathcal{C}_{pn}$, and a set of interfering users specific to n and p, $\mathcal{U}_n$ and $\mathcal{U}_p$ respectively. These sets have the following properties:

$$\begin{gathered} \mathcal{I}_p = \mathcal{C}_{pn} \cup \mathcal{U}_p, \quad \mathcal{C}_{pn} \cap \mathcal{U}_p = \emptyset, \\ \mathcal{I}_n = \mathcal{C}_{pn} \cup \mathcal{U}_n, \quad \mathcal{C}_{pn} \cap \mathcal{U}_n = \emptyset, \\ \mathcal{U}_p \cap \mathcal{U}_n = \emptyset. \end{gathered}$$

Depending on whether s_n or s_p is considered for decoding in (4.4) or (4.5) respectively, the interfering signals are interchanged between s_n and s_p. s_n is the useful signal in (4.4), and becomes the interfering signal in (4.5), whereas s_p is the interfering signal in (4.4) and the useful signal in (4.5). Therefore, the interfering user sets depend on the signal being decoded and their notation is defined accordingly, hence the terms $\mathcal{I}_p^{s_n}, \mathcal{I}_p^{s_p}, \mathcal{I}_n^{s_p}$, and $\mathcal{I}_n^{s_n}$. Also, since s_n is a common interfering signal to users p and n in (4.5), n belongs to $\mathcal{C}_{pn}$ when decoding s_p, thus the notation $\mathcal{C}_{pn}^{s_p}$ with $n \in \mathcal{C}_{pn}^{s_p}, n \notin \mathcal{U}_n, n \notin \mathcal{U}_p$. The same applies in (4.4) where s_p is a common interfering signal when decoding s_n, leading to the notation $\mathcal{C}_{pn}^{s_n}$ with $p \in \mathcal{C}_{pn}^{s_n}, p \notin \mathcal{U}_n, n \notin \mathcal{U}_p$. Then, it follows that $\mathcal{U}_n$ and $\mathcal{U}_p$ are not affected by the signal that is being decoded between s_n and s_p. This being said, the partition of the global interfering set for user n for ($\mathcal{I}_n^{s_p}$) is made relatively to the other user p whose signal is studied for decoding at the level of n. When the mutual SIC of n is studied with another user $p^{'}$, the global interfering set of n is changed but, more importantly, its partition is modified, thus affecting $\mathcal{U}_n$. To illustrate that with an example, if we consider the mutual SIC between n and p, where a third user m of the cluster has been previously decoded by p but not by n, then it would seem natural to state that m belongs to $\mathcal{U}_n$. However, when studying the SIC procedure between m and n, it is clear that m cannot belong to $\mathcal{U}_n$ since it is included in the common interfering sets of n and m. This means that the interfering set specific to user n $\mathcal{U}_n$ depends on the other user p considered for the application of mutual SIC; thus the needed notation $\mathcal{U}_{n(p)}$ (and $\mathcal{U}_{p(n)}$ for user p). To sum up, the user specific sets of n and p are independent of the signal being decoded (s_n and s_p), but they are at the same time defined according to the other user considered to have mutual SIC. The complete notations with the properties mentioned above are as follows:

for the decoding of s_n at p in (4.4)
$$\begin{cases} \mathcal{I}_p^{s_n} = \mathcal{C}_{pn}^{s_n} \cup \mathcal{U}_{p(n)}, \\ \mathcal{I}_n^{s_n} = \mathcal{C}_{pn}^{s_n} \cup \mathcal{U}_{n(p)}, \\ p \in \mathcal{C}_{pn}^{s_n}, \end{cases}$$

for the decoding of s_p at n in(4.5)
$$\begin{cases} \mathcal{I}_p^{s_p} = \mathcal{C}_{pn}^{s_p} \cup \mathcal{U}_{p(n)}, \\ \mathcal{I}_n^{s_p} = \mathcal{C}_{pn}^{s_p} \cup \mathcal{U}_{n(p)}, \\ n \in \mathcal{C}_{pn}^{s_p}, \end{cases}$$

By taking $r_1 = 1$ and $r_2 = 2$, the expression of $SINR_p^{s_n}$ can be written as:

$$\begin{aligned} SINR_p^{s_n} &= \frac{P_{n,1}h_{p,1} + P_{n,2}h_{p,2}}{\sum\limits_{i \in \mathcal{I}_p^{sn}} (P_{i,1}h_{p,1} + P_{i,2}h_{p,2}) + \sigma^2} \\ &= \frac{P_{n,1}h_{p,1} + P_{n,2}h_{p,2}}{\sum\limits_{i \in \mathcal{C}_{pn}^{sn}} (P_{i,1}h_{p,1} + P_{i,2}h_{p,2}) + \sum\limits_{i \in \mathcal{U}_{p(n)}} (P_{i,1}h_{p,1} + P_{i,2}h_{p,2}) + \sigma^2}. \end{aligned}$$

With these notations, the mutual SIC conditions that derive from (4.4) and (4.5) can now be developed by comparing $SINR_p^{s_n}$ with $SINR_n^{s_n}$ in (4.4), and $SINR_n^{s_p}$ with $SINR_p^{s_p}$ in (4.5). The SINR condition for the decoding of s_n at the level of p is: $SINR_p^{s_n} > SINR_n^{s_n}$. By subtracting $SINR_n^{s_n}$ from $SINR_p^{s_n}$ we get:

$$SINR_p^{s_n} - SINR_n^{s_n} = \frac{P_{n,1}h_{p,1} + P_{n,2}h_{p,2}}{\sum\limits_{i \in \mathcal{I}_p^{sn}} (P_{i,1}h_{p,1} + P_{i,2}h_{p,2}) + \sigma^2} - \frac{P_{n,1}h_{n,1} + P_{n,2}h_{n,2}}{\sum\limits_{i \in \mathcal{I}_n^{sn}} (P_{i,1}h_{n,1} + P_{i,2}h_{n,2}) + \sigma^2} > 0,$$

which leads to

$$\begin{aligned}A =&(P_{n,1}h_{p,1} + P_{n,2}h_{p,2})\Big[\sum_{i\in\mathcal{I}_n^{s_n}}(P_{i,1}h_{n,1} + P_{i,2}h_{n,2}) + \sigma^2\Big]\\ &-(P_{n,1}h_{n,1} + P_{n,2}h_{n,2})\Big[\sum_{i\in\mathcal{I}_p^{s_n}}(P_{i,1}h_{p,1} + P_{i,2}h_{p,2}) + \sigma^2\Big] > 0,\end{aligned}$$

where A is the numerator of $SINR_p^{s_n} - SINR_n^{s_n}$, whose expression can be further rearranged as:

$$\begin{aligned}A =&h_{n,1}h_{p,1}P_{n,1}\Big[\sum_{i\in\mathcal{I}_n^{s_n}}P_{i,1} - \sum_{i\in\mathcal{I}_p^{s_n}}P_{i,1}\Big]\\ &+h_{n,2}h_{p,2}P_{n,2}\Big[\sum_{i\in\mathcal{I}_n^{s_n}}P_{i,2} - \sum_{i\in\mathcal{I}_p^{s_n}}P_{i,2}\Big] + \sigma^2\Big[P_{n,1}(h_{p,1} - h_{n,1}) + P_{n,2}(h_{p,2} - h_{n,2})\Big]\\ &\underbrace{+h_{p,1}h_{n,2}\Big[P_{n,1}\sum_{i\in\mathcal{I}_n^{s_n}}P_{i,2} - P_{n,2}\sum_{i\in\mathcal{I}_p^{s_n}}P_{i,1}\Big] + h_{p,2}h_{n,1}\Big[P_{n,2}\sum_{i\in\mathcal{I}_n^{s_n}}P_{i,1} - P_{n,1}\sum_{i\in\mathcal{I}_p^{s_n}}P_{i,2}\Big]}_{B}.\end{aligned}$$

By detailing B, we get:

$$\begin{aligned}B =& h_{p,1}h_{n,2}\Big[P_{n,1}\Big(\sum_{i\in\mathcal{C}_{pn}^{s_n}}P_{i,2} + \sum_{i\in\mathcal{U}_{n(p)}}P_{i,2}\Big) - P_{n,2}\Big(\sum_{i\in\mathcal{C}_{pn}^{s_n}}P_{i,1} + \sum_{i\in\mathcal{U}_{p(n)}}P_{i,1}\Big)\Big]\\ &+ h_{p,2}h_{n,1}\Big[P_{n,2}\Big(\sum_{i\in\mathcal{C}_{pn}^{s_n}}P_{i,1} + \sum_{i\in\mathcal{U}_{n(p)}}P_{i,1}\Big) - P_{n,1}\Big(\sum_{i\in\mathcal{C}_{pn}^{s_n}}P_{i,2} + \sum_{i\in\mathcal{U}_{n(p)}}P_{i,2}\Big)\Big],\\ B =& (h_{p,1}h_{n,2} - h_{p,2}h_{n,1})\Big[P_{n,1}\sum_{i\in\mathcal{C}_{pn}^{s_n}}P_{i,2} - P_{n,2}\sum_{i\in\mathcal{C}_{pn}^{s_n}}P_{i,1}\Big] + h_{p,1}h_{n,2}\Big[P_{n,1}\sum_{i\in\mathcal{U}_{n(p)}}P_{i,2} - P_{n,2}\sum_{i\in\mathcal{U}_{p(n)}}P_{i,1}\Big]\\ &+ h_{p,2}h_{n,1}\Big[P_{n,2}\sum_{i\in\mathcal{U}_{n(p)}}P_{i,1} - P_{n,1}\sum_{i\in\mathcal{U}_{p(n)}}P_{i,2}\Big].\end{aligned}$$

In practical interference-limited systems, the background noise is negligible compared to the interfering signals [98, 99], i.e. $\sigma^2 << P_{k',r}h_{k,r}\ , \forall(k,k') \in \mathcal{K}^2, \forall r \in \{r_1, r_2\}$. Replacing B by its expression in A, we get the final expression of the SIC condition for the decoding of s_n at the level of user p:

$$\begin{aligned}&h_{n,1}h_{p,1}P_{n,1}\Big[\sum_{i\in\mathcal{U}_{n(p)}}P_{i,1} - \sum_{i\in\mathcal{U}_{p(n)}}P_{i,1}\Big] + h_{n,2}h_{p,2}P_{n,2}\Big[\sum_{i\in\mathcal{U}_{n(p)}}P_{i,2} - \sum_{i\in\mathcal{U}_{p(n)}}P_{i,2}\Big]\\ &+(h_{p,1}h_{n,2} - h_{p,2}h_{n,1})\Big[P_{n,1}\sum_{i\in\mathcal{C}_{pn}^{s_n}}P_{i,2} - P_{n,2}\sum_{i\in\mathcal{C}_{pn}^{s_n}}P_{i,1}\Big]\\ &+h_{p,1}h_{n,2}\Big[P_{n,1}\sum_{i\in\mathcal{U}_{n(p)}}P_{i,2} - P_{n,2}\sum_{i\in\mathcal{U}_{p(n)}}P_{i,1}\Big] + h_{p,2}h_{n,1}\Big[P_{n,2}\sum_{i\in\mathcal{U}_{n(p)}}P_{i,1} - P_{n,1}\sum_{i\in\mathcal{U}_{p(n)}}P_{i,2}\Big] > 0.\end{aligned} \tag{4.6}$$

To determine the condition for the decoding of s_p at the level of user n, n and p are simply swapped in (4.6).

Having defined the general conditions of a mutual SIC between two random users of a NOMA cluster of size M, we consider the special cases $M = 2$ and $M = 3$ in the following sections. We explore the specific properties of every case allowing different mutual SIC scenarios, we establish the corresponding set of PMCs and we discuss their significance and implications, before describing the followed methodology to efficiently assess the performance of each scenario.

4.4 Mutual SIC in a Two-User System

To determine the mutual SIC conditions in a two-user system, also referred to as Dual Mutual SIC (DMSIC), we first have to identify the interfering user sets for each user. Without loss of generality, we consider in this section that only users 1 and 2 from Fig. 4.1 are present in the system. However, the same reasoning can be developed for any couple of users, whether it includes a cell-center user and a cell-edge user or two cell-center users, leading to the same conditions with different indexes. Since users 1 and 2 constitute the whole NOMA cluster, the interfering sets specific to each user, $\mathcal{U}_{1(2)}$ and $\mathcal{U}_{2(1)}$, are empty and the interfering sets $\mathcal{I}_1$ and $\mathcal{I}_2$ are identical. Thus, by letting $p = 1$ and $n = 2$, we get $C_{12}^{s_2} = \{1\}$, and condition (4.6) under which user 1 is capable of decoding the signal s_2 of user 2 becomes:

$$(h_{1,1}h_{2,2} - h_{1,2}h_{2,1})[P_{2,1}P_{1,2} - P_{2,2}P_{1,1}] > 0. \tag{4.7}$$

Also, by letting $p = 2$ and $n = 1$, we get $C_{21}^{s_1} = \{2\}$, and the condition under which user 2 is capable of decoding the signal s_1 of user 1 is:

$$(h_{2,1}h_{1,2} - h_{1,1}h_{2,2})[P_{1,1}P_{2,2} - P_{1,2}P_{2,1}] > 0.$$

These two SIC conditions are equivalent and form a unique constraint. Therefore, if one user satisfies the constraint of interference cancellation, the other one does as well, and if one cannot perform SIC, the other user cannot either. This result is radically different from that of classic SIC in CAS [28, 104, 105], or a DAS with the paired signals powered by a common RRH (see chapter 2, section 2.5.1), where only one user out of the two performs interference cancellation.
Next, we investigate DMSIC in DPS and JT scenarios. We highlight the PMCs that differentiate each case as well as the corresponding formulation of the power limit constraints, before defining the new user-RRH association and power allocation strategy in each case.

4.4.1 Two-User System with Dynamic Point Selection

4.4.1.1 DPS-DMSIC

The use of multiple antennas to power the signals of multiplexed users is what rendered feasible the mutual SIC procedure that we introduced in chapter 2. The only transmission scenario considered in chapter 2 is in fact an intra-site CoMP with dynamic point selection only. As stated earlier, the calculation developed here considers the general case of JT-served users. To obtain the underlying DPS constraints, the signal must be transmitted from one antenna only. This translates into canceling out either $P_{m,1}$ or $P_{m,2}$ for any user

m (m = 1 or 2). By setting $P_{2,1}$ and $P_{1,2}$ to 0, the context of condition (2.32) (chapter 2, section 2.5.2.1) is met where user 1 is assigned to r_1 and user 2 to r_2, and the DMSIC condition (4.7) becomes solely dependent on the channel coefficients of the users:

$$P_{2,2}P_{1,1}(h_{1,2}h_{2,1} - h_{1,1}h_{2,2}) > 0 \Rightarrow h_{1,2}h_{2,1} > h_{1,1}h_{2,2}. \tag{4.8}$$

In other terms, the ability of the system to perform DMSIC when user 1 is powered by r_1 and user 2 by r_2 is uniquely determined by the channel characteristics of the system, since the power factors are necessarily positive. However, if $h_{1,1}h_{2,2} > h_{1,2}h_{2,1}$, DMSIC can still be achieved in the system by switching the serving antennas of the users. Indeed, if $P_{1,1} = P_{2,2} = 0$ in (4.7), then user 1 is served by r_2 and user 2 by r_1, satisfying the new corresponding mutual SIC constraints as follows:

$$P_{2,1}P_{1,2}(h_{1,1}h_{2,2} - h_{1,2}h_{2,1}) > 0 \Rightarrow h_{1,1}h_{2,2} > h_{1,2}h_{2,1}. \tag{4.9}$$

As a conclusion, in a two-user system using DPS, the channel characteristics are the only factors that determine the antenna association of each user: if $h_{1,2}h_{2,1} > h_{1,1}h_{2,2}$, user 1 is served by r_1 and user 2 by r_2; if not, the antenna association is simply reversed. Note that in either case, the users are not necessarily assigned their best antenna, from the channel gain perspective. For example, considering the case where $h_{1,2}h_{2,1} > h_{1,1}h_{2,2}$, it is impossible to have $h_{1,1} > h_{1,2}$ and $h_{2,2} > h_{2,1}$ at the same time, meaning that at least one user will not be served by its best antenna. Either only one user is assigned to its most preferable antenna, or neither user is served by its best antenna. Therefore, the DMSIC procedure goes against the usual practice of associating the user to its closest/best antenna. While this might seem counter-intuitive at first, it should be understood that the rate gain provided by interference cancellation greatly overcomes the channel gain "deficit", as it will be shown in the performance assessment section (section 4.6).

Moving on to the PMCs, the PA must ensure that the power level of the signal to be decoded (at the level of a given user) is higher than the combined power levels of the remaining signals that have not been decoded yet. Table 4.1 presents the PMCs and power limit constraints for every user according to the channel characteristics. P_{L_1} and P_{L_2} are the transmit power limits of RRHs r_1 and r_2, respectively.

Table 4.1 – PMCs and power limit constraints for two-user DPS clusters

	Channel gain conditions	
	$h_{1,1}h_{2,2} < h_{1,2}h_{2,1}$	$h_{1,1}h_{2,2} > h_{1,2}h_{2,1}$
User 1 PMC	$P_{2,2}h_{1,2} > P_{1,1}h_{1,1}$	$P_{2,1}h_{1,1} > P_{1,2}h_{1,2}$
User 2 PMC	$P_{1,1}h_{2,1} > P_{2,2}h_{2,2}$	$P_{1,2}h_{2,2} > P_{2,1}h_{2,1}$
Power limit	$P_{1,1} \leq P_{L_1}$ $P_{2,2} \leq P_{L_2}$	$P_{2,1} \leq P_{L_1}$ $P_{1,2} \leq P_{L_2}$

The two PMCs of the first case in Table 4.1 can be summed up in the same form as in (2.31):

$$\frac{h_{2,2}}{h_{2,1}} < \frac{P_{1,1}}{P_{2,2}} < \frac{h_{1,2}}{h_{1,1}}. \tag{4.10}$$

Note that a necessary and sufficient condition for the existence of a PA scheme that satisfies (4.10) is to have (4.8). Indeed if (4.8) is true, the right member of (4.10) is greater than the left one. The same holds for the second case in Table 4.1, when (4.9) is true.

The objective function of the optimization problem (4.3a) presented in section 4.2 is the sum of the user rates as expressed in (4.1). The DMSIC constraints determine the user-antenna association, and this affects the expressions of the PMCs and power limits as shown in Table 4.1. The corresponding strategy is referred to as DPS-DMSIC.

4.4.1.2 DPS-NoSIC

To assess the efficiency of DPS-DMSIC, we also consider a benchmark scenario, namely DPS-NoSIC, in which the mutual SIC procedure is excluded at both user sides. Thus, the imposed PMCs for DPS-DMSIC are dropped. In DPS-NoSIC, users may be served by the same antenna as there is no more obligation to satisfy the mutual SIC conditions. Then, for any given channel realization, two additional user-antenna associations are identified when both users are served by the same antenna r_1 or r_2, which raises to four the number of possible user-antenna associations. The expressions of the users rates now include the interfering term from every other user:

$$R_1 = \log_2\left(1 + \frac{P_{s_1,r(s_1)}h_{1,r(s_1)}}{P_{s_2,r(s_2)}h_{1,r(s_2)} + \sigma^2}\right),$$
$$R_2 = \log_2\left(1 + \frac{P_{s_2,r(s_2)}h_{2,r(s_2)}}{P_{s_1,r(s_1)}h_{2,r(s_1)} + \sigma^2}\right),$$

where $r(s_k)$ denotes the antenna powering the signal of user k. For every channel realization, the problems corresponding to the four user-antenna associations are solved, and the scheme yielding the highest throughput is retained.

4.4.2 Two-User System with Joint Transmission

4.4.2.1 JT-DMSIC

Users subject to JT receive their information signals from multiple RRHs which can be affiliated to different cells. In that regard, a user is not associated to a specific cell, and the idea of switching the user-antenna association as in the DPS case becomes irrelevant. The validity of the DMSIC constraint is a function of the channel and power variables, contrary to DPS. The BBUs must therefore adapt the PA in order to ensure the following condition:

$$(h_{1,1}h_{2,2} - h_{1,2}h_{2,1})[P_{2,1}P_{1,2} - P_{2,2}P_{1,1}] > 0. \tag{4.7}$$

By inspecting (4.7), we can see that if $h_{1,1}h_{2,2} > h_{1,2}h_{2,1}$, the PA must ensure that $P_{2,1}P_{1,2} > P_{2,2}P_{1,1}$; otherwise, the power condition must be reversed.

Regarding the PMCs, the power level of s_1 at the level of user 1 is the sum of the signal powers from r_1 and r_2 and it amounts to $P_{1,1}h_{1,1} + P_{1,2}h_{1,2}$. Therefore, the PMC for the decoding of s_2 at the level of user 1 is given by :

$$P_{2,1}h_{1,1} + P_{2,2}h_{1,2} > P_{1,1}h_{1,1} + P_{1,2}h_{1,2}. \tag{4.11}$$

Similarly, the PMC for the decoding of s_1 at the level of user 2 is:

$$P_{1,1}h_{2,1} + P_{1,2}h_{2,2} > P_{2,1}h_{2,1} + P_{2,2}h_{2,2}. \tag{4.12}$$

Proposition 4.1. *If the PMCs of a two-user system are valid, the DMSIC condition is necessarily valid as well.*

Proof. Let us rewrite the PMCs (4.11) and (4.12) in the following form:

$$(P_{2,2} - P_{1,2})h_{1,2} > (P_{1,1} - P_{2,1})h_{1,1},$$
$$(P_{1,1} - P_{2,1})h_{2,1} > (P_{2,2} - P_{1,2})h_{2,2}.$$

Then, the terms $P_{1,1} - P_{2,1}$ and $P_{2,2} - P_{1,2}$ have the same sign. If they are both positive, we get the following inequality:

$$\frac{h_{1,1}}{h_{1,2}} < \frac{P_{2,2} - P_{1,2}}{P_{1,1} - P_{2,1}} < \frac{h_{2,1}}{h_{2,2}},$$

which leads to $h_{2,2}h_{1,1} < h_{1,2}h_{2,1}$ (actually, the channel constraint imposes the positive sign of $P_{1,1} - P_{2,1}$ and $P_{2,2} - P_{1,2}$, not the other way around). However, since $P_{1,1} - P_{2,1}$ and $P_{2,2} - P_{1,2}$ are assumed positive, $P_{2,2}P_{1,1} > P_{2,1}P_{1,2}$. The DMSIC condition (4.7) is thus verified, since power term $(P_{2,1}P_{1,2} - P_{2,2}P_{1,1})$ and channel term $(h_{1,1}h_{2,2} - h_{1,2}h_{2,1})$ have the same sign.
Similarly, assuming the negativity of $P_{1,1} - P_{2,1}$ and $P_{2,2} - P_{1,2}$ implies opposite channel conditions $(h_{1,1}h_{2,2} > h_{1,2}h_{2,1})$ and transmit power relations ($P_{2,1} > P_{1,1}$ and $P_{1,2} > P_{2,2} \implies P_{1,2}P_{2,1} > P_{1,1}P_{2,2}$), which makes (4.7) a product of two positive terms. This concludes our proof. □

Therefore, both PMCs at the level of users 1 and 2 encompass their common DMSIC condition, hence the number of constraints in the PA problem of sum-throughput maximization through DMSIC is reduced. The last two constraints account for the transmit power limits of each RRH given by:

$$P_{1,1} + P_{2,1} \leq P_{L_1}, \tag{4.13}$$
$$P_{1,2} + P_{2,2} \leq P_{L_2}. \tag{4.14}$$

On a side note, even though user-antenna association is irrelevant to the JT context, the power allocation is similar to the user-antenna selection in DPS: when $h_{1,2}h_{2,1} > h_{1,1}h_{2,2}$, the dominant signal transmitted by r_1 is s_1 (since $P_{1,1} > P_{2,1}$) and the dominant signal transmitted by r_2 is s_2 (since $P_{2,2} > P_{1,2}$), taking us back to user-antenna association in DPS when $h_{1,2}h_{2,1} > h_{1,1}h_{2,2}$. The same analysis applies when $h_{1,2}h_{2,1} < h_{1,1}h_{2,2}$: s_2 is dominant at the level of r_1 and s_1 is dominant at the level of r_2. This showcases how DPS is a special case of JT and implies that JT is naturally richer in potential and properties. For this reason, in section 4.5, we consider only JT scenarios for a three-user system, as it inherently encompasses all the DPS cases and many others.

At last, the problem formulation for the JT case can be summed up as: maximize sum rate $R_1 + R_2$ expressed using (4.2), under power limit constraints (4.13) and (4.14) and PMCs (4.11) and (4.12).

4.4.2.2 JT-NoSIC

JT-NoSIC is introduced to assess the efficiency of the DMSIC procedure when applied to JT-users. It serves as a benchmark for the performance of JT-DMSIC. The problem structure in JT-NoSIC remains globally unchanged except that the PMCs are dropped, and the rate expressions of R_1 and R_2 are given by:

$$R_1 = \log_2\left(1 + \frac{P_{1,1}h_{1,1} + P_{1,2}h_{1,2}}{P_{2,1}h_{1,1} + P_{2,2}h_{1,2} + \sigma^2}\right),$$
$$R_2 = \log_2\left(1 + \frac{P_{2,1}h_{2,1} + P_{2,2}h_{2,2}}{P_{1,1}h_{2,1} + P_{1,2}h_{2,2} + \sigma^2}\right).$$

4.5 Mutual SIC in a Three-User System

In this section, mutual SIC is studied for a three-user NOMA cluster. The conventional technique for serving users in CoMP is presented first, then a new scheme based on a full JT system is introduced. At last, a middle-ground strategy combining the proposed and conventional serving methods is proposed to enable a fair comparison between the methods.

4.5.1 The Conventional Approach (CellEdgeJT-CellCenterSIC)

The conventional way of employing JT was first thought of as a way to improve the signal quality of weak cell-edge users that suffer the most from inter-cell interference. Inner-cell users are generally considered to be more interference-immune given their proximity to the serving antenna, and their relative distance from the interfering ones. In that sense, the study in [30] sought to improve the system spectral efficiency by serving the cell-edge user (user 3 in Fig. 4.1) by both RRHs r_1 and r_2, while user 1 and 2 are served uniquely by their closest antennas, r_1 and r_2 respectively. In that setup, the cell-edge user suffers from the interference of both user 1 and user 2; however, it is the only user taking advantage of cell coordination in JT. Users 1 and 2 are able to successfully decode the signal of user 3 but cannot remove each other's signals. From a classic single-antenna single-SIC point of view, the cell-edge user is the weak user both in cell 1 with user 1, and in cell 2 with user 2. We refer to this method as CellEdgeJT-CellCenterSIC.
Let us determine the SIC conditions at the level of user 1 and user 2 respectively to remove the signal of user 3 (these conditions were not considered in [30]). Since both users $p = 1$ (resp. $p = 2$) and $n = 3$ suffer from the interference of user $m = 2$ (resp. $m = 1$), we have $\mathcal{C}_{pn}^{s_n} = \{m, p\} = \{2, 1\}$ (resp. $\{1, 2\}$), $\mathcal{U}_{p(n)} = \mathcal{U}_{n(p)} = \emptyset$. After replacing each variable by its value in (4.6), and keeping in mind that $P_{1,2} = P_{2,1} = 0$, the SIC conditions for the decoding of s_3 at the level of users 1 and 2 are respectively:

$$(h_{1,1}h_{3,2} - h_{1,2}h_{3,1})[P_{3,1}P_{2,2} - P_{3,2}P_{1,1}] > 0, \tag{4.15}$$
$$(h_{2,1}h_{3,2} - h_{2,2}h_{3,1})[P_{3,1}P_{2,2} - P_{3,2}P_{1,1}] > 0. \tag{4.16}$$

These conditions imply that the common power factor and the two channel factors must have the same sign:

$$\begin{aligned}\text{sign}\ (h_{1,1}h_{3,2} - h_{1,2}h_{3,1}) &= \text{sign}\ (h_{2,1}h_{3,2} - h_{2,2}h_{3,1}) \\ &= \text{sign}\ (P_{3,1}P_{2,2} - P_{3,2}P_{1,1}).\end{aligned} \tag{4.17}$$

The validity of this SIC procedure is mainly based on the channel properties: if both channel factors are not of the same sign, SIC is not applicable.
The problem formulation of CellEdgeJT-CellCenterSIC resides in the maximization of the sum rate of R_1, R_2 and R_3 given by:

$$R_1 = \log_2\left(1 + \frac{P_{1,1}h_{1,1}}{P_{2,2}h_{1,2} + \sigma^2}\right),\ R_2 = \log_2\left(1 + \frac{P_{2,2}h_{2,2}}{P_{1,1}h_{2,1} + \sigma^2}\right),$$
$$R_3 = \log_2\left(1 + \frac{P_{3,1}h_{3,1} + P_{3,2}h_{3,2}}{P_{1,1}h_{3,1} + P_{2,2}h_{3,2} + \sigma^2}\right),$$

having (4.17) as SIC constraints, and the following PMC and power limit constraints :

$$P_{3,1}h_{1,1} + P_{3,2}h_{1,2} > P_{1,1}h_{1,1} + P_{2,2}h_{1,2},$$
$$P_{3,1}h_{2,1} + P_{3,2}h_{2,2} > P_{2,2}h_{2,2} + P_{1,1}h_{2,1},$$
$$P_{1,1} + P_{3,1} \leq P_{L_1},$$
$$P_{2,2} + P_{3,2} \leq P_{L_2}.$$

4.5.2 Triple Mutual SIC in a Joint Transmission System (FullJT-TMSIC)

In this subsection, we propose the use of JT for the whole NOMA cluster. This is driven by three main reasons:

1. The densification of the network topology implies smaller distances between users and antennas, but also between RRHs of different cells. This proximity of RRHs brings back into question the ICI-immune character of cell-center users, hence the potential use of JT for these users.

2. Inspired by the results of section 4.4.1, the ideas of weak and strong users no longer hold in the presence of a mutual SIC procedure. Therefore, exploring the mutual SIC capabilities of the system for all three users and not just the cell-edge user is an idea worth investigating.

3. The use of JT maximizes the chances of successful Triple Mutual SIC (TMSIC), since all possible DPS combinations are only special cases of joint transmission as pointed out in section 4.4.2.

We propose a new method to enable a complete mutual SIC procedure at the level of every user, through the use of JT. This means that every user must be able to decode and subtract the signals of both other users. The mutual SIC conditions, in this case, strongly depend on the decoding order undergone at the level of each user, as previously discussed in section 4.3. This decoding order is related to the PMCs: user p cannot decode the signal of user n unless the power level of s_n is dominant at p. At the level of every user, two decoding orders are possible, raising to eight the total number of decoding orders combinations in the system, as shown in Table 4.2. The second row labels each combination by an identifying number. The cells of the table indicate, for any user (row), and any selected combination (column), the decoding order followed at the level of the user. For example, in the first combination, user 1 starts by decoding the signal of user 2 before proceeding to that of user 3.

Table 4.2 – The eight potential decoding orders of TMSIC

	Decoding order ID							
	1	2	3	4	5	6	7	8
u_1	u_2-u_3	u_2-u_3	u_2-u_3	u_2-u_3	u_3-u_2	u_3-u_2	u_3-u_2	u_3-u_2
u_2	u_1-u_3	u_1-u_3	u_3-u_1	u_3-u_1	u_1-u_3	u_1-u_3	u_3-u_1	u_3-u_1
u_3	u_1-u_2	u_2-u_1	u_1-u_2	u_2-u_1	u_1-u_2	u_2-u_1	u_1-u_2	u_2-u_1

Let m, n, and p be the three users of the system. For any selected pair of users (p, n), and for a given decoding order, their mutual SIC constraints fall in one of the three following categories of mutual SIC:

1. Users p and n did not manage to decode the signal of user m prior to decoding their respective signals. The users-decoding ID triplets (p,n,ID) that fall into this category are: (u_1,u_2,1), (u_1,u_2,2), (u_2,u_3,4), (u_1,u_3,5), (u_1,u_3,7), and (u_2,u_3,8).

2. User p managed to decode the signal of user m prior to decoding the signal of user n, while n did not manage to decode s_m before proceeding with s_p. The corresponding ordered triplets (p,n,ID) are: (u_1,u_3,1), (u_2,u_3,2), (u_1,u_3,3), (u_3,u_2,3), (u_2,u_1,3) (u_2,u_1,4), (u_1,u_2,5), (u_2,u_3,6), (u_1,u_2,6), (u_3,u_1,6), (u_3,u_2,7), and (u_3,u_1,8).

3. Both users p and n successfully decoded the signal of user m prior to decoding each others signals. The corresponding triplets are: (u_2,u_3,1), (u_1,u_3,2), (u_1,u_3,4), (u_2,u_3,5), (u_1,u_2,7), and (u_1,u_2,8).

For every scenario, we start by identifying the interference sets $\mathcal{C}_{pn}^{s_n}$, $\mathcal{C}_{pn}^{s_p}$, $\mathcal{U}_{p(n)}$ and $\mathcal{U}_{n(p)}$, and then derive the mutual SIC conditions between n and p. From section 4.3, we recall that $p \in \mathcal{C}_{pn}^{s_n}$ and $n \in \mathcal{C}_{pn}^{s_p}$.

Scenario 1

Users p and n did not decode s_m before canceling each other's interference. In this case, m is a common interfering signal to p and n. Therefore $\mathcal{C}_{pn}^{s_n} = \{m, p\}$, $\mathcal{C}_{pn}^{s_p} = \{m, n\}$, and $\mathcal{U}_{n(p)} = \mathcal{U}_{p(n)} = \emptyset$. Using (4.6), we get the following condition at the level of user p:

$$(h_{p,1}h_{n,2} - h_{p,2}h_{n,1})[P_{n,1}(P_{p,2} + P_{m,2}) - P_{n,2}(P_{p,1} + P_{m,1})] > 0.$$

The SIC condition at the level of user n is simply obtained by interchanging p and n in the previous expression:

$$(h_{n,1}h_{p,2} - h_{n,2}h_{p,1})[P_{p,1}(P_{n,2} + P_{m,2}) - P_{p,2}(P_{n,1} + P_{m,1})] > 0.$$

By letting $H_{pn} = h_{p,1}h_{n,2} - h_{p,2}h_{n,1}$, the mutual SIC conditions can be written in the following form:

$$\begin{cases} H_{pn}[P_{n,1}(P_{p,2} + P_{m,2}) - P_{n,2}(P_{p,1} + P_{m,1})] > 0, & (4.18) \\ H_{pn}[P_{p,2}(P_{n,1} + P_{m,1}) - P_{p,1}(P_{n,2} + P_{m,2})] > 0. & (4.19) \end{cases}$$

Note that, contrary to the two-user JT system, the SIC condition to remove s_p at the level of n is no longer the same as the SIC condition to cancel s_n at the level of p. This means

that it may happen that only one of the users succeeds in decoding the signal of the other one. The PMCs for the removal of s_n then s_m at the level of user p are respectively:

$$P_{n,1}h_{p,1} + P_{n,2}h_{p,2} > (P_{p,1} + P_{m,1})h_{p,1} + (P_{p,2} + P_{m,2})h_{p,2},$$
$$P_{m,1}h_{p,1} + P_{m,2}h_{p,2} > P_{p,1}h_{p,1} + P_{p,2}h_{p,2}.$$

The PMCs for the removal of s_p then s_m at the level of user n are respectively:

$$P_{p,1}h_{n,1} + P_{p,2}h_{n,2} > (P_{n,1} + P_{m,1})h_{n,1} + (P_{n,2} + P_{m,2})h_{n,2},$$
$$P_{m,1}h_{n,1} + P_{m,2}h_{n,2} > P_{n,1}h_{n,1} + P_{n,2}h_{n,2}.$$

Scenario 2

User p decoded s_m and user n did not decode s_m before canceling their respective signals. In this scenario, m only affects the interfering set of user n, therefore we have $\mathcal{U}_{n(p)} = \{m\}$, $\mathcal{U}_{p(n)} = \emptyset$, $C_{pn}^{s_n} = \{p\}$, $C_{pn}^{s_p} = \{n\}$. Let A be the expression of the SIC condition at the level of user p. Using (4.6), we have:

$$A = h_{n,1}h_{p,1}P_{n,1}P_{m,1} + h_{n,2}h_{p,2}P_{n,2}P_{m,2} + h_{p,1}h_{n,2}P_{n,1}P_{m,2} + h_{p,2}h_{n,1}P_{n,2}P_{m,1}$$
$$+(h_{p,1}h_{n,2} - h_{p,2}h_{n,1})(P_{n,1}P_{p,2} - P_{n,2}P_{p,1}) > 0.$$

By adding and subtracting $h_{p,2}h_{n,1}P_{n,1}P_{m,2}$ and $h_{p,1}h_{n,2}P_{n,2}P_{m,1}$ to A, we get:

$$\begin{aligned} A = {} & h_{n,1}h_{p,1}P_{n,1}P_{m,1} + h_{n,2}h_{p,2}P_{n,2}P_{m,2} \\ & + h_{p,2}h_{n,1}P_{n,1}P_{m,2} + h_{p,1}h_{n,2}P_{n,2}P_{m,1} \\ & - h_{p,2}h_{n,1}P_{n,1}P_{m,2} - h_{p,1}h_{n,2}P_{n,2}P_{m,1} \\ & + h_{p,1}h_{n,2}P_{n,1}P_{m,2} + h_{p,2}h_{n,1}P_{n,2}P_{m,1} \\ & + (h_{p,1}h_{n,2} - h_{p,2}h_{n,1})(P_{n,1}P_{p,2} - P_{n,2}P_{p,1}). \end{aligned}$$

Grouping the terms in the first two rows and factoring them yields: $(P_{n,1}h_{n,1} + P_{n,2}h_{n,2}) \times (P_{m,1}h_{p,1} + P_{m,2}h_{p,2})$. Grouping the third and forth rows together and taking out common factors yields: $H_{pn}(P_{n,1}P_{m,2} - P_{n,2}P_{m,1})$. Therefore A, becomes:

$$A = [P_{n,1}(P_{p,2} + P_{m,2}) - P_{n,2}(P_{p,1} + P_{m,1})]H_{pn} + (P_{n,1}h_{n,1} + P_{n,2}h_{n,2})[P_{m,1}h_{p,1} + P_{m,2}h_{p,2}] > 0.$$

There is an additional positive term compared to (4.18). This means that the condition that must be satisfied to ensure SIC of s_n at the level of p is less stringent when p has previously removed the message of user m. This result is shown here through calculation, but it is also intuitive, since removing the interference term of user m enhances $SINR_p^{s_n}$ compared to $SINR_n^{s_n}$ in (4.4). On the other hand, this dissymmetry of the interfering user sets degrades the chances of n to perform SIC of s_p when compared to (4.19), as its $SINR_n^{s_p}$ suffers from an interference that is not present in $SINR_p^{s_p}$ in (4.5). This can be verified by deriving the SIC condition at the level of n. To obtain the SIC conditions at the level of user n, n and p must be interchanged in the initial SIC condition in (4.6) before making any replacement in the interfering sets leading to the current expression of

A. By letting B be the expression of the SIC condition we get:

$$B = h_{p,1}h_{n,1}P_{p,1}\Bigg[\sum_{i\in\mathcal{U}_{p(n)}} P_{i,1} - \sum_{i\in\mathcal{U}_{n(p)}} P_{i,1}\Bigg] + h_{p,2}h_{n,2}P_{p,2}\Bigg[\sum_{i\in\mathcal{U}_{p(n)}} P_{i,2} - \sum_{i\in\mathcal{U}_{n(p)}} P_{i,2}\Bigg]$$
$$+(h_{n,1}h_{p,2} - h_{n,2}h_{p,1})\Bigg[P_{p,1}\sum_{i\in C_{pn}^{sp}} P_{i,2} - P_{p,2}\sum_{i\in C_{pn}^{sp}} P_{i,1}\Bigg]$$
$$+h_{n,1}h_{p,2}\Bigg[P_{p,1}\sum_{i\in\mathcal{U}_{p(n)}} P_{i,2} - P_{p,2}\sum_{i\in\mathcal{U}_{n(p)}} P_{i,1}\Bigg] + h_{n,2}h_{p,1}\Bigg[P_{p,2}\sum_{i\in\mathcal{U}_{p(n)}} P_{i,1} - P_{p,1}\sum_{i\in\mathcal{U}_{n(p)}} P_{i,2}\Bigg].$$

Replacing $\mathcal{U}_{n(p)}$ by $\{m\}$, $\mathcal{U}_{p(n)}$ by $\emptyset$, and C_{pn}^{sp} by $\{n\}$, B becomes:

$$\begin{aligned}B =& - h_{n,1}h_{p,2}P_{p,2}P_{m,1} - h_{n,2}h_{p,1}P_{p,1}P_{m,2} - h_{p,1}h_{n,1}P_{p,1}P_{m,1} - h_{p,2}h_{n,2}P_{p,2}P_{m,2}\\ &+(h_{n,1}h_{p,2} - h_{n,2}h_{p,1})(P_{p,1}P_{n,2} - P_{p,2}P_{n,1}).\end{aligned}$$

By adding and subtracting $h_{n,2}h_{p,1}P_{p,2}P_{m,1}$ and $h_{n,1}h_{p,2}P_{p,1}P_{m,2}$ to B, we get:

$$\begin{aligned}B =& h_{n,2}h_{p,1}P_{p,2}P_{m,1} + h_{n,1}h_{p,2}P_{p,1}P_{m,2} - h_{n,1}h_{p,2}P_{p,2}P_{m,1} - h_{n,2}h_{p,1}P_{p,1}P_{m,2}\\ &-h_{n,2}h_{p,1}P_{p,2}P_{m,1} - h_{n,1}h_{p,2}P_{p,1}P_{m,2} - h_{p,1}h_{n,1}P_{p,1}P_{m,1} - h_{p,2}h_{n,2}P_{p,2}P_{m,2}\\ &+(h_{n,1}h_{p,2} - h_{n,2}h_{p,1})(P_{p,1}P_{n,2} - P_{p,2}P_{n,1}).\end{aligned}$$

Combining the terms of the first row together and those of the second row gives:

$$\begin{aligned}B =& (h_{n,2}h_{p,1} - h_{n,1}h_{p,2})P_{p,2}P_{m,1} + (h_{n,1}h_{p,2} - h_{n,2}h_{p,1})P_{p,1}P_{m,2}\\ &- h_{p,1}P_{m,1}(h_{n,1}P_{p,1} + h_{n,2}P_{p,2}) - h_{p,2}P_{m,2}(h_{n,2}P_{p,2} + h_{n,1}P_{p,1})\\ &+ (h_{n,1}h_{p,2} - h_{n,2}h_{p,1})(P_{p,1}P_{n,2} - P_{p,2}P_{n,1}).\end{aligned}$$

Finally, grouping the common factors leads to the final channel and power conditions are given by:

$$B = [P_{p,2}(P_{n,1} + P_{m,1}) - P_{p,1}(P_{n,2} + P_{m,2})]H_{pn} - (h_{n,1}P_{p,1} + h_{n,2}P_{p,2})[h_{p,1}P_{m,1} + h_{p,2}P_{m,2}] > 0.$$

It is therefore confirmed that the new SIC condition at the level of n has an additional negative term compared to (4.19).
The PMCs for the removal of s_m then s_n at the level of user p are respectively:

$$\begin{aligned}P_{m,1}h_{p,1} + P_{m,2}h_{p,2} &> (P_{p,1} + P_{n,1})h_{p,1} + (P_{p,2} + P_{n,2})h_{p,2},\\ P_{n,1}h_{p,1} + P_{n,2}h_{p,2} &> P_{p,1}h_{p,1} + P_{p,2}h_{p,2}.\end{aligned}$$

Also, the PMCs for the removal of s_p then s_m at the level of user n are respectively:

$$\begin{aligned}P_{p,1}h_{n,1} + P_{p,2}h_{n,2} &> (P_{n,1} + P_{m,1})h_{n,1} + (P_{n,2} + P_{m,2})h_{n,2},\\ P_{m,1}h_{n,1} + P_{m,2}h_{n,2} &> P_{n,1}h_{n,1} + P_{n,2}h_{n,2}.\end{aligned}$$

Scenario 3

Users p and n decoded s_m before canceling each other's interference. In this scenario, the conditions of mutual SIC between p and n are exactly the same as in the two-user system

since the third user, m, is taken out of the equation for the two users. Therefore, the mutual SIC constraint is the same as (4.7):

$$(h_{n,1}h_{p,2} - h_{n,2}h_{p,1})\left[P_{p,1}P_{n,2} - P_{p,2}P_{n,1}\right] > 0.$$

The signal of the third user m must be the dominant one at both users p and n. The PMCs of p are as follows:

$$P_{m,1}h_{p,1} + P_{m,2}h_{p,2} > (P_{p,1} + P_{n,1})h_{p,1} + (P_{p,2} + P_{n,2})h_{p,2},$$
$$P_{n,1}h_{p,1} + P_{n,2}h_{p,2} > P_{p,1}h_{p,1} + P_{p,2}h_{p,2}.$$

The PMCs for the removal of s_m then s_p at the level of user n are:

$$P_{m,1}h_{n,1} + P_{m,2}h_{n,2} > (P_{n,1} + P_{p,1})h_{n,1} + (P_{n,2} + P_{p,2})h_{n,2},$$
$$P_{p,1}h_{n,1} + P_{p,2}h_{n,2} > P_{n,1}h_{n,1} + P_{n,2}h_{n,2}.$$

At last, the total power constraints are the same for all eight configurations and they are given by:

$$P_{1,1} + P_{2,1} + P_{3,1} \leq P_{L_1}, \tag{4.20}$$
$$P_{1,2} + P_{2,2} + P_{3,2} \leq P_{L_2}. \tag{4.21}$$

To sum up, our proposed method, namely FullJT-TMSIC, serves all three users using joint transmission and seeks to achieve an interference-free NOMA cluster. For every channel realization, the method solves the problem of sum-rate maximization ($\max R_1 + R_2 + R_3$, where R_i, $i = 1, 2, 3$, are given in (4.2)) eight times with the PMCs and mutual SIC conditions for every corresponding decoding order, while respecting the power limits imposed by the system in (4.20) and (4.21). The algorithm retains the results of the best performing decoding order configuration per channel realization.

4.5.3 Enhancement over the Conventional Approach (CellEdgeJT-TMSIC)

Two major aspects differentiate FullJT-TMSIC from the CellEdgeJT-CellCenterSIC conventional approach: the use of the mutual SIC procedure at all users, and the employment of JT to serve all users. However, the FullJT context is not necessary for achieving TMSIC. Therefore, to assess separately the benefits of JT from those of TMSIC, we propose to use the TMSIC procedure in CellEdgeJT-CellCenterSIC configurations, when possible, calling it CellEdgeJT-TMSIC. In this case, only the cell-edge user is served using JT, while all three users may cancel their mutual interferences.

Compared to CellEdgeJT-CellCenterSIC, CellEdgeJT-TMSIC presents the advantage of using a TMSIC while FullJT-TMSIC presents the advantage of using a complete JT system compared to CellEdgeJT-TMSIC. Moreover, the use of mutual SIC, and more precisely TMSIC, allows the algorithm to reach a solution when the initial CellEdgeJT-CellCenterSIC technique fails because the SIC conditions strongly depend on the channel conditions in (4.17): if the signs of the channel differences don't match, SIC is not possible irrespectively of the power distribution. The PMCs and mutual SIC conditions are directly derived from the ones obtained in section 4.5.2 by letting either $P_{p,1}$ or $P_{p,2}$

(resp. $P_{n,1}$ or P_{n2}, P_{m1} or P_{m2}) be equal to zero, when a cell-center user p (resp. n, m) is concerned. The eight scenarios are then evaluated. However, because of the decrease in the degrees of freedom in the system (in terms of non-zero power variables), the chances of successive triple mutual SIC are lower with CellEdgeJT-TMSIC, compared to FullJT-TMSIC. Therefore, the CellEdgeJT-TMSIC technique first applies TMSIC when possible. If no solution is found, CellEdgeJT-CellCenterSIC is applied. If neither strategy leads to a solution, SIC is abandoned at all users levels, i.e. all the SIC and PMC constraints are relaxed and the rate maximization problem involves the sum rate of interference-full users. Their rates are given by:

$$R_1 = \log_2\left(1 + \frac{P_{1,1}h_{1,1}}{P_{3,1}h_{1,1} + (P_{3,2} + P_{2,2})h_{1,2} + \sigma^2}\right),$$
$$R_2 = \log_2\left(1 + \frac{P_{2,2}h_{2,2}}{(P_{3,1} + P_{1,1})h_{2,1} + P_{3,2}h_{2,2} + \sigma^2}\right),$$
$$R_3 = \log_2\left(1 + \frac{P_{3,1}h_{3,1} + P_{3,2}h_{3,2}}{P_{1,1}h_{3,1} + P_{2,2}h_{3,2} + \sigma^2}\right).$$

4.5.4 On Successful SIC in FullJT-TMSIC and CellEdgeJT-TMSIC

Achieving a complete TMSIC in three-user NOMA clusters using two serving antennas is no longer guaranteed as it was the case for DPS-DMSIC and JT-DMSIC in two-user clusters. In such situations, it is possible to evaluate the alternatives where a smaller number of users operate in mutual SIC while the rest may benefit from single SIC or not. However, this is not the idea of the chapter since rate maximization is not by itself the aim of our work but just a means to measure the effectiveness of combining mutual SIC with CoMP. That is why we revert directly to the NoSIC alternative, as we are only interested in the cases of full triple mutual SIC. It is therefore clear that all three methods are not guaranteed to yield a successful TMSIC implementation for all simulations, and that the effectiveness of the proposed methods are to be measured with respect to both the rate gain provided by TMSIC, and the statistics of TMSIC occurrences.

4.6 Performance Evaluation

Simulations are conducted to evaluate the performance of the presented scenarios and techniques, under the following practical conditions: The outer cell radius of each hexagonal cell is $R_d = 500$ m. The penetration depth of the user 3 zone is of 30 m in each cell (Cf. Fig. 4.1). Three out of the four RRHs (per cell) are spread across the cell, uniformly positioned on a circle of radius $2R_d/3$, while the fourth is located at each cell center. Users are independently positioned, their positions being randomly generated with a uniform probability distribution over their respective regions. The transmission channel model includes a distance-dependent path-loss of decay factor 3.76, and a zero-mean lognormal shadowing with an 8 dB variance. The total bandwidth is $B = 10$ MHz, subdivided over $S = 64$ subbands to yield a subband bandwidth of $B/S = 156.250$ kHz. The power spectral density of the additive background white noise is $N_0 = 4.10^{-18}$ mW/Hz, and the noise power on each subband is $\sigma^2 = N_0 B/S$. The power limit constraints over the serving antennas r_1 and r_2 are varied such that the total available system power P_L (excluding other non-serving RRHs) remains constant throughout the simulations. Unless specified

otherwise, the total power $P_L = P_{L_1} + P_{L_2}$ is 4 W. MATLAB software is used to generate the numerical results and *fmincon* from the optimization toolbox is used to solve the optimization problems in each technique.

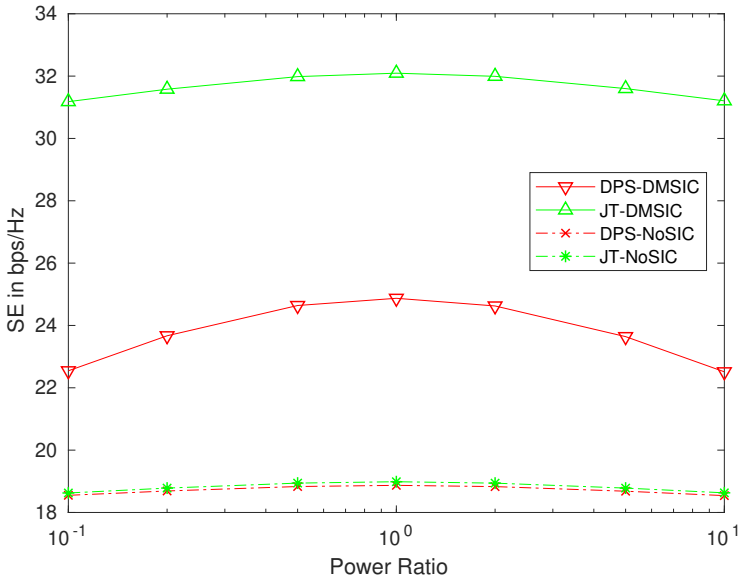


Figure 4.2 – Spectral Efficiency of a two-user system as a function of P_{L_1}/P_{L_2}.

In Fig.4.2, the system Spectral Efficiency (SE) for the different two-user strategies is presented as a function of the antennas power limit ratio P_{L_1}/P_{L_2}. Although antenna power limits of different cells are not usually linked, the chosen representation provides a useful analysis, for network planning, of the best power distribution between the antennas. A first noticeable property is the shape of the curves: all the techniques seem to reach their performance peak at the unity power ratio, implying that the system better handles different user distributions when $P_{L_1} = P_{L_2}$. It should be pointed out, though, that this observation is only true on an average basis, i.e. the optimal power ratio is not necessarily one for every single channel realization.

At the common performance peak ($P_{L_1} = P_{L_2} = 2$ W), an important SE gap between DMSIC and NoSIC algorithms is observed for both the JT and DPS cases. SE gains of 13.1 bps/Hz (69% increase) and 6 bps/Hz (32% increase) are achieved between JT-DMSIC and JT-NoSIC, and between DPS-DMSIC and DPS-NoSIC respectively. This clearly showcases the superiority of the mutual SIC procedure with respect to the common practice of automatically assigning the users to their best antennas which is implicitly done in the NoSIC algorithms as discussed hereafter.

The JT algorithms dominate their DPS counterparts in both DMSIC and NoSIC scenarios. However, the performance gap between DPS-NoSIC and JT-NoSIC is nearly imperceptible. To understand this behavior, we first recall the four possible DPS-NoSIC scenarios of section 4.4.1.2, where users can be served either by the same antenna or by different ones. Any of these four cases can be derived from the JT version of this algorithm by simply setting the appropriate power variables to zero. Once again, JT encompasses all the different DPS scenarios into a broader one. The simulation results reveal that the power allocation scheme that maximizes the total rate for DPS-NoSIC almost always resides in allocating to the user with the best channel gain the entire power P_L of the serving

RRH. The signal of the second user is switched off, whether it is served by the same RRH or not, avoiding thereby the interference that would be caused by that user. In such cases, the enhancement brought by the JT scenario is in the addition of a new signal coming from the second antenna that enhances the reception quality of the user, and thus its rate as well as the total system rate. The increase in power level translates into a marginal rate improvements when the user rate vs. SNR (Signal-to-Noise Ratio) curve is already at a saturation point in DPS. In contrast, the more equitable power distribution between users that takes place in DPS-DMSIC makes the working point of every user quite far from the saturation region of their rate vs SNR curves. This effect is very similar to the waterfilling algorithm where maximizing the total rate is done through shifting some of the available power away from the best link towards another, rather than focusing the whole power on the best link. The only difference here is that, instead of having multiple subbands allocated to one user, the same subband is allocated to two different users at the same time. In this regard, the effect of the DMSIC procedure is virtually doubling the bandwidth of the system without adding interference. Not only does this achieve a much greater fairness and more important sum rates, but it also yields a significant rate improvement when moving from DPS-DMSIC to JT-DMSIC, as shown in Fig. 4.2.

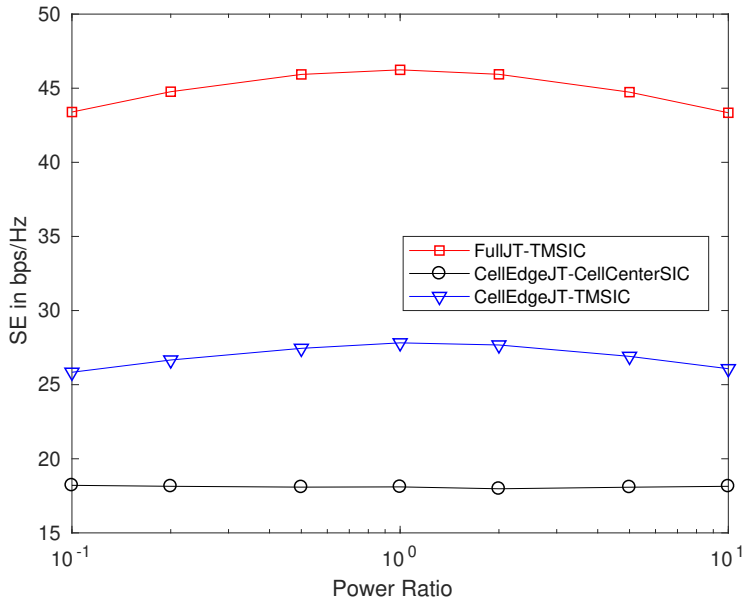


Figure 4.3 – Comparison of the rate maximization procedures for a three-user system.

The performance of the discussed methods for three-user clusters is presented in Fig. 4.3. As stated earlier, a complete mutual SIC procedure is no longer guaranteed in three-user systems, and different techniques lead to different success rates for TMSIC. For our setup, a statistical analysis of the obtained results yields 95% chances of successful mutual SIC in FullJT-TMSIC and 46% in CellEdgeJT-TMSIC. The analysis also shows that even the easier single SIC conditions in CellEdgeJT-CellCenterSIC are not always feasible, with 44% success rate for SIC of the signal of user 3, s_3, at the level of user 1 and user 2.

Comparing CellEdgeJT-CellCenterSIC and CellEdgeJT-TMSIC showcases the enhancements brought by adopting the triple mutual SIC strategy: 18.2 bps/Hz vs. 27.8 bps/Hz

at the peak. Indeed, the rate gain is entirely due to the use of the TMSIC procedure, as no change is carried to the system configuration: in both cases, user 3 is served with JT and users 1 and 2 are served by a single antenna in DPS. This shows that the occurrence of TMSIC is not exclusively bound to a full JT NOMA cluster, and it highlights the ability of TMSIC to increase the total throughput without requiring any technical change in the system. On the other hand, comparing FullJT-TMSIC and CellEdgeJT-TMSIC sheds the light on the importance of a fully JT-based system in enhancing the throughput. This time, the use of JT to serve every user distinguishes FullJT-TMSIC from CellEdgeJT-TMSIC. As in Fig. 4.2, the rate gain due to JT towards DPS is magnified by triple mutual SIC where a rate gain of 18.4 bps/Hz is achieved (66% increase).

Table 4.3 – Jain fairness measurement for three-user systems for $P_{L_1}/P_{L_2} = 1$

	Jain fairness
FullJT-TMSIC	0.97
CellEdgeJT-CellCenterSIC	0.40
CellEdgeJT-TMSIC	0.62

A fairness measurement of the three-user techniques is provided in Table 4.3 for a unit power ratio ($P_{L_1}/P_{L_2} = 1$). The Jain fairness index is used [31]. This index is upper bounded by 1 for absolute fairness scenarios (i.e. all users achieve the same rate on average), and lower bounded by 1/3 which corresponds to the worst case scenario (i.e. a single user is holding all of the system throughput). The fairness index achieved by FullJT-TMSIC approaches the upper bound whereas the CellEdgeJT-CellCenterSIC technique has a poor fairness index (0.40). This shows that not only does FullJT-TMSIC perform best with regards to SE, but it also achieves the highest values of fairness among users. Thanks to the mutual SIC procedure, FullJT-TMSIC achieves a higher system throughput through a fairer distribution of the available power to the users. To better showcase this behavior, the individual rates of users are presented for both FullJT-TMSIC and CellEdgeJT-CellCenterSIC as a function of the total system power in Fig. 4.4. Instead of showing the average individual rate achieved by each user, the averages of the minimum, maximum and middle rates achieved in every simulation are put forward, in order to better emphasize the throughput disparity for the different methods.

It can be seen from Fig. 4.4 that most of the throughput achieved by CellEdgeJT-CellCenterSIC comes from the highest rate user. Indeed, for a total system power of 8W, the minimum and middle rate users account for only 8.3% of the total throughput, compared to the 60% for FullJT-TMSIC. The rate distribution in FullJT-TMSIC is much fairer, each user actively contributing to the system throughput.

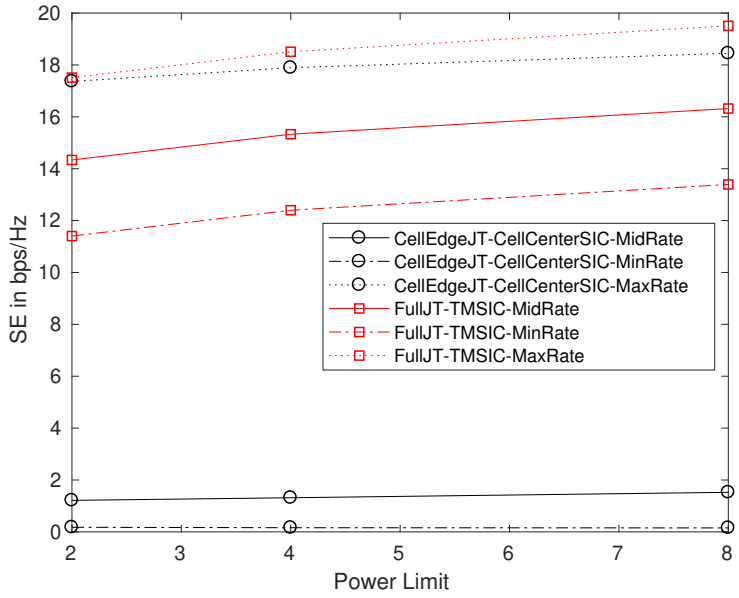


Figure 4.4 – Minimum, maximum and middle individual user rates as a function of the system power for a power ratio equal to one in a three-user system.

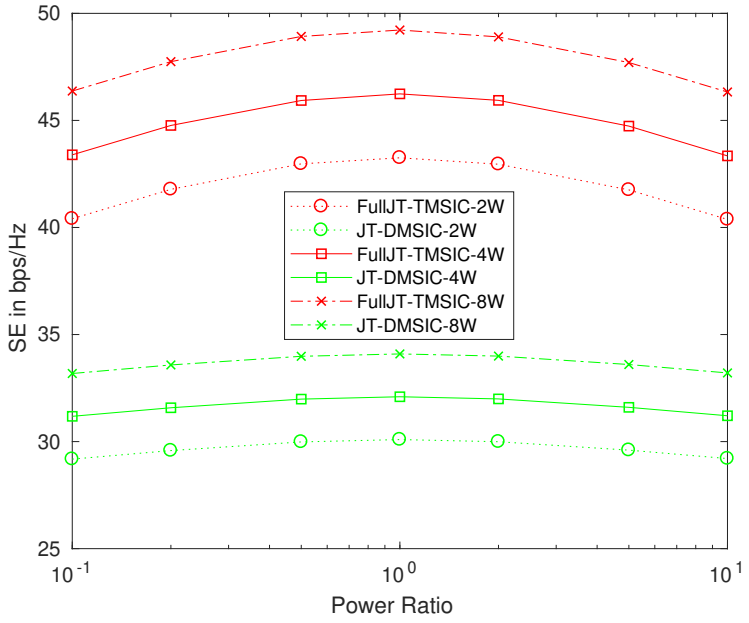


Figure 4.5 – Comparison of the best performing scenario for 2-user vs. 3-user clusters, for $P_L = 2, 4$ and 8 W.

In Fig. 4.5, the best performing approach for two and three-user clusters are compared in the same conditions of power ratios and total system power. Also, to allow a fair comparison, the user deployment is kept unvaried for the two initial users: for every channel realization, users 1 and 2 are randomly deployed according to the system model in Fig. 4.1, and the third user is added to the system without affecting the initial distribution of the two other users. Even under these conditions, the rate gain provided by the third

user accounts to a 44% increase in SE, for a power ratio equal to one when $P_L = 4$ W. This significant increase is not only due to the exploitation of the added diversity by the third user. In fact, being able to serve a third user without causing interference - which is the core of TMSIC - is equivalent to adding to the system an additional virtual subband for exploitation. This was the case for JT-DMSIC compared to JT-NoSIC, and it is also the case of FullJT-TMSIC in comparison with JT-DMSIC. Also, this result contrasts with the general knowledge inherent to classical single-SIC NOMA systems, such as in [27], where it is shown that the performance gain of three vs. two collocated users per subband and powered by the same antenna is rather minor. With a judicious NOMA-DAS employing mutual SIC, the number of users per cluster could be efficiently extended to the limit that can be allowed by both the SIC complexity constraint at receivers, and the large but yet limited computational power available at the BBU for scheduling. Note that for a general NOMA cluster of size M, $M-1$ signals must be decoded at the level of a given user, which can be done in $(M-1)!$ possible orders. The total number of possible decoding orders in the entire cluster is therefore given by $(M-1)!^M$. Due to the exponential increase of the scheduling complexity with the cluster size, the best trade-off is usually attained for two or three-user clusters.

Comparing the rates for different power values, it appears that a linear increase in the throughput occurs for a geometric progression in the total power. This is to be expected given the logarithmic relation between the serving power and the rate (cf. equations (4.1) and (4.2)). Furthermore, it can be observed that rate curves for different power limits are parallel which reinforces the idea that maximum throughput is achieved, on average, for unit power ratios.

As a conclusion, FullJT-TMSIC is by far the best performing technique. Even though particularly restrictive measures on antenna selection were set in our study by limiting the serving antenna choices to r_1 and r_2 in the configuration of Fig. 4.1, an important success rate to establish triple mutual SIC was observed with 95% chances. Furthermore, it is expected that taking advantage of the spatial diversity of each cell by fully exploiting the DAS system would yield even higher percentages of triple mutual SICs. Moreover, when subcarrier assignment is considered, the frequency diversity of the system can be leveraged, enhancing even further the chances of triple mutual SIC. In fact, having observed the efficiency of TMSIC, a new way of user-clustering can be envisioned in which the selection of user 1 and user 2, RRHs r_1 and r_2, and the subband, are based on the cell-edge user, in order to guarantee a TMSIC implementation.

4.7 Conclusion

This chapter focused on the combination of NOMA with CoMP systems to enhance cell-edge user experience as well as the global system performance. We first explored the conditions for a mutual SIC procedure for a general NOMA cluster with two coordinated antennas. The mutual SIC procedure was then applied to two-user and three-user clusters in both DPS and JT. Important performance enhancements were shown in the system throughput (up to 70%) and the user fairness which validate the potential of this technology in reaching current and future challenges imposed by 5G and beyond systems.

The considerable gains of TMSIC suggest building resource allocation schemes of user-antenna-subbands associations that favor TMSIC feasibility above other considerations. Therefore, in the next chapter, antenna positioning problems are considered from the

perspective of TMSIC application.

The contributions of this chapter led to the publication of the following journal paper:

A. Kilzi, J. Farah, C. Abdel Nour and C. Douillard,"Mutual Successive Interference Cancellation Strategies in NOMA for Enhancing the Spectral Efficiency of CoMP Systems," in *IEEE Trans. Commun.*, vol. 68, no. 2, pp. 1213-1226, Feb. 2020.

Chapter 5

Analysis of Drone Placement Strategies for Complete Interference Cancellation in Two-Cell NOMA CoMP Systems

The use of Unmanned Aerial Vehiculess as flying base stations is rapidly growing in the field of wireless communications to leverage the capacity of congested cells. This chapter considers a two-cell system where one of the cells is saturated, i.e. can no longer serve its users, and is supported by a UAV. The UAV positioning procedures are proposed to best alleviate the load on the congested cell with a particular attention directed at enhancing system SE through a fairer serving of cell-edge users as well as cell-centered users of the two adjacent cells. From the experience of the previous chapter, achieving an interference-free user cluster through the application of TMSIC allowed for better system fairness and SE. Therefore, the driving idea of UAV placement, in this part of the study, is to enable TMSIC while taking into account the characteristics of Air-to-Ground (A2G) links in terms of random LoS and NLoS realizations between users and the UAV.

This chapter is organized as follows: section 5.1 discusses the importance of resorting to flying base stations in the context of mobile networks and presents a review of previous work on UAV positioning. Section 5.2 describes the system model and formulates the general UAV placement problem. Section 5.3 introduces the mathematical framework for modeling the UAV positioning problem on a probabilistic basis. In section 5.4, the proposed UAV positioning techniques are presented, while power allocation strategies are described in section 5.5. In section 5.6, the performance results are assessed, and section 5.7 draws the major conclusions of this chapter.
The contributions of this chapter can be summarized as follows:

- We study the UAV positioning problem while taking into account the specificity of LoS/NLoS propagation between users and the UAV, instead of the mean path loss model used in most of the literature.
- We introduce a probabilistic framework that enables the calculation of the TMSIC probability associated to the UAV position. This enables the formulation of a UAV positioning problem to maximize the chances of TMSIC between users.
- We investigate several positioning techniques based on the probabilistic framework

with different optimization criteria, and we compare them to positioning techniques based on the traditional mean path loss consideration. We also highlight the existing trade-offs between system capacity, fairness, and computational complexity of the investigated approaches.

5.1 Related Works

UAVs have lately been gathering interest as a growing research topic for mobile communication networks [117–121]. The major capabilities of UAVs reside in their fast and cost-effective setup and their virtually unconstrained mobility in the aerial space, largely improving the probability of LoS communication. Unlike terrestrial mobile base stations that are bound by road maps and traffic light constraints for circulation, UAVs can move freely through space to cope with the evolving demand for service or network reconfiguration. Many applications require such key capabilities, ranging from natural disaster scenarios like floods, hurricanes and tornadoes, to public safety communication, and temporary crowded events like concerts or festivals in large arenas, sports events in football stadiums, etc. While deploying additional small base stations in anticipation to planned events such as festivals could be profitable for the case of long lasting events (expanding over a few days), it is not suited for dealing with temporary and unpredictable emergency situations typically spanning over the course of a couple of minutes to a few hours. Such scenarios could be rooted to exceptional events like for the cases of disaster relief and service recovery, as well as to much more common congestion scenarios like antenna failure or energy shortage, actual traffic jamming resulting in uneven data traffic loads, etc. Deploying additional small cells especially for that matter equates to large expenditure costs for small periods of time, hence the inefficiency of such approaches. Relying on UAVs for these systems is an appealing feature thanks to their on-demand service capabilities (they can be released and retrieved after use), their adjustable position in real time which can cope with high data traffic variation, and their cost-effective and fast deployment. Therefore, the use of UAVs in the system provides greater flexibility and better preparedness to respond to all sorts of wireless demands occurring in a rather difficult-to-predict manner [122].

Much work has been done on the integration of NOMA into UAV-assisted networks. The authors in [123] study the case of a UAV BS serving a large number of users using NOMA. A simultaneous optimization of the UAV height, the bandwidth allocation to users, the transmit antenna beamwidth and PA is conducted to solve the max-min rate problem using inner convex approximations. The results show that NOMA outperforms OMA in this context, achieving results close to dirty paper coding. However, the UAV's horizontal position is fixed at the center of the cell and the user pairing strategy is based on the Euclidean distance between a far-user and a nearby-user. In contrast, the work in [124] proposes a heuristic pairing strategy for multi-user systems inspired by the optimal PA and UAV placement solution for rate maximization of a single NOMA pair. Bisection search is used afterwards to determine the optimal PA and UAV placement for the maximization of the minimum sum rate of user pairs. A UAV-assisted NOMA network is proposed in [125] where a fixed BS and a UAV cooperate to serve users. The sum rate maximization is accomplished by optimizing the rate of UAV-served users through trajectory and scheduling optimization first, then NOMA precoding is optimized to maximize BS-served user rates. In [126], a UAV is dispatched to upload specific information

to ground BSs that serve uplink users with rate constraints. The objective for the UAV fly path is to complete its mission as quickly as possible. To that end, a *fly-hover-fly* procedure is proposed, coupled with successive convex approximation and uplink NOMA serving is used. The results show that mission completion time is significantly minimized with the proposed NOMA scheme compared to OMA.
The work in [127] focused on studying the performance of the UAV downlink command and control (C&C) channel, for which the 3GPP has defined minimum rate requirements. The study compared the deployment of a UAV for two network architectures: a traditional three-sector BS operating in OMA, and a massive MIMO cellular system operating in multi-user mode (i.e. multiple users scheduled per time-frequency resource). The use of MIMO with UAV improved reliability compared to traditional cells when supporting the data rate requirements of a C&C channel, thus allowing for higher altitude placement of UAVs compared to traditional cells. However, the study also revealed that UAVs can severely degrade the performance achieved by Ground User Equipments (GUE) in MIMO if an uplink power control policy is not applied to protect the GUEs, which stresses the need for coordination between the aerial and ground networks.

Indeed, the integration of UAVs as aerial base stations supporting the ground network will require a better management of the system resources in time and frequency, since the backhaul link between the UAV and the network needs to be established and the hand-off procedures as well as low-latency control need to be guaranteed. Therefore, in the following, we consider a CoMP framework to best evaluate the potential gains provided by UAVs. More specifically, JT-CoMP is assumed where signals are transmitted to each user from multiples TPs.

In the last chapter, we studied the combination of NOMA with CoMP for a two-cell system. A full JT system over NOMA clusters of two and three users was studied showing significant advantages over partial JT (i.e., where JT is only used for cell-edge users and DPS is used for cell-center users). Sending the NOMA signals from different TPs enabled mutual SIC between users, which led to defining the conditions of DMSIC and TMSIC for two or three-user clusters respectively. The obtained interference-free NOMA clusters provided significantly better performance results than classical NOMA schemes in terms of spectral efficiency as well as fairness among users, which suggests positioning the UAV with the aim of favoring TMSIC application. Thus, coupling the interference cancellation capabilities of NOMA with CoMP and the mobility of UAVs aims for an effective ICI cancellation. This ICI cancellation is all the more possible thanks to the management of the UAV mobility and power levels. Indeed, compared to fixed ground base stations, the UAV allows for both a reduction in the needed transmit power (by ensuring higher link qualities than conventional ground-to-BS channels) as well as a localization of interference in the region the UAV is hovering over while serving users.

5.2 System Model

A two-cell system is considered where each cell is originally served by a unique BS located at its center. However, one of the cells is congested in a way that its BS can no longer serve additional users. A UAV is deployed to assist the congested system as shown in Fig. 5.1. The UAV may be controlled by an external controller or the BS of the non-congested cell (cell 1 in Fig. 5.1), which communicates to the UAV its flight path information and power allocation through a backhaul link. The management of the backhaul link to the BS is

not considered in this work and was studied in [128]. In such scenarios, UAV placement generally tends to favor the cell-edge users [129] that suffer from poor channel gains as well as significant potential interference due to the neighboring cell. However, while focusing exclusively on such users tends to boost the inter-user fairness within the cell, system throughput is not optimized and only marginal enhancements would occur on the throughput performance. To strike a balance between fairness and system throughput, cell-edge as well as cell-center users must be considered for the UAV placement problem. Moreover, to take advantage of the cooperation between the cells and to properly manage inter-cell interference, cell-center users from cell 1 and 2 should be considered as well. The interference management can be done through NOMA pairing of users from both cells, as was done in chapter 4. For this purpose, we focus our study on a three-user NOMA cluster formed by a triplet of users selected from regions 1, 2 and 3 of the two cells, as shown in Fig. 5.1, where each user can be representative of a user agglomeration from its respective region.

The fixed BS a_1 serves the users and is assisted by a UAV working as a Mobile Base Station (MBS). The BS and MBS are both equipped with a single transmit antenna. It is assumed that the information to be transmitted for each user is made available at the level of the BS and MBS through the backhaul link, enabling DPS and JT serving in the system. JT-mode is used in the remainder of this chapter, given its superior performance to DPS, as shown in chapter 4.

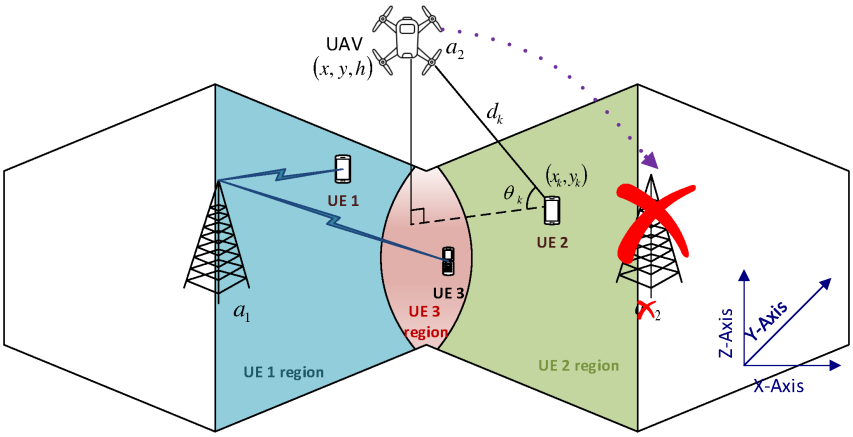


Figure 5.1 – Illustration of the two-cell JT system with the functional base station a_1, the saturated BS in cell 2, the UAV working as MBS a_2, and the three colored user regions.

The objective of this study is to serve the three users such that the resulting channel gains from the UAV position allow the application of TMSIC on their subband. By doing so, system throughput and fairness would be optimized. Note that other users in the system are assumed to be served on different subbands, without causing interference on the considered user triplet. However, the UAV positioning only involves the user triplet that includes the cell-edge user. Additionally, note that despite Fig. 5.1 depicting a CAS, the proposed problem formulation provided next is applicable to distributed network architectures (DAS, small cells, etc.), in which a_1 and a_2 of Fig. 5.1 play the role of two nearby antennas of adjacent cells.

In the following, the path loss model is presented, followed by a reminder on the TMSIC conditions, then the TMSIC solution space is discussed. Afterwards, the UAV placement problem is formulated.

5.2.1 Path Loss Model

The A2G links between users and the UAV are either LoS or NLoS with some probability. Assuming that the UAV is located at position (x, y, h), and that user k is located at position (x_k, y_k) in the ground plane, the path loss for the LoS and NLoS links in dB is given respectively by [130]:

$$L_{\text{LoS}} = 20\log\left(\frac{4\pi f_c d_k}{c}\right) + \eta_{\text{LoS}}, \tag{5.1}$$

$$L_{\text{NLoS}} = 20\log\left(\frac{4\pi f_c d_k}{c}\right) + \eta_{\text{NLoS}}, \tag{5.2}$$

where c is the speed of light in vacuum, f_c is the carrier frequency, d_k is the distance between the UAV and user k (cf. Fig.5.1), η_{LoS} and η_{NLoS} are the average additional losses for LoS and NLoS transmissions. The probability of having a LoS link, P_{LoS}, depends on the angle θ_k formed by the UAV-user k segment and its projection on the ground plane: $\theta_k = \tan^{-1}(h/\sqrt{(x-x_k)^2 + (y-y_k)^2})$. P_{LoS} is modeled as:

$$P_{\text{LoS}} = \frac{1}{1 + \alpha e^{-\beta\left(\frac{180}{\pi}\theta_k - \alpha\right)}}, \tag{5.3}$$

where α and β are constants that depend on the environment (suburban, urban, dense-urban, etc.) parameters [130, 131] such as the ratio of built-up land area to total land area, the number of buildings per unit area, and a scale parameter describing the building's heights distribution. Let $h_{k,i}$ be the squared channel gain between user k and BS a_i. The squared channel gain $h_{k,2}$ between the UAV and user k can be obtained from the experienced path loss L by:

$$h_{k,2} = 10^{-L/10} = \frac{c^2}{(4\pi f_c d_k)^2} \times \begin{cases} 10^{-\eta_{\text{LoS}}} & \text{for } L = L_{\text{LoS}}, \\ 10^{-\eta_{\text{NLoS}}} & \text{for } L = L_{\text{NLoS}}, \end{cases} \tag{5.4}$$

$h_{k,2}$ is then a function of the UAV position as well as the random channel realization regarding the LoS/NLoS nature of the user-UAV link.

5.2.2 Signal Model and TMSIC Conditions

An adequate UAV placement is one that delivers channel links such that TMSIC is rendered feasible in that position. Recall that to enable TMSIC, a set of constraints must be satisfied including PMCs and rate constraints. If we take back the three users notation m, n, and p, the fundamental result from chapter 4 on the condition for decoding a signal s_n, at the level of user p, is to have:

$$\begin{aligned} &h_{n,1}h_{p,1}P_{n,1}\left[\sum_{i\in\mathcal{U}_{n(p)}} P_{i,1} - \sum_{i\in\mathcal{U}_{p(n)}} P_{i,1}\right] + h_{n,2}h_{p,2}P_{n,2}\left[\sum_{i\in\mathcal{U}_{n(p)}} P_{i,2} - \sum_{i\in\mathcal{U}_{p(n)}} P_{i,2}\right] \\ &\quad + (h_{p,1}h_{n,2} - h_{p,2}h_{n,1})\left[P_{n,1}\sum_{i\in C_{pn}^{sn}} P_{i,2} - P_{n,2}\sum_{i\in C_{pn}^{sn}} P_{i,1}\right] \\ &+ h_{p,1}h_{n,2}\left[P_{n,1}\sum_{i\in\mathcal{U}_{n(p)}} P_{i,2} - P_{n,2}\sum_{i\in\mathcal{U}_{p(n)}} P_{i,1}\right] + h_{p,2}h_{n,1}\left[P_{n,2}\sum_{i\in\mathcal{U}_{n(p)}} P_{i,1} - P_{n,1}\sum_{i\in\mathcal{U}_{p(n)}} P_{i,2}\right] > 0. \end{aligned}$$

For a particular decoding order u, six similar rate constraints (two at the level of each user) must be verified to enable TMSIC, hence the corresponding set of SIC constraints $\mathcal{SIC}(u)$ accounting for the decoding order u. However, recall from (5.4) that $h_{k,2}$ depends on the LoS/NLoS realization of the A2G channel of user k. Thus, even for the same decoding order u, the SIC constraints change according to the LoS/NLoS A2G state of the three users in the cluster. These eight possible LoS/NLoS configurations among the users, coupled with the eight potential decoding orders, lead to a total of 64 possible combinations of decoding orders/random channel realizations. The SIC constraints are then denoted by $\mathcal{SIC}(i,u)$ for the i^{th} LoS/NLoS combination and u^{th} decoding order. On the other hand, the PMCs stipulate that if signal s_p is to be decoded prior to the other signals s_n and s_m at the level of a given user (user m or n), the power level of s_p must be greater than the sum of power levels of s_n and s_m. Since in TMSIC every user decodes the signal of the two others before retrieving its own signal, six PMCs must be verified (for any LoS/NLoS combination i and any decoding order u), constituting the set of PMCs denoted by $\mathcal{PMC}(i,u)$. The rate achieved by each user k, when JT-CoMP is used to apply TMSIC between the user triplet, is given by:

$$R_k = B\log_2\left(1+\frac{\sum_{i=1}^{2}P_{k,i}h_{k,i}}{N_0 B}\right), \tag{5.5}$$

where B is the subband bandwidth, and N_0 is the power spectral density of additive white Gaussian noise. A final set of constraints is to account for the transmit power limits of a_1 and a_2 referred to as P_{L_1} and P_{L_2}:

$$\begin{aligned} P_{1,1}+P_{2,1}+P_{3,1} &\leq P_{L_1},\\ P_{1,2}+P_{2,2}+P_{3,2} &\leq P_{L_2}. \end{aligned} \tag{5.6}$$

The first inequality accounts for the sum of the users powers over antenna a_1, and the second one accounts for the sum power over antenna a_2. The problem then resides in finding the positions of the UAV such that: 1) the PMCs, 2) the mutual SIC constraints, and 3) the total transmit power constraints are satisfied.

5.2.3 TMSIC Solution Space

When TMSIC feasibility is targeted, the problem at hand can be seen as admitting several constraints with no objective function, and is therefore a Constraint Satisfaction Problem (CSP) [132]. In other words, one would seek the set of UAV positions where TMSIC is feasible while respecting the constraints. We denote by $r_{i,n}$ the region of space in which the UAV can be placed such that TMSIC is possible, for the n^{th} decoding order, and the i^{th} LoS/NLoS configuration. If we let $\mathbb{D}$ be the allowed space region for UAV positioning, then the CSP for a combination (i,n) can be cast as:

$$\mathcal{CSP}_{i,n}: r_{i,n} = \{\boldsymbol{pos} \in \mathbb{D} / \mathcal{PMC}(i,n), \mathcal{SIC}(i,n), (5.6)\},$$

with $\boldsymbol{pos}$ being the UAV position. Note that the search is explicitly done over the UAV position, but also implicitly over the power variables which are included in $\mathcal{PMC}(i,n), \mathcal{SIC}(i,n)$. In order to determine the entire region in which TMSIC is guaranteed, the procedure below must be followed:

- Solve the CSP for all 64 combinations. Let r_i be the solution space corresponding to the i^{th} LoS/NLoS configuration, obtained by $r_i = \bigcup_{n=1}^{8} r_{i,n}$.
- The space in which a TMSIC is guaranteed to occur is a region where TMSIC is possible for any channel realization of LoS/NLoS combinations. Therefore, the solution of the CSP shall be obtained by $\mathbb{S} = \bigcap_{i=1}^{8} r_i = \bigcap_{i=1}^{8} \left(\bigcup_{n=1}^{8} r_{i,n} \right)$.

Although these resolution steps provide meaningful insights into the spatial representation of TMSIC-enabled regions, note that solving these CSPs is done by Set Inversion Via Interval Analysis (SIVIA) [133]. The latter operates on set intervals using the branch-and-prune method, leading to an exponential complexity on the search space dimensions (nine in our case: three UAV position variables and six power variables) and the required resolution error. For our system parameters, the CSP resolution is practically inapplicable. Most importantly, the existence of a TMSIC-guaranteed space region is not guaranteed due to the r_i intersections which may yield an empty space $\mathbb{S}$. In fact, not only TMSIC may not be guaranteed ($\mathbb{S} = \emptyset$), but the regions $r_{i,n}$ themselves might be empty. If $\forall n, r_{i,n} = \emptyset$, then TMSIC cannot be achieved when the i^{th} LoS/NLoS combination occurs. If this is the case for all LoS/NLoS combinations, then TMSIC application is impossible for the considered user triplet and antenna power limits.

While the CSP complexity can be worked around by turning the CSP into an optimization problem, the problem of TMSIC feasibility has to be addressed. Between the extreme cases of impossible TMSIC application and TMSIC-guaranteed application, there is a middle ground in which it is best to assess the TMSIC application in probabilistic terms. To that end, in the next sections the UAV placement problem is first remodeled into an optimization problem, then the probabilistic TMSIC framework is developed.

5.2.4 UAV Placement Problem Formulation

Optimization problems are at the core CSPs associated to an objective function. A family of optimization problems having different objective functions but with the same constraints (the core CSP) leads to different solutions from one another, but within the same solution space of the aforementioned CSP. For the case at hand, setting the optimization problem with constraints $\mathcal{PMC}(i,n), \mathcal{SIC}(i,n)$ and (5.6) automatically leads to a solution within the desired region $r_{i,n}$ without requiring the knowledge of the whole region. Let f be the optimization function to be carefully selected by the system administrator; the generic formulation of the UAV placement problem becomes:

$$\mathcal{OP}^1_{i,n}: \quad \{\boldsymbol{pos}^*_{i,n}\} = \underset{\boldsymbol{pos}, P_{k,r}}{\arg\max}\, f(\boldsymbol{pos}, P_{k,r}), \tag{5.7}$$

such that: $\mathcal{PMC}(i,n), \mathcal{SIC}(i,n)$ and (5.6) are verified.

Then, the best UAV position is retained:

$$\begin{aligned} \{i^*, n^*\} = & \underset{(i,n)\in[\![1..8]\!]\times[\![1..8]\!]}{\arg\max} f(\boldsymbol{pos}^*_{i,n}), \\ & \boldsymbol{POS} = \boldsymbol{pos}^*_{i^*,n^*}. \end{aligned} \tag{5.8}$$

While this approach does not deliver the entire $r_{i,n}$, it guarantees that $\boldsymbol{pos}^*_{i,n}$ is inside $r_{i,n}$. However, if no solution exists, then it can be affirmed that $r_{i,n}$ is empty, i.e. TMSIC

is impossible to achieve for combination (i, n). This is true independently of f since the optimization function does not affect the feasibility of the problem that is set by its constraints. Therefore, performing UAV positioning by trying to solve $\mathcal{OP}^1_{i,n}$ for all the combinations is by itself a TMSIC feasibility check.

It is worth mentioning that regions $r_{i,n}$ and r_i are not assumed to be known, but their use in the discussion is for modeling convenience and for a better understanding of the problem characteristics through a spatial representation of the discussed properties.

In the next section, the probabilistic framework is discussed in order to provide meaningful insights for the selection of the optimization function by the system administrator.

5.3 Probabilistic Framework for TMSIC-Based UAV Positioning

To determine the TMSIC probability associated with the position of a UAV, let $l = 1$ be the value assigned to the state of a LoS link and $l = 0$ to that of an NLoS state. Given the three-bit binary vector (l_1, l_2, l_3) representing the state of the A2G links of users 1, 2, and 3 respectively, we denote by c_i the i^{th} combination that corresponds to its three-bit binary vector in base two plus one, $i = (l_1, l_2, l_3)_2 + 1$. For instance, the all LoS state is represented by c_8, and the all NLoS state is represented by c_1. The space region r_i corresponds then to combination c_i. To define and evaluate the probability of TMSIC for a UAV position, let us consider the probability of achieving TMSIC through c_i:

$$Pr(\text{TMSIC} \cap c_i/\boldsymbol{pos}) = Pr(c_i/\boldsymbol{pos}) \times Pr(\text{TMSIC}/c_i, \boldsymbol{pos}).$$

Analyzing these terms, we state that knowing c_i and $\boldsymbol{pos}$, the probability of having TMSIC is given by:

$$p_i(\boldsymbol{pos}) \triangleq Pr(\text{TMSIC}/c_i, \boldsymbol{pos}) = \begin{cases} 1, \text{ if } \boldsymbol{pos} \in r_i \\ 0, \text{ else.} \end{cases}$$

In other words, for a fixed c_i and a known UAV position, TMSIC is deterministic and not random, it is either feasible or not according to the belonging of $\boldsymbol{pos}$ to r_i. On the contrary, for a fixed UAV position and fixed user positions, c_i is random and any of the eight link states is possible; however, some LoS/NLoS configurations are more likely to occur than others. Since user positions are mutually independent, the probability of having c_i knowing $\boldsymbol{pos}$ is the product of the probabilities of having the channel state of each user matching that of c_i:

$$\begin{aligned} Pr(c_i/\boldsymbol{pos}) &= Pr(l_1/\boldsymbol{pos}) \times Pr(l_2/\boldsymbol{pos}) \times Pr(l_3/\boldsymbol{pos}) \\ &= \prod_{j=1}^{3} \left[l_j P_{\text{LoS}}(\theta_j) + (1 - l_j) P_{\text{NLoS}}(\theta_j) \right], \end{aligned}$$

where $P_{\text{NLoS}}(\theta_j) = 1 - P_{\text{LoS}}(\theta_j)$. Then by applying the law of total probability, the probability of having a TMSIC for a given UAV position is:

$$Pr(\text{TMSIC}/\boldsymbol{pos}) = \sum_{i=1}^{8} Pr(c_i/\boldsymbol{pos}) \times p_i(\boldsymbol{pos}). \tag{5.9}$$

This clearly shows that the probabilistic nature of TMSIC is bound to the random A2G channel realization and not to the TMSIC procedure itself. Indeed, if the UAV position is fixed and the LoS/NLoS realization c_j is known, the TMSIC procedure is either possible (for at least one decoding order) or it is not (for any of the decoding orders). Therefore, the UAV position directly affects the TMSIC probability through $Pr(c_i/\boldsymbol{pos})$, provided that $p_i(\boldsymbol{pos}) = 1$, which is translated into the satisfaction of the constraints of $\mathcal{OP}^1_{i,n}$. We conclude that the TMSIC probability expression in (5.9) shows that the UAV placement can be made to optimize the TMSIC probability by incorporating this probability into the optimization function f. Based on this fact, the UAV positioning strategies are presented in the next section.

On another hand, once in position, the UAV can determine the actual channel realization c_j through channel estimation by comparing the actual channel gains with the theoretical one in (5.4). Furthermore, if the obtained c_j is different from the channel realization c_{i^*} that yields $\boldsymbol{POS}$ in (5.8), not much can be said about the feasibility of TMSIC for c_j. Indeed, the only available information regarding the TMSIC applicability in $\boldsymbol{POS}$ is that $p_{i^*}(\boldsymbol{POS}) = 1$, but $p_j(\boldsymbol{POS})$ is not known. This can only be determined once the UAV position is fixed and the optimization in (5.7) is rerun for all the decoding orders. This justifies thereby the separation between the UAV placement phase, dealt with in section 5.4, from the power allocation phase which is presented in section 5.5.

5.4 Proposed UAV Positioning Techniques (UPT) based on TMSIC

In this section, we present the different strategies that can be used to position the UAV. The approaches derived from the LoS/NLoS path loss model are presented first, in sections 5.4.1 to 5.4.3. Alternatively, the UPT based on the mean path loss model is presented in section 5.4.4. In both cases, TMSIC positioning is attempted, if TMSIC turns out to be impossible, we revert to a common positioning technique in section 5.4.5.

5.4.1 Maximum Probability Positioning (MPP)

In order to maximize the TMSIC probability, the objective function should be set equal to (5.9). Since the r_i regions cannot be known, $p_i(\boldsymbol{pos})$ is not available for any UAV position $\boldsymbol{pos}$. This causes a problem to the TMSIC probability expression since we don't know which LoS/NLoS combinations to account for in (5.9). Nonetheless, following the constraints of $\mathcal{OP}^1_{i,n}$, the only region the UAV is guaranteed to be in after optimization is $r_{i,n}$, thus f is set to $Pr(c_i/\boldsymbol{pos})$ instead of the total TMSIC probability $Pr(\text{TMSIC}/\boldsymbol{pos})$. Therefore, the original optimization using objective function (5.9) is replaced by an optimization over a lower bound of (5.9). The UAV placement problem is then written as follows:

$$\mathcal{OP}^{1,a}_{i,n} : \{\mathcal{OP}^1_{i,n}, f = Pr(c_i/\boldsymbol{pos})\} \tag{5.10}$$

such that: $\mathcal{PMC}(i,n), \mathcal{SIC}(i,n)$ and (5.6) are verified.

The final UAV position is obtained from (5.8). Given that the remaining combinations ($c_i \neq c_{i^*}$) are not taken into account in $Pr(\text{TMSIC}/\boldsymbol{POS})$, the computed TMSIC probability $Pr(c_{i^*}/\boldsymbol{POS})$ is only a lower bound to the actual TMSIC probability $Pr(\text{TMSIC}/\boldsymbol{POS})$. The obtained lower bound achieves optimality, i.e. $Pr(\text{TMSIC}/\boldsymbol{POS})$ equals $Pr(c_{i^*}/\boldsymbol{POS})$,

when $\boldsymbol{POS} \in r_{i^*}$ and $\boldsymbol{POS} \notin \cup_{i=1,i\neq i^*}^{8} r_i$. However, since the combination c_{i^*} leading to $\boldsymbol{POS}$ is not known in advance, the only situation where the solution to $\mathcal{OP}_{i,n}^{1,a}$ is guaranteed to achieve optimality is when the r_i regions are pairwise disjoint.

5.4.2 Maximum Rate Positioning (MRP)

When solving $\mathcal{OP}_{i,n}^{1,a}$, the obtained UAV position guarantees the highest TMSIC probability without taking into account the resulting achievable throughput. Another approach to UAV positioning is based on the maximum achievable throughput via TMSIC. That way, if a UAV position enabling TMSIC exists for the given user cluster, both MPP and MRP deliver UAV positions enabling TMSIC, but with different values of the associated throughput and of the lower bound on TMSIC probability. Let $R_{TMSIC} = \sum_{k=1}^{3} R_k$ be the total throughput achieved when TMSIC is enabled. The MRP problem takes the following form:

$$\mathcal{OP}_{i,n}^{1,b} : \{\mathcal{OP}_{i,n}^{1}, f = R_{TMSIC}\} \tag{5.11}$$

such that: $\mathcal{PMC}(n,i), \mathcal{SIC}(n,i)$ and (5.6) are verified,

and the final UAV position is obtained from (5.8).

5.4.3 Maximum Probability and Rate Positioning (MPRP)

In section 5.4.1, the position obtained through MPP yields the highest TMSIC probability; however, it does not hold any guarantee with regards to the achievable throughput. In contrast, when the system throughput is favored, as in section 5.4.2, the results may give UAV positions with high throughput but low TMSIC probability. Therefore, instead of aiming at maximizing the chances of TMSIC or the system throughput alone, the UAV is positioned such that the product of the rate by the associated probability is maximized:

$$\mathcal{OP}_{i,n}^{1,c} : \{\mathcal{OP}_{i,n}^{1}, f = Pr(c_i/\boldsymbol{pos})R_{TMSIC}\} \tag{5.12}$$

s.t: $\mathcal{PMC}(n,i), \mathcal{SIC}(n,i)$ and (5.6) are verified.

Compared to other UAV positioning techniques seeking TMSIC, this approach has the advantage of accounting for both the throughput associated to a combination c_i, as well as its probability of occurence. On the other hand, the obtained position does not favor TMSIC as much as MPP solutions. Another approach to position the drone relying on the mean path loss instead of the LoS/NLoS combination is developed next as an alternative to MPP, MRP and MPRP.

5.4.4 Mean Path Loss Positioning (MPLP)

Most works on flying base stations [120, 128, 134] are based on the mean path loss of A2G channels to perform scheduling tasks. The mean path loss of A2G links is given by:

$$L_{av} = P_{\text{LoS}} L_{\text{LoS}} + P_{\text{NLoS}} L_{\text{NLoS}}.$$

The A2G links in this case are no longer defined by the three-bit vector (l_1, l_2, l_3) of LoS/NLoS combinations. The whole concept of LoS/NLoS combinations (c_i) and regions (r_i) becomes irrelevant since a unique expression is available for every user-UAV link. Therefore, the PMCs and SIC conditions depend only on the decoding order, hence the

notations $\mathcal{SIC}(n), \mathcal{PMC}(n)$. Achieving TMSIC cannot be formulated as a probability maximization problem that depends on the different LoS/NLoS combinations: for the given user triplet, either TMSIC is achieved or it is not. However, to avoid running into another CSP, we consider the system throughput objective function and search for the UAV position that maximizes it as follows:

$$\mathcal{OP}_n^2 : \quad \{\boldsymbol{pos}_n^*\} = \underset{\boldsymbol{pos}, P_{k,r}}{\arg\max}(R_{TMSIC})$$

such that: $\mathcal{SIC}(n), \mathcal{PMC}(n)$, and (5.6) are verified.

Even though the objective function does not compromise the feasibility of the solution in any way (no additional constraints are involved), it affects the position of the UAV and therefore the performance of the obtained solution in terms of achieved TMSIC probability and throughput. In fact, this issue is not specific to rate maximization, i.e. any other objective function would have been subject to the same inconvenience. The reason for that is the use of an average channel model to obtain the UAV position. Having obtained the drone position for every decoding order (when the system admits a solution), the position yielding the maximum value of the objective function is selected:

$$\begin{aligned} \{n^*\} &= \underset{n \in [\![1..8]\!]}{\arg\max}(R_{TMSIC}), \\ \boldsymbol{POS} &= \boldsymbol{pos}_{n^*}^*. \end{aligned} \tag{5.13}$$

When comparing the procedure for $\boldsymbol{POS}$ assignment in (5.13) to the procedure used for MPP, MRP and MPRP in (5.8), an eight-fold complexity decrease is observed using the mean path loss model in MPLP. The 64 combinations of decoding orders and LoS realizations that had to be solved turn into 8 combinations of the unique channel realization – i.e. the mean path loss channel – with the decoding orders. This difference will be accounted for when discussing the selection of the best UAV positioning technique in the performance assessment (section 5.6).

5.4.5 Probabilistic Approach Based on Subband Splitting Positioning (SSP)

When TMSIC proves to be impossible (cf. section 5.2.4), an alternative UAV positioning technique shall be used. Its expected properties are the guarantee of a solution for any user positions and a reduced complexity compared to TMSIC. In the previous chapter, DMSIC on the same subband was shown to be always possible when serving the users with two different BSs. Therefore, in case of TMSIC impossibility, we propose to divide the subband into two equal half subbands (supposed to have equal channel gains), and then to pair the cell-edge user (UE 3 of Fig. 5.1) with one of the cell-center users (UE 1 or UE 2 of Fig. 5.1) on each half subband. This leads to two independent pairs of users applying DMSIC separately on each subband. Their PMCs are:
PMCs for DMSIC between (UE 1,UE 3)

$$\begin{cases} P_{3,1,1}h_{1,1} + P_{3,2,1}h_{1,2} > P_{1,1,1}h_{1,1} + P_{1,2,1}h_{1,2} \\ P_{1,1,1}h_{3,1} + P_{1,2,1}h_{3,2} > P_{3,1,1}h_{3,1} + P_{3,2,1}h_{3,2} \end{cases} \tag{5.14}$$

PMCs for DMSIC between (UE 2,UE 3)

$$\begin{cases} P_{3,1,2}h_{2,1} + P_{3,2,2}h_{2,2} > P_{2,1,2}h_{2,1} + P_{2,2,2}h_{2,2} \\ P_{2,1,2}h_{3,1} + P_{2,2,2}h_{3,2} > P_{3,1,2}h_{3,1} + P_{3,2,2}h_{3,2} \end{cases} \tag{5.15}$$

where the additional index d in the power terms $P_{k,r,d}$ refers to the used half subband. UEs 1 and 3 are paired on the first half subband ($d = 1$), and users 2 and 3 are paired on the second half subband ($d = 2$). Note that DMSIC constraints are met when the PMCs are satisfied, as it was proven in chapter 4, section 4.4.2. Moreover, a single decoding order is possible at the level of every user in the respective half subband, hence the positioning problem needs to be solved only once for every c_i. Then, similarly to MPRP, the UAV placement aims at maximizing the product of the DMSIC throughput by the c_i probability. The following problem is solved for the eight c_i channel realizations and, then, the resulting position of the combination leading to the highest value is selected.

$$\mathcal{OP}_i^3 : \quad \{\boldsymbol{pos}\}^* = \underset{\boldsymbol{pos}, P_{k,a,m}}{\arg\max} \left(R_{DMSIC} \times Pr(c_i/\boldsymbol{pos})\right), \tag{5.16}$$

such that: (5.14), (5.15) and $\begin{cases} P_{1,1,1} + P_{2,1,2} + P_{3,1,1} + P_{3,1,2} \leq P_{L_1} \\ P_{1,2,1} + P_{2,2,2} + P_{3,2,1} + P_{3,2,2} \leq P_{L_2}, \end{cases}$

where the system throughput R_{DMSIC} is given by:

$$R_{DMSIC} = \sum_{k\in\{1,3\}} \frac{B}{2} \log_2 \left(1 + \frac{\sum_{r=1}^{2} P_{k,r,1}h_{k,r}}{N_0 B/2}\right) + \sum_{k\in\{2,3\}} \frac{B}{2} \log_2 \left(1 + \frac{\sum_{r=1}^{2} P_{k,r,2}h_{k,r}}{N_0 B/2}\right).$$

This positioning technique is only used when the chosen TMSIC positioning technique (MPP, MPRP, or MPLP) fails to provide a solution.

5.5 Power Allocation Strategy

We present hereafter the global PA approach that is applied at the level of the BS of cell 1 and instructed to the UAV to maximize system throughput. The approach resides in applying TMSIC when possible, otherwise alternative non-TMSIC PAs are used. In the following, we detail how the global PA approach is adapted according to the Alternative Power Allocation Technique (APAT) when TMSIC is not feasible, and the UPT. The flow chart describing the complete power allocation strategy is presented in Fig. 5.2.

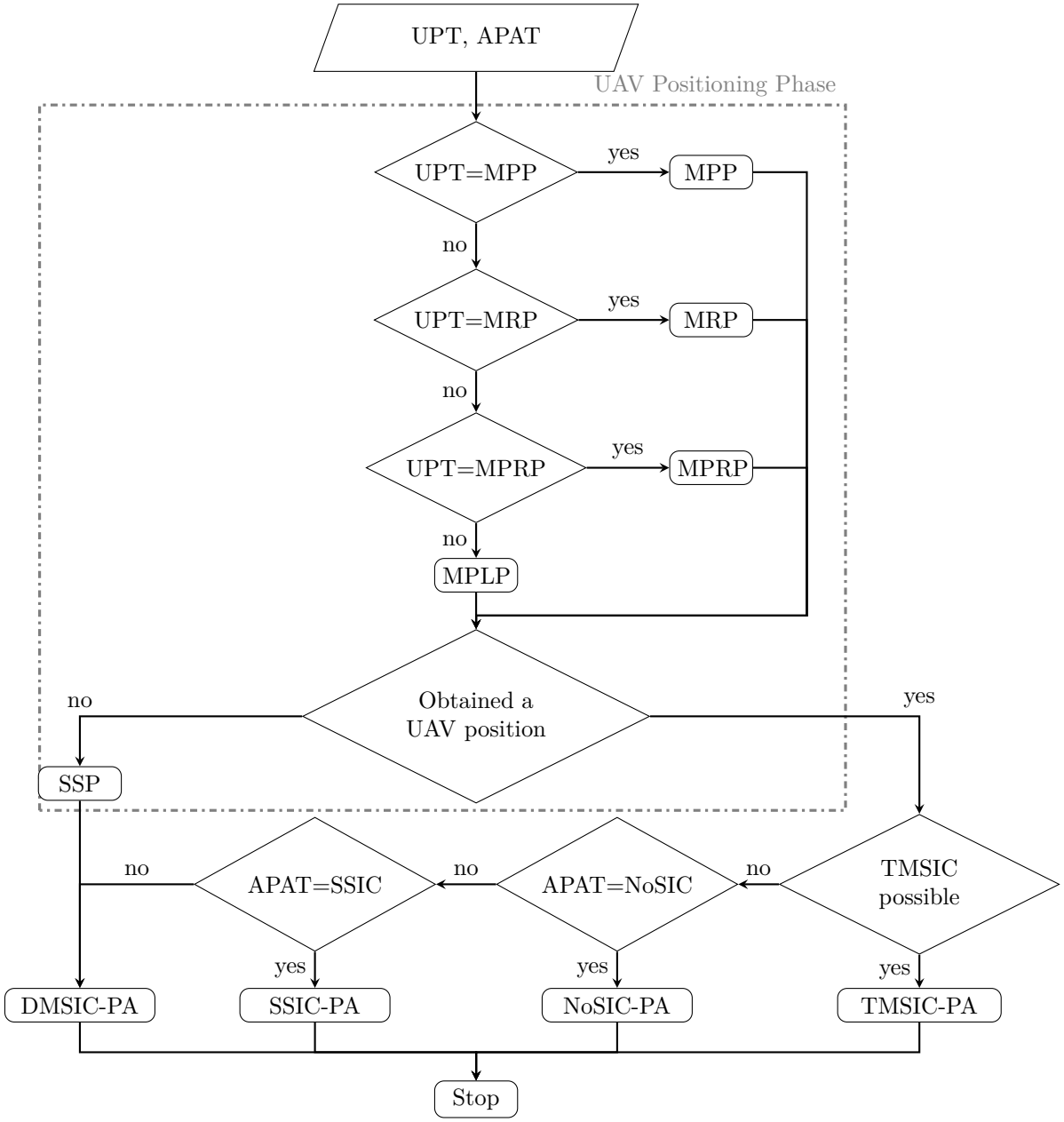


Figure 5.2 – Flow chart of the global strategy for the different UPT-APAT pairs selected by the system administrator.

In section 5.2.4, we explained that performing UAV positioning by trying to solve the variants of $\mathcal{OP}^1_{i,n}$ is a TMSIC feasibility check. Through that check, empty $r_{i,n}$ regions are determined. If all the regions are empty, i.e. if no UAV position is obtained, the check fails and the non-TMSIC PAs of section 5.5.2 are applied. If a UAV position is obtained, then TMSIC PA might be feasible, thus TMSIC PA is attempted.

5.5.1 TMSIC Power Allocation and TMSIC Testing

If a UAV position is obtained, three cases are identified according to three quantities:

- c_{i^*}, the channel realization which leads to ***POS***,

- c_j, the actual channel realization obtained after positioning the UAV,
- $\mathcal{N}$, the set of decoding orders for which $r_{j,n}$ exists: $\mathcal{N} = \{n \in [\![1,8]\!] / r_{j,n} \neq \emptyset\}$.

The three cases are:

1) $c_j = c_{i^*}$: TMSIC-PA is feasible, and the PA problem $\mathcal{OP}^4_{j,n}$ is solved for the decoding orders in $\mathcal{N}$ (which cannot be empty).

2) $c_j \neq c_{i^*}, \mathcal{N} \neq \emptyset$: TMSIC-PA might be feasible, and we need to solve $\mathcal{OP}^4_{j,n}$ for $n \in \mathcal{N}$ to check its feasibility.

3) $c_j \neq c_{i^*}, \mathcal{N} = \emptyset$: TMSIC-PA is not feasible; in this case, we must revert to non-TMSIC PAs (section 5.5.2).

Note that, once the UAV is positioned, the PA does not affect the TMSIC probability, so the optimization function is the same for MPP, MRP, MPRP and MPLP which targets throughput maximization:

$$\mathcal{OP}^4_{j,n}: \quad \{P^*_{k,r}\} = \arg\max_{P_{k,r}} \left(R_{TMSIC}\right),$$

such that: $\mathcal{PMC}(j,n), \mathcal{SIC}(j,n)$ and (5.6) are verified.

In the second case, achieving TMSIC is not guaranteed because $\boldsymbol{POS}$ might be outside of the $r_{j,n}$ ($n \in \mathcal{N}$) regions. That is why $\mathcal{OP}^4_{j,n}$ needs to be solved to determine if TMSIC is feasible. In the MPLP case, the existence of the $r_{i,n}$ regions has not been tested during the UAV positioning phase (as it is the case for MPP, MRP and MPRP), hence $\mathcal{OP}^4_{j,n}$ is solved/checked for all the decoding orders. These differences are pictured in the flowchart of Fig. 5.3.

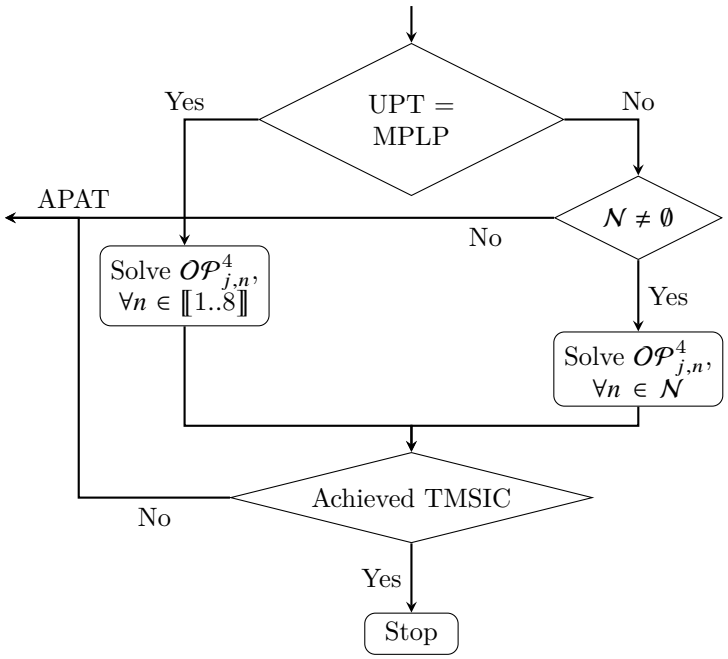


Figure 5.3 – Detailed flow chart of the testing and the TMSIC-PA blocks of Fig. 2.

5.5.2 Alternative Power Allocation Techniques

In the case where TMSIC is not feasible, several PA alternatives for system throughput maximization are possible. Based on the principle that the achieved rate increases when interference cancellation is successfully conducted, it is natural to seek the highest number of SIC procedures between the three users. Since TMSIC corresponds to 6 SICs, two at the level of each user, when one SIC fails, we can apply a 5-SIC procedure. Following this pattern, 5 SICs, than 4, 3, 2 and a single SIC must be all tried in that order until the first setup that leads to a valid PA solution. This ideal strategy counts 112 potential problems to be solved when taking into account all possible decoding orders for every case. Due to the number and complexity of these problems, this strategy is disregarded. Besides, the success of this strategy is not guaranteed, just like it was not the case for TMSIC. Three alternative non-TMSIC PAs are proposed. When a TMSIC procedure is declared infeasible after undergoing the tests in section 5.5.1, the BS of cell 1 executes one of the following PA schemes: DMSIC, NoSIC, or Single SIC (SSIC).

5.5.2.1 DMSIC

Following the reasoning of section 5.4.5, we resort to subband division followed by DMSIC, with the difference that DMSIC is now used for PA and not for UAV positioning. The DMSIC-PA problem takes the following form:

$$\mathcal{OP}^{5,a}: \quad \{P_{k,r,d}\}^* = \underset{P_{k,r,d}}{\arg\max} \left(R_{DMSIC}\right), \tag{5.17}$$

such that the constraints of $\mathcal{OP}_i^3$ are verified.

Since the UAV position has been fixed previously, $\mathcal{OP}^{5,a}$ is solved only once for the obtained configuration c_j (unlike $\mathcal{OP}_i^3$ that is solved for all combinations in section 5.4.5) and the resulting power allocation is instructed to the UAV by the BS.

5.5.2.2 NoSIC

Without dividing the subband, a simpler alternative to TMSIC resides in abandoning all SIC procedures and solving the new rate maximization problem without any other system constraints than the total transmit power of BSs. Users signals interfere on one another and the problem formulation is given by:

$$\mathcal{OP}^{5,b}: \quad \{P_{k,r}\}^* = \underset{P_{k,r}}{\arg\max} \left(\sum_{k=1}^{3} B \log_2(1 + \frac{\sum_{r=1}^{2} P_{k,r} h_{k,r}}{\sum\limits_{k'=1,k'\neq k}^{3} P_{k',r} h_{k',r} + N_0 B}) \right)$$

such that (5.6) is satisfied.

5.5.2.3 SSIC

Standard NOMA SIC procedures may also be used when TMSIC is impossible. In this case, the strong users in the two cells, i.e. UE 1 and UE 2, successfully decode the signal of the weak user UE 3 that cannot perform SIC. This interference cancellation scheme is similar to the NOMA-CoMP system adopted in [30] (and used as benchmark for chapter

4), with the difference that in our system all users are served through JT-CoMP (and not only the cell-edge user). The corresponding optimization problem is:

$$\mathcal{OP}^{5,c} : \{P_{k,r}\}^* = \arg\max_{P_{k,r}} \left[B\log_2(1 + \frac{P_{3,1}h_{3,1} + P_{3,2}h_{3,2}}{(P_{1,1}+P_{2,1})h_{3,1} + (P_{1,2}+P_{2,2})h_{3,2} + N_0B}) \right.$$
$$\left. + B\log_2(1 + \frac{P_{1,1}h_{1,1} + P_{1,2}h_{1,2}}{P_{2,1}h_{1,1} + P_{2,2}h_{1,2} + N_0B}) + B\log_2(1 + \frac{P_{2,1}h_{2,1} + P_{2,2}h_{2,2}}{P_{1,1}h_{2,1} + P_{1,2}h_{2,2} + N_0B}) \right]$$

such that:

- SIC of the signal of U3 is guaranteed at the level of U1 and U2 respectively:
  $(h_{1,1}h_{3,2} - h_{1,2}h_{3,1})[P_{3,1}(P_{2,2}+P_{1,2}) - P_{3,2}(P_{1,1}+P_{2,1})] > 0$
  $(h_{2,1}h_{3,2} - h_{2,2}h_{3,1})[P_{3,1}(P_{2,2}+P_{1,2}) - P_{3,2}(P_{1,1}+P_{2,1})] > 0$

- PMC constraints are verified at the level of U1 and U2 respectively:
  $P_{3,1}h_{1,1} + P_{3,2}h_{1,2} > (P_{1,1}+P_{2,1})h_{1,1} + (P_{2,2}+P_{1,2})h_{1,2}$
  $P_{3,1}h_{2,1} + P_{3,2}h_{2,2} > (P_{2,2}+P_{1,2})h_{2,2} + (P_{1,1}+P_{2,1})h_{2,1}$

- Power limit constraints are satisfied as in (5.6).

Note that the SIC and PMC derivations for this case are directly derived form equations (4.15) and (4.16) of section 4.5.1, but without canceling out $P_{2,1}$ and $P_{1,2}$ thanks to JT serving. We note also that the condition in (4.17) on the identical sign of the channel terms to enable SIC still holds:

$$\text{sign}(h_{1,1}h_{3,2} - h_{1,2}h_{3,1}) = \text{sign}(h_{2,1}h_{3,2} - h_{2,2}h_{3,1}).$$

If the users channel gains do not comply with this condition, the single SIC procedure cannot work, and the PA scheme reverts to NoSIC-PA.

As stated in the beginning of this section, the first aim of the presented PA procedures is the accomplishment of a successful TMSIC. In other words, APAT is applied as a backup solution just like SSP was for MPP, MRP, MPRP and MPLP. In the performance assessment section, the nomenclature of the resource allocation techniques is done according to the selected TMSIC-based positioning, and to the selected APAT.

5.6 Performance Assessment Procedure and Simulation Results

5.6.1 Performance Assessment

In the previous section, the global PA strategy was detailed to determine the throughput associated to a given user combination. As already explained, even when the users positions are fixed and the UAV position has been found, c_j cannot be determined in advance before placing the UAV and measuring the obtained A2G links. Due to the random nature of LoS/NLoS links, any combination can occur and a fair comparison in the simulation results can only be made when the throughput associated to the UAV position is averaged

over all possible combinations. Section 5.5 presented the PA steps followed at the level of BS a_1 in real time, whereas this section presents the followed procedure to simulate and assess the performance of each UPT-APAT couple. Let $\mathcal{R}$ be the rate vector associated to every combination c_i; the expected achieved rate for the determined UAV position is given by:

$$\overline{R} = \sum_{i=1}^{8} \mathcal{R}(i) Pr(c_i/\boldsymbol{POS}) \tag{5.18}$$

To estimate $\mathcal{R}$, the procedure followed in section 5.5 (sections 5.5.1 and 5.5.2 successively) is iterated for every channel combination. By doing so, the TMSIC testing procedure (Fig. 5.3) is undergone for every c_i, and the probability $p_i(\boldsymbol{POS})$ of having TMSIC (or not) knowing c_i and $\boldsymbol{POS}$ is determined. Thus, the exact TMSIC probability is retrieved from (5.9).

5.6.2 Simulation Results

To evaluate the performance of the presented UPTs and APATs, 1000 simulations were conducted with different user positionings according to Fig. 5.1. The outer cell radius of each hexagonal cell is R_d = 500 m. User 3 region has a maximum width of 60 m along the x axis. Users are assumed to have low mobility, they are independently positioned, their positions being randomly generated with a uniform probability distribution over their respective regions. The transmission channel model between the fixed BS and the users includes a distance-dependent path loss of decay factor 3.76, and a zero-mean lognormal shadowing with an 8 dB variance. The working frequency is 2 GHz, and the parameters of the A2G model are $\alpha = 9.61$, $\beta = 0.16$, $\eta_{LoS} = 1$ dB and $\eta_{NLoS} = 19$ dB, corresponding to an urban environment [130]. The search region for UAV positioning is a rectangular box delimited along the x axis by the cell diameters at the edges of regions 1 and 2 respectively, with the UAV height varying between 50 m and 100 m above the ground. The considered subband bandwidth is $\boldsymbol{B}$ = 156.25 kHz (equivalent to a total bandwidth of 10 MHz subdivided into 64 subbands). The power spectral density of the additive background white noise is $N_0 = -174$ dBm/Hz, and the noise power in a subband is $\sigma^2 = N_0 \boldsymbol{B}$. The power limit constraint over the fixed BS (a_1) is varied between 0.5 W and 5 W, and the MBS power limit assigned to the user cluster is 0.5 W. MATLAB software is used to generate the numerical results and *fmincon* from the optimization toolbox is used to solve the optimization problems in each proposed technique.

The TMSIC probability of the UAV positioning techniques is independent of the used APAT, hence the methods presented in Fig. 5.4 are named after the UPT. The Lower Bound (LB) curves of MPP, MRP and MPRP represent the probability of achieving TMSIC through c_{i^*}, $Pr(c_{i^*}/\boldsymbol{POS})$. The exact probability curves add to the LBs the probability of other combinations that enable TMSIC when the UAV is in $\boldsymbol{POS}$. As expected, MPP-LB delivers the best TMSIC probability between the three methods with 89.9% TMSIC success rate, with MPRP coming second with 88%, and MRP is last with 6.9% for P_{L_1} = 5 W. This important deficit in probability of MRP compared to the two other methods is explained by the absence of the probability term in its objective function: the UAV position is selected according to the throughput it could provide irrespective of the associated probability. This being said, the probability that truly matters is the exact probability, since it reflects the experienced TMSIC probability. We first point out the remarkable closeness between MPP, MRP and MPRP-exact despite the

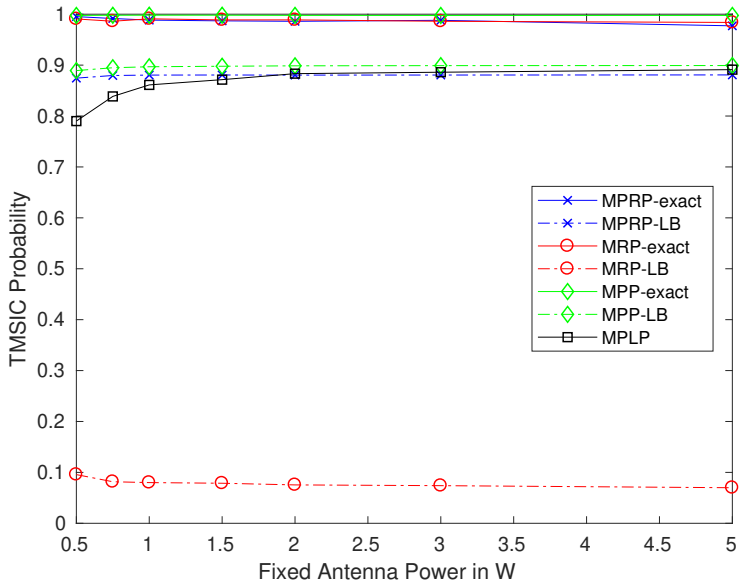


Figure 5.4 – TMSIC probability of the UAV positioning techniques as a function of the fixed antenna power P_{L_1}.

relatively important differences in the lower bounds. While a 10% increase in the TMSIC success rate of MPP and MPRP due to the contribution of the remaining configurations is an intuitive result, it is less evident to explain the substantial increase in probability observed for MRP (from 6.9% to 98%). In fact, the small lower bound probability for MRP translates into a low probability of occurrence of c_{i^*}, then other configurations have higher probabilities of occurrence. If they lead to a TMSIC, their contribution to the total probability will be dominant with respect to c_{i^*}. This was confirmed by a statistical analysis of the number of configurations leading to TMSIC per simulation, which showed that, on average, 7.68 configurations out of the eight yield a TMSIC for MRP. The same analysis can be transposed to MPLP, since it does not account for the TMSIC probability when positioning the MBS (the technique is transparent to the LoS/NLoS combination paradigm). Nevertheless, an average of seven combinations out of the eight enable TMSIC, which explains the relatively high TMSIC probability 89.1%. However, this probability is the lowest among that of all UPTs.

Fig. 5.5 shows a comparison of the system performance in terms of the average SE for all PA and positioning techniques. The achieved SE when the two fixed BSs are available to serve users is added for comparison; DMSIC is used as PA in this case. The performance improvement due to UAV mobility, compared to fixed BSs, is clearly observed for all positioning techniques. Also, the consideration of LoS/NLoS combinations efficiently increases the SE by 3 to 5 bps/Hz for MRP and MPRP compared to MPLP. However, the average MPP performance is lagging behind, as it only surpasses MPLP for small P_{L_1} values before going below for power limit values above 1.5 W. This suggests that the evolution of the UAV position with the growing value of P_{L_1} affects the A2G links in a way that the increase rate of the MPP throughput is lower than that of MPLP. Indeed, an analysis of UAV positioning in MPP and its evolution with the power limit shows that high P_{L_1} values tend to place the UAV at the edges of the search region, resulting in poor channel gains, which explains the lower throughput compared to MPLP

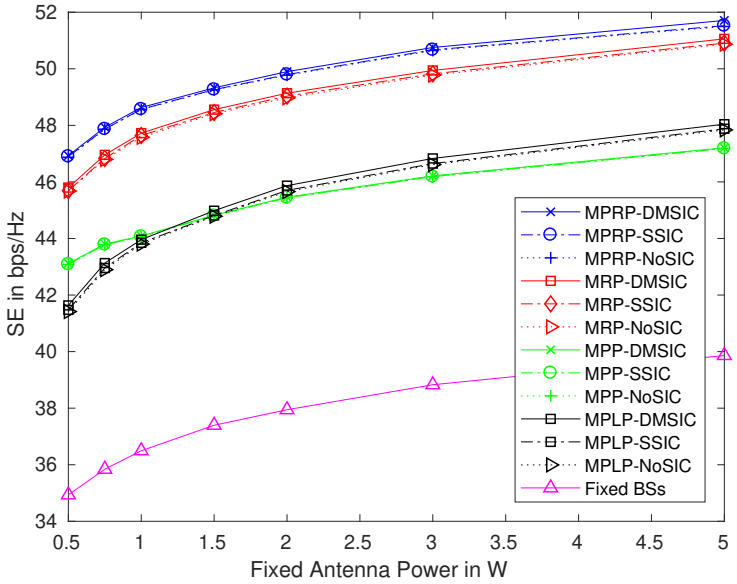


Figure 5.5 – Spectral efficiency of the different UAV positioning techniques and PA strategies.

at $P_{L_1} = 5$ W. More details on the reasons behind this placement, its interaction with the user positioning and the effect it has on user throughput are given later for all the positioning techniques in the analysis of the individual user rates shown in Fig. 5.7.

Nonetheless, we can sum up the results of Fig. 5.5 by stating that focusing exclusively on the TMSIC probability can mislead the UAV placement into areas with poor A2G links and poor achievable throughput. The introduction of the throughput in the objective function provides the qualitative edge for MRP over MPP, since throughput is accounted for during positioning, while the TMSIC probability difference between the two is negligible (cf. Fig. 5.4). This being said, combining the throughput and the probability in MPRP provides even better results since both objectives are accounted for from the start of the positioning process. However, the performance gain of MPRP and MRP comes at the cost of an additional complexity compared to MPLP, since 64 combinations need to be checked for MRP and MPP compared to the eight decoding orders assessed by MPLP.

Regarding the NoSIC, SSIC, and DMSIC APAT variants for every UPT, small performance differences are observed for all techniques. This is due to the fact that, most of the time, TMSIC is successfully applied and non-TMSIC PAs are summoned for only a small proportion of LoS/NLoS combinations not leading to a TMSIC (around 0.3/8 or less for all UPTs when $P_{L_1} = 0.5$ W). Nonetheless, DMSIC is the best APAT in terms of throughput and is therefore used by default from hereinafter. The methods names are selected according to the selected UPT in the following results. In Fig. 5.6, the Jain fairness index [31] is used to assess the fairness of the contribution of each user to the total throughput. The index is upper bounded by 1 for absolute fairness and lower bounded by 1/3 for the worst case scenario. It is first observed that MPLP presents the lowest fairness index with a maximum of 0.84 for $P_{L_1} = 5$ W. The other techniques present much higher fairness indices. This is due to the significantly higher probability of achieving TMSIC which was shown in chapter 4 to provide better throughput through better fairness. The remaining

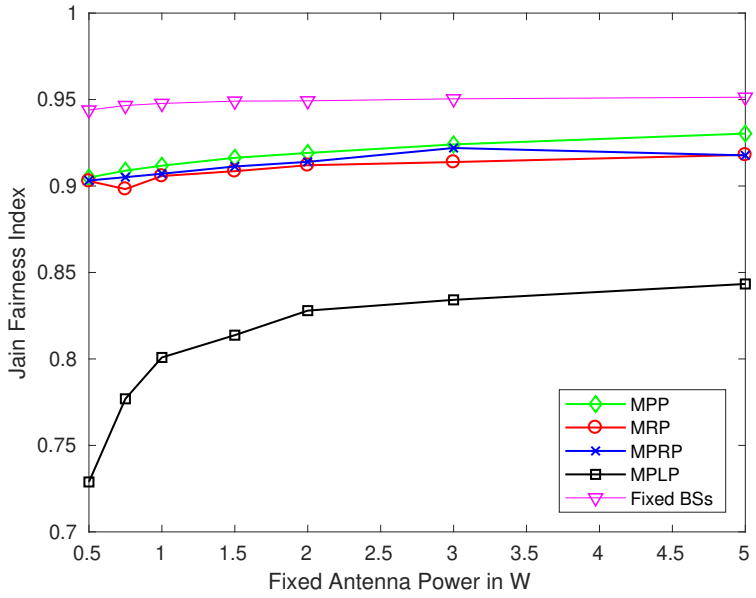


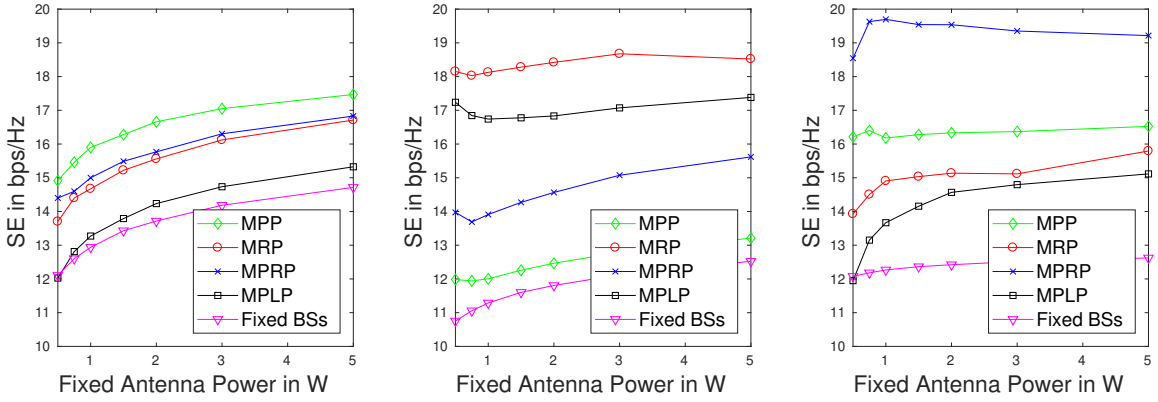
Figure 5.6 – Fairness comparison of the positioning techniques as a function of the fixed antenna power.

UPTs have a quite similar fairness, with the MPP presenting an overall better fairness, especially for high P_{L_1} values. The fixed BSs scenario presents a slightly better fairness compared to MPP, MRP and MPRP. In fact, as in MPP where the UAV placement is pushed back towards the limit of the search region for high power limits, the fixed BSs correspond to a_2 being further away from the user cluster, compared to other UPTs. This translates into a smaller achieved throughput, as shown in Fig. 5.5, but it also leads to a greater fairness due to the symmetry of the user cluster with respect to a_1 and former fixed a_2.

So far, MPP has been shown to provide the best TMSIC probability and fairness from Figs. 5.4 and 5.6, whereas MPRP was shown to yield the highest sum-throughput in Fig. 5.5. Although a trade-off does exist between throughput and fairness, the closeness of the fairness measures and TMSIC probability between MPRP and MPP (0.03 units of difference in the fairness index, and one percentage point difference in probability), compared to the large gap in throughput (around 4.5 bps/Hz, i.e. a 10% difference) does tend to promote MPRP as the best trade-off. However, when having a closer look at the individual user rates for every UPT, other factors come into play which affect the choice of the positioning technique as seen from the results of Fig. 5.7.

In Fig. 5.7, we present the individual throughput for every user category, for all positioning techniques. The separate contribution of each user in the cluster throughput is analyzed for each UPT. Starting with the two fixed BSs, we can observe that the influence on throughput of the growing power limit is more pronounced for user 1 than for user 3, and for user 3 more than for user 2. The closer the user to a_1 on average, the more it benefits from the additional power of a_1. However, user 3 globally presents the lowest user throughput in the cluster, because of its geographical position on the cells edges.

To analyze the performance of positioning techniques, we must first discuss the effect



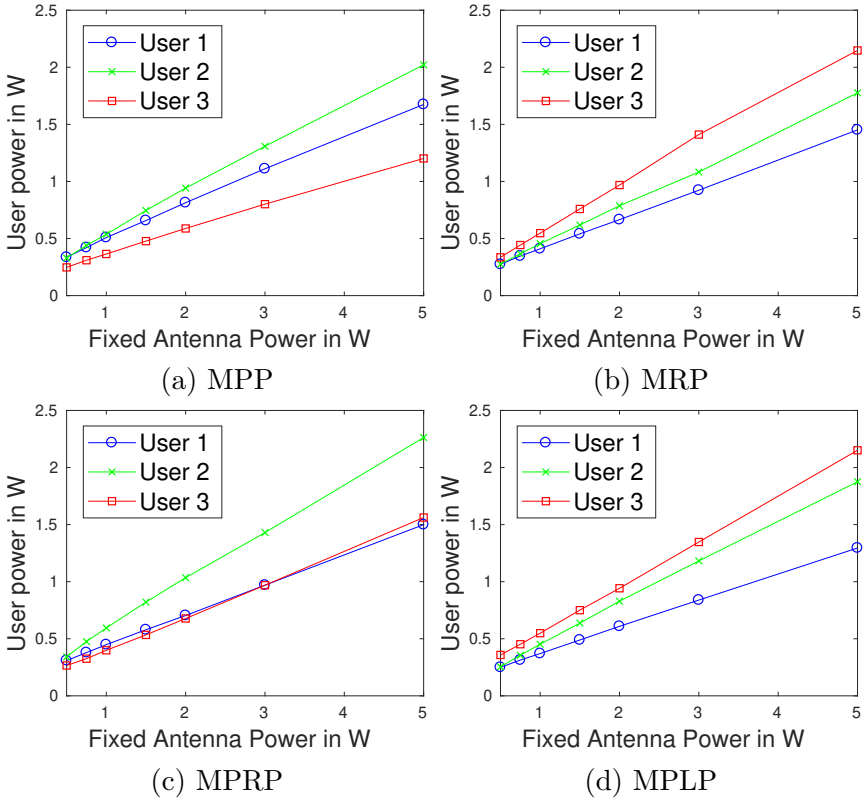
(a) User 1 (b) User 3 (c) User 2

Figure 5.7 – Throughput distribution over the three-user NOMA cluster.

of the UAV position on the channel gains as well as how the objective functions affect this position. We first focus on TMSIC probability as the objective function. According to (5.3), the LoS link with user k has its chances maximized when θ_k tends to 90° (the UAV is on top of k), and the NLoS link is favored when θ_k tends to 0° (the UAV and k are far apart on the xy plane). If the users are (very) close to one another, then placing the UAV on top of the three of them leads to the largest c_8 probability (the all-LoS case). If not, the c_is achieving the best probability are when the UAV is placed almost at the top of one user, establishing a LoS link with that user and favoring NLoS links with the other two. In that scenario, user 3 is the least likely to have the UAV on top of it: being a cell-edge user, the distance separating it from the other two users (which would be the distance separating them from the UAV in the xy plane) is rather small compared to the distance that separates user 1 from user 2 if the UAV was placed on top of one of these two. This smaller distance reduces the chances of NLoS with users 1 and 2 when the UAV is on top of user 3, that is why c_5 and c_3 are favored (i.e. either user 1 or user 2 being in LoS). This explains why the rate of user 3 in MPP is below those of user 1 and 2, with an average rate difference of 4 bps/Hz. Also, if the users are far enough from one of the corners of the search region, the all-NLoS combination (c_1) becomes the most probable combination, under the condition of a possible TMSIC for the UAV position at this corner. This is aided by the growing power limit which enables more locations to achieve TMSIC. However, placing the UAV at the corners of the search region with higher powers induces poorer channel gains due to the free space path loss and to the high NLoS probability, which explains the behavior of MPP in Fig. 5.5.

When the throughput is considered in the objective function, a significant advantage is given for user 3 over users 1 and 2 because of its location in between the two cell-center users. When only the throughput is considered (as in MRP and MPLP), LoS dominant combinations are favored due to their better channel gains yielding a higher throughput. However, for the resulting position, the combination which yielded the UAV position is rarely the most favorable one (as discussed previously for Fig. 5.4) and the actual combination contributing the most to the TMSIC probability is c_2. In other terms, the UAV ends up in between the three users, favoring thereby a LoS link only with user 3, enhancing its rate as shown in Fig. 5.7 for both MPLP and MRP. Regarding MPRP, the fact that it takes into account both throughput and probability enabled it to deliver the

best solutions from the average throughput perspective. Such solutions usually reside in placing the UAV relatively close to user 2 (by favoring c_3) so that the system throughput is maximized. Obviously, doing so profits most to user 2: its average rate is around 19 bps/Hz when user 1 and user 3 rates vary between 14 and 16 bps/Hz.



(a) MPP (b) MRP

(c) MPRP (d) MPLP

Figure 5.8 – User power allocation according to the selected UPT.

The presented results from Fig. 5.7 can also be looked at from the perspective of the average power allocated to each user for every UPT. Figs. 5.8b and 5.8d show that rate-focused techniques like MRP and MPLP, which tend to place the UAV over user 3 (favoring c_2), end up loading user 3 with the highest power level, translating into a higher throughput of user 3 compared to users 1 and 2. On the other hand, it is also clear from Figs. 5.8a and 5.8c for MPP and MPRP that the power allocated to users 1 and 2 is more important than for user 3. As mentioned previously for Fig. 5.7 regarding these methods, the UAV placement favors UAV locations over user 1 (c_5) and user 2 (c_3), leading to a higher achieved throughput for users 1 and 2 compared to user 3. To go even further, since MPRP favors c_3 exclusively, a greater gap is observed between the powers of user 1 and user 2 in MPRP compared to MPP. In fact, combining the analyses of Figs. 5.7, 5.8a and 5.8c, we can say that MPP delivers similar rate and power allocations to users 1 and 2 with user 3 lagging behind, whereas MPRP delivers similar rate and power allocations to users 1 and 3 with user 2 ahead of both users.

This great diversity in the performance results at the level of every different user provides a broad selection choice depending on system priorities. If cell-edge user's performance is prioritized (and thereby cell-edge user groups) over the total system throughput, going with MRP is the most suitable choice. On the other hand, if cell-center user's performance is the priority, then MPRP and MPP can be employed in such cases, while keeping

in mind that MPRP delivers the best overall throughput performance. Finally, MPLP can be also used to favor the cell-edge user, while maintaining a good global throughput and also reducing the optimization complexity compared to MRP due to the simpler mean path loss model. This wide panel of selection also provides the network planner with a multitude of answers to face the variations in time of the users traffic requirements, where the user priorities could change and therefore the UPT strategy can change accordingly.

5.7 Conclusion

In this chapter, we addressed the problem of UAV placement for supporting an overloaded BS in a two-cell NOMA CoMP system. The UAV positioning seeks the application of TMSIC which provides great fairness and throughput performance. The proposed approach considers the LoS/NLoS channel combinations of users, instead of using the mean path loss, which proved its efficiency in both TMSIC probability and system throughput. Exclusive attention to TMSIC probability over system throughput showed its shortcomings regarding high power limit values, whereas the combination of probability and throughput information best captures the features of the problem and delivers the best performance results. The presented techniques have a great diversity and can be selected at will according to which group of users is prioritized (cell-edge vs. cell-center) with negligible compromise on system performance.

In the last chapter of this thesis, we turn our attention towards the context of D2D communications, while still seeking the applicability of mutual SIC NOMA. Enabling devices in proximity to communicate in a peer-to-peer fashion is expected to offload the surging demand in throughput from the network backbone, decentralizing it over the network front-end. That being said, the anticipated leaps in capacity will require the use of multi-factorial solutions. Therefore, we will be looking to combine our proposed NOMA techniques to the D2D scenario while also resorting to full-duplex communications.

The contributions of this chapter led to the publication of the following journal paper:

A. Kilzi, J. Farah, C. Abdel Nour and C. Douillard, "Analysis of Drone Placement Strategies for Complete Interference Cancellation in Two-Cell NOMA CoMP Systems," in *IEEE Access*, vol. 8, pp. 179055-179069, Sept. 2020.

Chapter 6

NOMA Mutual SIC for Full-Duplex D2D Systems Underlaying Cellular Networks

As the number of connected devices keeps on growing, paradigms shifts need to be undertaken to keep up with the explosive demand. D2D communication is one such solution which can increase the number of connections, reduce latency, and offload traffic from MNOs without requiring any additional network infrastructures. That is why it has received a growing interest from both academia and industry in the last couple of years [32–36]. In this chapter, we propose to study the interplay of NOMA mutual SIC with the D2D ecosystem to further improve system performance. Assuming a pre-established cellular network, the aim will be to operate the D2D-Cellular User (CU) pairing and power control such that the sum-throughput of the D2D underlay system is maximized without affecting the QoS of CUs.

We first start by presenting state-of-the-art research on inband underlay D2D with NOMA (section 6.1). Then, the system model is presented in section 6.2 and the joint channel and power allocation problem is formulated, where it is shown that the resource allocation problem could be separated into disjoint PA and channel assignment problems. The PA problems of Full Duplex (FD) and Half Duplex (HD) without SIC (FD-NoSIC, HD-NoSIC) are solved in section 6.3, while the PA problem with SIC is reformulated for HD and FD (HD-SIC, FD-SIC) in section 6.4. Mutual SIC PA is solved for the case of HD transmission in section 6.5. In sections 6.6, 6.7, and 6.8 the conditions of mutual SIC for FD-D2D are derived, the problem constraint reduction is performed, and then a geometrical resolution is proposed, allowing for a cost-effective resolution of the FD-SIC PA problem. The channel allocation procedure is discussed in section 6.9. Simulation results are presented in section 6.10, and conclusions are drawn in section 6.11.

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

- We derive the PMCs and SIC conditions allowing for mutual interference cancellation between D2D and CU users.
- We show that PMCs imply the SIC conditions for both HD and FD transmission modes, which greatly reduces the PA problem complexity for the case of FD-SIC.
- We solve analytically the PA problem for all transmission strategies, especially for the case of FD-SIC where an efficient procedure is provided to optimally solve the

D2D rate maximization problem with constant time complexity.

- The complementarity between D2D and mutual SIC NOMA is highlighted. It is discussed how NOMA integration can extend the applicability of D2D to broader user configurations and channel scenarios.

6.1 Related Works

Recently, considerable attention was directed at the combination of NOMA with D2D communications in underlay mode. The study in [135] considers resource block assignment and PA in a downlink NOMA system with D2D. HD is used in the D2D pairs, and CUs are grouped in NOMA clusters. The influence of the HD-D2D users over the SIC decoding orders of CUs is accounted for in both the block assignment and the PA phase, because the interference they generate may change the decoding order. However, NOMA SIC is not used to decode the interfering signals of the collocated D2D pairs. The same is true for [136], but additional power constraints are introduced on the D2D pairs to maintain the same SIC decoding orders at CUs as for the case of D2D-disabled systems. The work in [137] introduces the concept of D2D group, where a D2D transmitter communicates with multiple D2D receivers via NOMA. To maximize the network sum-throughput, sub-channel allocation is conducted using many-to-one matching for CU-D2D grouping, and optimal PA is approximated iteratively via successive convex approximation. When limiting the number of multiplexed D2Ds to one per CU user, the work in [138] provides a joint D2D-CU grouping and PA strategy for energy efficiency maximization: the Kuhn-Munkres technique is applied successively for channel allocation, while optimal PA is obtained using the Karush-Kuhn-Tucker conditions. In all the preceding studies, NOMA is applied either between the CU users [136], or between users of the same D2D group [137, 138], but the interference cancellation of the D2D signals at the level of CU users (and inversely) is not considered. At most, attention is given towards managing the SIC decoding order at the level of the CUs in [135, 136], or at the level of the D2D receivers in [137, 138].

The work in [139] tackles the problem of HD-D2D throughput maximization in an uplink system where NOMA is used between D2D and CU users. If the D2D causes strong interference on the BS, its signal can be decoded and then subtracted before retrieving the CU signal. However, FD-D2D is not studied and SIC occurs only at the level of the BS, i.e. D2D devices suffer from CU interference. Besides, the information-theoretic conditions for SIC feasibility are not considered in the study. In [140], an efficient graph-based scheme is proposed to maximize the D2D sum-rate of an uplink system. To that end, an interlay mode is introduced to HD-D2D communication where a D2D pair can join a NOMA group to remove the interference between it and the cellular NOMA users. However, the conditions for applying SIC - and thus for determining the SIC decoding order - are only conditioned by the ascending order of channel gains between the senders and the receivers. In other words, the interfering signals that can be canceled are the ones that are attributed channel gains better than that of the useful signal, regardless of their power level at reception. This may lead to outage probabilities of one if no PA measures are taken to guarantee SIC stability as shown in [105]. The work in [141] incorporates NOMA into D2D cellular networks to maximize system connectivity. Unlike [140], the D2D NOMA-aided

modes are defined according to the SIC orders at the level of the D2D and the BS. The SIC decoding orders are governed by the strong interfering signal which is bound to the channel conditions as well as the used PA. The optimal PA and mode selection are solved in the presence of decoding SINR threshold constraints, then the user pairing problem is turned into a min-cost max-flow problem which is solved by the Ford–Fulkerson algorithm. However, the channel and power conditions enabling the SIC procedure are not developed beyond the SINR conditions, and furthermore, the case of FD-D2D NOMA-aided network was not addressed neither in this study, nor in the entire literature combining NOMA and D2D. Therefore, we study in this chapter the combination of NOMA with D2D systems in general, and FD-D2D systems in particular. Also, differently from previous works, great attention is directed towards deriving necessary and sufficient channel and power conditions enabling mutual SIC application between D2D and CUs.

6.2 System Model

In this work, we consider the integration of a D2D underlay system into a pre-established cellular network. The base network consists of K CU users transmitting over their assigned UpLink (UL) channels separately, with a maximum of one channel per CU user. The system bandwidth is divided into $N \geq K$ channels and the D2D system is constituted of D D2D pairs ($D \leq K$) exchanging data over a subset of the K UL cellular channels, with a single D2D pair per cellular channel. The pairs can exchange data either in HD or FD mode, while the CUs are always in HD. The objective of the study is to perform optimal D2D channel allocation and power control, such that the obtained D2D-CU pairs yield maximum D2D sum-throughput while guaranteeing the required rates of the collocated CU users. To that end, let $\mathbb{C}$ denote the set of CUs, $\mathbb{C} = \{u_1, u_2, \ldots, u_K\}$, and $\mathbb{D}$ the set of D2D pairs, $\mathbb{D} = \{(d_{1,1}, d_{1,2}), (d_{2,1}, d_{2,2}), \ldots, (d_{D,1}, d_{D,2})\}$. A schematic of the network is presented in Fig. 6.1, where $d_{n,1}$ and $d_{n,2}$ form the n^{th} pair transmitting in FD mode.

The interference channel power gains between a CU u_i, on the one hand, and $d_{n,1}$ and $d_{n,2}$ on the other hand, are denoted by $h_{d_{n,1},u_i}$ and $h_{d_{n,2},u_i}$ respectively. The direct link between CU u_i and BS b has a squared channel gain denoted by h_{b,u_i}. The message m_{u_i}, transmitted by u_i with power P_{u_i}, reaches the BS with a power level $P_{u_i} h_{b,u_i}$, and causes an interference level of $P_{u_i} h_{d_{n,1},u_i}$ and $P_{u_i} h_{d_{n,2},u_i}$ at $d_{n,1}$ and $d_{n,2}$ respectively. Each device $d_{n,j}$ ($j \in \{1, 2\}$) of the n^{th} D2D pair can transmit a message $m_{n,j}$ with power $P_{n,j}$ to the other D2D user and suffers from both the interference of user u_i and its residual self interference power $\eta_{n,j} P_{n,j}$, with $\eta_{n,j}$ denoting its Self Interference (SI) cancellation capability. The D2D inter-user channel gain is denoted by h_{d_n} and the interference channel gains from $d_{n,1}$ and $d_{n,2}$ to the BS are denoted by $h_{b,d_{n,1}}$ and $h_{b,d_{n,2}}$ respectively. In this study, a frequency-non-selective channel is assumed, so that the channel gains are independent from the sub-band frequency and account only for large scale fading including path loss and shadowing.

In this work, it is assumed that, prior to resource allocation and data exchange, a D2D discovery phase [14, 32, 142] takes place in the system, during which the D2D devices inform the BS about their desire to initiate a D2D link, and forward to the BS their estimates of the D2D-CU links ($h_{d_{n,1},u_i}, h_{d_{n,2},u_i}$), as well as the D2D links (h_{d_n}). Therefore, the BS is assumed to have perfect knowledge of the long-term evolution of the different channel gains, through signaling exchange between the different entities. The BS then performs resource allocation based on these estimated channel gains to optimally pair

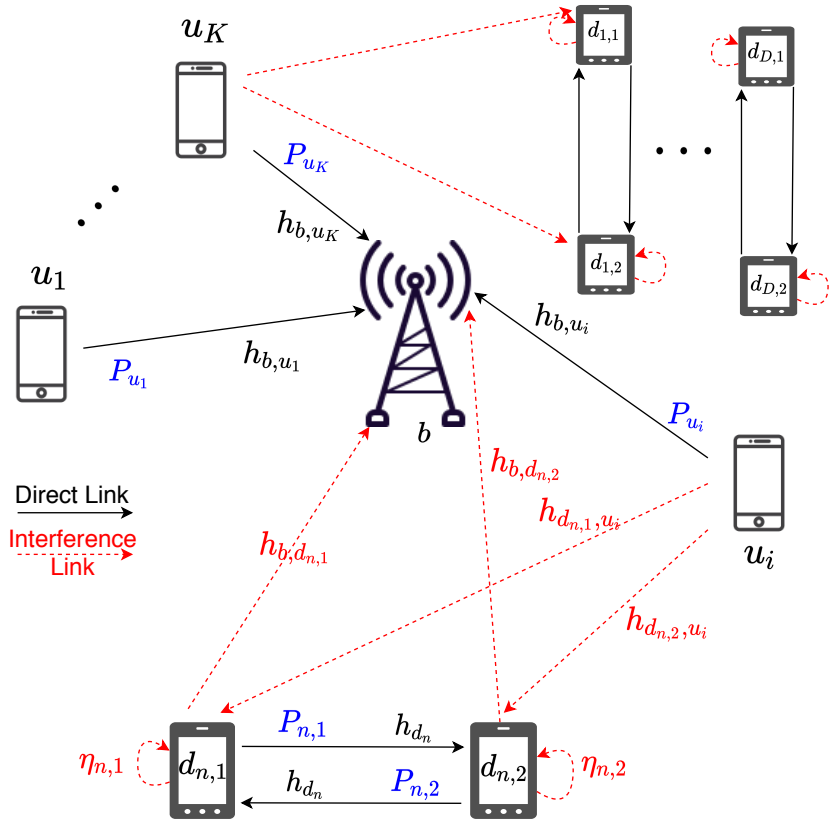


Figure 6.1 – FD-D2D system with D pairs underlaying a cellular network with K CUs.

the D2Ds to CUs and to instruct D2D-CU pairs of the required transmit powers on their collocated channels, according to the selected transmission scenario.

6.2.1 Formulation of the Joint Channel and Power Allocation Problem

Let O be the channel allocation matrix, with element $o(n,i)$ at the n^{th} row and i^{th} column equaling one if D2D pair n is collocated with CU u_i and zero otherwise. Also, let $\mathcal{R}_{D2D}(n,i)$ be the maximum achievable D2D rate of pair n when collocated with u_i. Channel allocation is performed such that a D2D pair is multiplexed over a single UL channel, on the one hand, and that a maximum of one D2D pair is multiplexed over a UL channel, on the other hand. The joint channel and power allocation problem for the maximization of the total D2D throughput can be cast as:

$$\max_{\{O,P_{n,1},P_{n,2},P_{u_i}\}} \left(\sum_{i=1}^{K} \sum_{n=1}^{D} o(n,i) \times \mathcal{R}_{D2D}(n,i) \right)$$
$$\text{s.t.} \quad \sum_{i=1}^{K} o(n,i) = 1, \forall n \in \{1,\ldots,D\}, \sum_{n=1}^{D} o(n,i) \leq 1, \forall i \in \{1,\ldots,K\}, \tag{6.1}$$

where $\mathcal{R}_{D2D}(n,i)$ is the solution to:

$$\max_{\{P_{n,1},P_{n,2},P_{u_i}\}} R_{D2D}(n,i), \tag{6.2}$$

such that:

$$R_{u_i} \geq R_{u_i,min}, \tag{6.2a}$$
$$P_{n,1} \leq P_{n,1,M}, \tag{6.2b}$$
$$P_{u_i} \leq P_{u_i,M}, \tag{6.2c}$$
$$P_{n,2} \leq P_{n,2,M}. \tag{6.2d}$$

$P_{u_i,M}, P_{n,1,M}, P_{n,2,M}$ are the maximum transmit powers of $u_i, d_{n,1}$ and $d_{n,2}$ respectively, $R_{u_i,min}$ is the minimum target rate of u_i, R_{u_i} its achieved rate, and $R_{D2D}(n,i)$ the D2D rate, i.e. the sum of the rates achieved by $d_{n,1}$ ($R_{d_{n,1}}$) and $d_{n,2}$ ($R_{d_{n,2}}$).
From the structure of Problem (6.1), and since CUs are allocated orthogonal channels, the performance of a given D2D-CU pair is independent from the network activity over the remaining channels in the system. Therefore, one can optimize the throughput of all possible D2D-CU pairs, constructing a $D \times K$ table of achievable rates, and then proceed to the optimal channel allocation phase which assigns the D2D-CU pairs based on their achievable rate, aiming to maximize the D2D sum-throughput in the system. The aim of the following sections is to obtain the optimal PAs of the four transmission methods FD-NoSIC, HD-NoSIC, HD-SIC and FD-SIC in order to build their corresponding tables of achievable rates $\mathcal{R}_{D2D}^{FD-NoSIC}, \mathcal{R}_{D2D}^{HD-NoSIC}, \mathcal{R}_{D2D}^{HD-SIC}$, and $\mathcal{R}_{D2D}^{FD-SIC}$ respectively. Based on these tables, optimal channel allocation is conducted in section 6.9.

6.3 Power Allocation for No-SIC Scenarios

From hereinafter, since the optimal D2D rate of all (n,i) couples is to be computed and because the resolution of the PAs is independent of the elected D2D-CU couple, we drop the indices relative to a specific D2D pair and CU. Hence, user u designates the CU at hand, and d_1 and d_2 are the corresponding D2D pair. The involved channels gains are therefore denoted by h_d, h_{b,d_1}, h_{b,d_2}, $h_{d_1,u}$, $h_{d_2,u}$ and $h_{b,u}$, and the transmit powers of d_1, d_2 and u are P_1, P_2, P_u, with their power limits $P_{1,M}, P_{2,M}, P_{u,M}$.

6.3.1 FD-NoSIC

In FD, d_1 and d_2 transmit simultaneously, thus they both suffer from Residual SI (RSI). Since, in this method, SIC is not attempted at the levels of d_1, d_2 and the BS, the SINRs at the level of the BS and the D2D users are given by:

$$SINR_b = \frac{P_u h_{b,u}}{P_1 h_{b,d_1} + P_2 h_{b,d_2} + \sigma^2},$$
$$SINR_{d_1} = \frac{P_2 h_d}{P_u h_{d_1,u} + \eta_1 P_1 + \sigma^2}, \quad SINR_{d_2} = \frac{P_1 h_d}{P_u h_{d_2,u} + \eta_2 P_2 + \sigma^2}, \tag{6.3}$$

with σ^2 being the additive Gaussian noise power. The achieved rates are expressed according to the Shannon capacity theorem:

$$R_u = B \log_2(1 + SINR_b), \tag{6.4}$$
$$R_{d_1} = B \log_2(1 + SINR_{d_1}), \quad R_{d_2} = B \log_2(1 + SINR_{d_2}), \tag{6.5}$$

with B the bandwidth of each UL channel resource. Due to the interference terms in (6.3), Problem (6.2) is non-convex. To solve it, a geometrical representation can be used, leading to the analytical global solution in [143]. This method is adopted in our work to derive the results of the FD-NoSIC scenario in the performance assessment section.

6.3.2 HD-NoSIC

The time slot is now divided into two equal half-time slots where d_1 and d_2 alternately transmit and receive information. To maximize the total D2D rate, the optimization is conducted separately in the two half-time slots. In the first half, d_1 transmits information ($P_2 = 0$). In Problem (6.2), the objective function and CU rate are now:

$$R_{D2D,1} = R_{d_2} = B\log_2(1 + \frac{P_1 h_d}{P_{u,1} h_{d_2,u} + \sigma^2}),$$
$$R_{u,1} = B\log_2(1 + \frac{P_{u,1} h_{b,u}}{P_1 h_{b,d_1} + \sigma^2}).$$

Also, Problem (6.2) is constrained only by eqs. (6.2a) to (6.2c). Note that $P_{u,1}$ is the transmit power of u during the first half-time slot. $R_{D2D,1}$ is strictly increasing with P_1 and decreasing with $P_{u,1}$; therefore, to maximize $R_{D2D,1} = R_{d_2}$, P_1 should be increased and $P_{u,1}$ decreased as long as $R_{u,1}$ satisfies the minimum rate condition of the CU. Consequently, P_1 should be increased as much as possible and then $P_{u,1}$ is obtained as a function of P_1 ($P_{u,1} = f(P_1)$) by enforcing an equality between $R_{u,1}$ and $R_{u,min}$. If, for $P_1 = P_{1,M}$, $f(P_{1,M}) \leq P_{u,M}$, couple $(P_{1,M}, f(P_{1,M}))$ is retained as the (P_1, $P_{u,1}$) solution; otherwise, couple $(f^{-1}(P_{u,M}), P_{u,M})$ delivers the best solution. This is summarized as follows:

$$P_1^* = \min\{P_{1,M}, f^{-1}(P_{u,M})\}, \quad P_{u,1}^* = f(P_1^*),$$

where $f^{-1}(P_{u,M})$ is given by:

$$f^{-1}(P_{u,M}) = \frac{1}{h_{b,d_1}}\left[\frac{P_{u,M} h_{b,u}}{2^{\frac{R_{u,min}}{B}} - 1} - \sigma^2\right].$$

The same reasoning is applied for the second half-time slot (where $P_1 = 0$) to maximize $R_{D2D,2} = R_{d_1}$. The total user u and D2D rates are given by:

$$R_u = \frac{1}{2}R_{u,1} + \frac{1}{2}R_{u,2}, \quad R_{D2D} = \frac{1}{2}R_{d_1} + \frac{1}{2}R_{d_2}.$$

6.4 Power Allocation Problem Modification for HD and FD with Mutual SIC (HD-SIC and FD-SIC)

Using a SIC receiver at the level of the BS and the D2D users, interfering messages can be decoded and then subtracted from the received message, canceling thereby the interference in both FD and HD scenarios. Let m_1 and m_2 be the messages transmitted by devices d_1 and d_2. In the case of FD, the BS can decode and subtract successively m_1 then m_2, or m_2 then m_1, before proceeding to the decoding of m_u (the message transmitted by the CU);

hence, two decoding orders are possible. Users d_1 and d_2 can also remove the interference of u, leading to the following SINR expressions:

$$SINR_{d_1} = \frac{P_2 h_d}{\eta_1 P_1 + \sigma^2},$$
$$SINR_{d_2} = \frac{P_1 h_d}{\eta_2 P_2 + \sigma^2},$$
$$SINR_b = \frac{P_u h_{b,u}}{\sigma^2}.$$

The SINR expressions are replaced in (6.4) and (6.5) to obtain R_u and $R_{D2D} = R_{d_1} + R_{d_2}$ that will be used in Problem (6.2). For the case of HD, the SINRs in the first half-time slot are:

$$SINR_{d_2} = \frac{P_1 h_d}{\sigma^2}, \qquad SINR_b = \frac{P_u h_{b,u}}{\sigma^2}.$$

In the second half-time slot, $SINR_b$ is the same as in the first half-time slot, and $SINR_{d_1}$ is given by $P_2 h_d/\sigma^2$. Problem (6.2) is now reformulated in each time slot by expressing the rates using the present SINRs. However, additional constraints relative to the SIC feasibility must be added to the problem. In the following sections, PMCs and SIC conditions are derived and then Problem (6.2) is solved for HD-SIC and FD-SIC successively.

6.5 Power Allocation for HD-SIC scenario

Consider the first half-time slot, where u and d_1 are transmitting and b and d_2 are receiving. Hereafter, we develop the mutual SIC constraints between b and d_2 (as a receiver). Let $SINR_i^{m_j}$ be the SINR of message m_j at the level of user i (i is either d_1, d_2 or b, and j is either 1, 2 or u). For b to successfully decode the message m_1 transmitted by d_1 to d_2, the received rate of m_1 at the level of b must be greater than the rate of m_1 at the level of d_2. Thus, we must have: $SINR_b^{m_1} > SINR_{d_2}^{m_1}$. Similarly, the rate condition for the decoding of m_u at the level of d_2 is derived from the condition $SINR_{d_2}^{m_u} > SINR_b^{m_u}$. This situation is equivalent to the case of two different RRHs transmitting both messages to two separate receivers, which was studied in chapter 2. The SINR conditions lead to:

$$h_{b,d_1} h_{d_2,u} > h_d h_{b,u}. \tag{6.6}$$

In addition to condition (6.6), the PMCs must be verified, in order to ensure that the message to be decoded first at the level of a receiver has a higher power level than that of the remaining message. The PMCs for the decoding of m_u and m_1 at the level of d_2 and b are given by:

$$\left.\begin{array}{l} P_{u,1} h_{d_2,u} > P_1 h_d \\ P_1 h_{b,d_1} > P_{u,1} h_{b,u} \end{array}\right\} \Rightarrow A = \frac{h_d}{h_{d_2,u}} < \frac{P_{u,1}}{P_1} < \frac{h_{b,d_1}}{h_{b,u}} = B. \tag{6.7}$$

Note that (6.6) is satisfied if (6.7) is satisfied, since (6.6) is equivalent to $A < B$. Therefore, the PMCs encompass the rate conditions while being more restrictive. Problem (6.2) now only includes the additional constraint (6.7) for the first time slot. The HD-SIC rate

expressions are as follows:

$$R_{D2D,1} = R_{d_2} = B\log_2(1 + \frac{P_1 h_d}{\sigma^2}),$$
$$R_{u,1} = B\log_2(1 + \frac{P_{u,1} h_{b,u}}{\sigma^2}).$$

Maximizing R_{d_2} lies in the increase of P_1. Also, guaranteeing the CU rate $R_{u,min}$ can be achieved by setting $P_{u,1}$ to $P_{u,m} = (2^{\frac{R_{u,min}}{B}} - 1)\sigma^2/h_{b,u}$. However, due to the PMCs, the increase in P_1 is very likely to increase $P_{u,1}$ according to the range of allowed values in (6.7), leading to an excess of CU rate. Since maximization of network throughput (i.e. sum of D2D and CU rates) is not our objective, we select from the range of admissible $R_{u,1}$ values, the one closest to $R_{u,min}$. With that criterion in mind, the PA problem for D2D rate maximization is solved by increasing P_1 as much as possible (possibly until $P_{1,M}$) and adjusting $P_{u,1}$ accordingly. The proposed PA procedure, illustrated in Fig. 6.2,

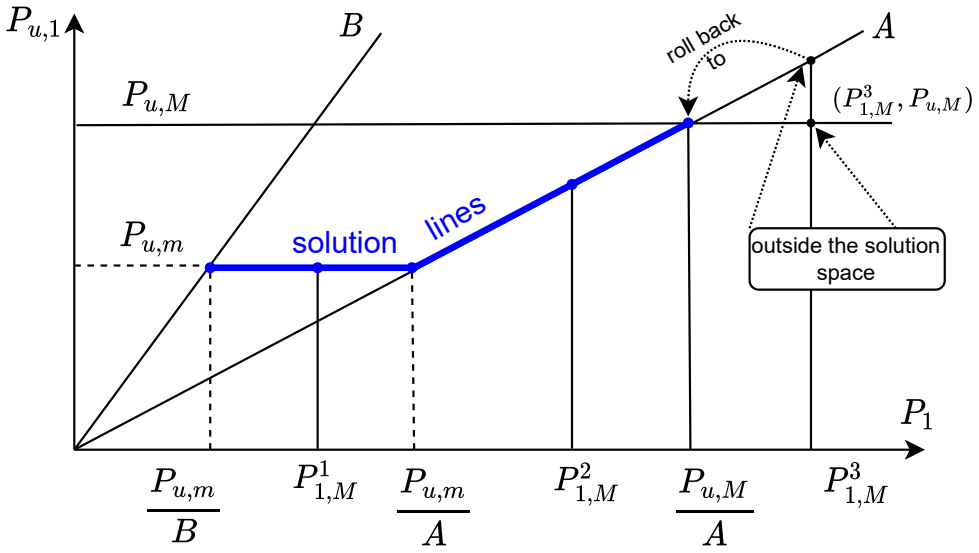


Figure 6.2 – Schematic of the solution space to the HD-SIC PA problem, for different $P_{1,M}$ values.

operates as follows: if $P_{1,M} < P_{u,m}/A$, keep the couple $(P_1 = P_{1,M}, P_{u,1} = P_{u,m})$ as the optimal solution. This case is represented by the example $P^1_{1,M}$ on the horizontal blue line in Fig. 6.2. If this is not the case, check if $AP_{1,M} > P_{u,M}$. If yes (cf. example $P^3_{1,M}$ in Fig. 6.2), the solution is $(P_{u,M}/A, P_{u,M})$; if not (cf. example $P^2_{1,M}$), the solution is $(P_{1,M}, AP_{1,M})$. Restricting the solution space to the blue lines in Fig. 6.2 guarantees that the CU always transmits at the minimum necessary power that respects the problem constraints. Note that if $P_{1,M}$ is too low ($< P_{u,m}/B$), the problem is not feasible even when (6.6) is verified.

For the second time slot, the same methodology is followed, where the PMCs and the new necessary and sufficient channel conditions are given by:

$$h_{d_1,u} h_{b,d_2} > h_{b,u} h_d, \tag{6.8}$$
$$A' = \frac{h_d}{h_{d_1,u}} < \frac{P_{u,2}}{P_2} < \frac{h_{b,d_2}}{h_{b,u}} = B'. \tag{6.9}$$

As a conclusion, in the HD-SIC scenario, the system checks for the validity of the channel condition corresponding to the half-time slot before going through the procedure described above. If the channel condition is not favorable or if no solution exists (i.e. $P_{u,m} > P_{u,M}$ or $P_{1,M} < P_{u,m}/B$ for the first half, and $P_{2,M} < P_{u,m}/B^{'}$ for the second half), the system reverts to the HD-NoSIC solution of section 6.3.2. This leads to four possible combinations of SIC/NoSIC procedures, two for every half-time slot, and they are all included in the HD-SIC algorithm.

6.6 Derivation of the SIC conditions for FD mutual SIC

In this scenario, we are looking for the conditions that allow d_1 to decode m_u, d_2 to decode m_u, and b to decode m_1 and m_2. As already mentioned, two decoding orders are possible at the level of b.

6.6.1 First decoding order: b decodes m_2 then m_1

We first start by studying the mutual SIC constraints between b and d_1 (as a receiver). For b to successfully decode message m_2 transmitted by d_2 to d_1, we must have:

$$SINR_b^{m_2} > SINR_{d_1}^{m_2},$$
$$\frac{P_2 h_{b,d_2}}{\sigma^2 + P_1 h_{b,d_1} + P_u h_{b,u}} > \frac{P_2 h_d}{\sigma^2 + P_1 \eta_1 + P_u h_{d_1,u}}.$$

Since practical systems are interference-limited [98, 99], the noise power is negligible compared to the interfering terms, which yields the SIC condition:

$$P_1(h_{b,d_2}\eta_1 - h_d h_{b,d_1}) + P_u(h_{d_1,u}h_{b,d_2} - h_d h_{b,u}) > 0. \tag{6.10}$$

In addition to condition (6.10), the PMCs must be verified. Since b decodes m_2 first, then we have the following PMC for the decoding of m_2:

$$P_2 h_{b,d_2} > P_1 h_{b,d_1} + P_u h_{b,u}. \tag{6.11}$$

However, the PMC for the decoding of m_1 at the level of b is given by:

$$P_1 h_{b,d_1} > P_u h_{b,u}, \tag{6.12}$$

since m_2 is subtracted prior to decoding s_1. For d_1 to be able to remove the interference of m_u prior to retrieving m_2, we must have $SINR_{d_1}^{m_u} > SINR_b^{m_u}$, which leads to:

$$\frac{P_u h_{d_1,u}}{\sigma^2 + P_1\eta_1 + P_2 h_d} > \frac{P_u h_{b,u}}{P_2 h_{b,d_2} + P_1 h_{b,d_1} + \sigma^2}$$
$$P_1(h_{d_1,u}h_{b,d_1} - h_{b,u}\eta_1) + P_2(h_{d_1,u}h_{b,d_2} - h_{b,u}h_d) > 0, \tag{6.13}$$

and the corresponding PMC is:

$$P_u h_{d_1,u} > P_2 h_d + P_1 \eta_1. \tag{6.14}$$

Regarding the mutual SIC between the receivers b and d_2, the decoding of m_1 at the level of b requires $SINR_b^{m_1}$ to be greater than $SINR_{d_2}^{m_1}$:

$$\frac{P_1 h_{b,d_1}}{\sigma^2 + P_u h_{b,u}} > \frac{P_1 h_d}{\sigma^2 + P_2\eta_2 + P_u h_{d_2,u}},$$
$$P_2 h_{b,d_1}\eta_2 > P_u(h_{b,u}h_d - h_{d_2,u}h_{b,d_1}). \tag{6.15}$$

Note that $SINR_b^{m_1}$ does not include P_2 since m_2 is decoded and canceled prior to m_1. The corresponding PMC is given by:

$$P_1 h_{b,d_1} > P_u h_{b,u}. \tag{6.16}$$

At the level of d_2, $SINR_{d_2}^{m_u}$ must be greater than $SINR_b^{m_u}$ to decode and subtract m_u before retrieving m_1. This yields the following condition:

$$SINR_{d_2}^{m_u} > SINR_b^{m_u}$$
$$\frac{P_u h_{d_2,u}}{\sigma^2 + P_2\eta_2 + P_1 h_d} > \frac{P_u h_{b,u}}{\sigma^2 + P_1 h_{b,d_1}}$$
$$P_1(h_{b,d_1}h_{d_2,u} - h_d h_{b,u}) > P_2\eta_2 h_{b,u} \tag{6.17}$$

Note that this new expression of $SINR_b^{m_u}$ does not include the interference term $P_2 h_{d_2,u}$ as it was the case in (6.13), because m_2's interference is cancelled prior to removing m_1. Finally, the PMC at the level of d_2 is given by:

$$P_u h_{d_2,u} > P_1 h_d + P_2\eta_2. \tag{6.18}$$

6.6.2 Second decoding order: b decodes m_1 then m_2

Following the same reasoning as in section 6.6.1, for the case where m_1 is decoded before m_2 at the level of b, the PMC and rate constraints for a full mutual SIC between d_1 and b, and d_2 and b, are obtained and listed below:

$$P_1\eta_1 h_{b,d_2} > P_u(h_{b,u}h_d - h_{b,d_2}h_{b,d_1}) \tag{6.19}$$
$$P_2(h_{d_1,u}h_{b,d_2} - h_{u,b}h_d) > h_{u,b}\eta_1 P_1 \tag{6.20}$$
$$P_2(h_{b,d_1}\eta_2 - h_{b,d_2}h_d) + P_u(h_{b,d_2}h_{b,d_1} - h_{b,u}h_d) > 0 \tag{6.21}$$
$$P_1(h_{d_2,u}h_{b,d_1} - h_d h_{u,b}) + P_2(h_{d_2 b}h_{d_2,u} - h_{u,b}\eta_2) > 0 \tag{6.22}$$
$$P_2 h_{b,d_2} > P_u h_{b,u} \tag{6.23}$$
$$P_u h_{d_1,u} > P_2 h_d + P_1\eta_1 \tag{6.24}$$
$$P_1 h_{b,d_1} > P_u h_{b,u} + P_2 h_{b,d_2} \tag{6.25}$$
$$P_u h_{d_2,u} > P_1 h_d + P_2\eta_2 \tag{6.26}$$

In addition to constraints eqs. (6.2a) to (6.2d), Problem (6.2) now includes eight new constraints that express the full SIC feasibility (either equations (6.10) to (6.18) or (6.19) to (6.26), depending on the decoding order). Solving this optimization problem with inequality constraints by means of the standard Karush–Kuhn–Tucker conditions implies exploring all the possible combinations of active/inactive constraints (an inequality constraint is active if it is verified with equality). This results in a total of $2^{12}-1$ combinations to be considered. To reduce this exorbitant complexity, the interplay between SIC rate conditions and PMCs is analyzed in the next section, targeting the removal of redundant constraints.

6.7 Power Allocation Problem Simplification of FD-SIC by Constraint Reduction

Consider the first decoding order at the level of b where m_2 is decoded before m_1. The PMCs for the decoding of m_1 at the level of b and of m_u at the level of d_2 are given by (6.16) and (6.18). By multiplying (6.16) by $h_{d_2,u}$ and adding it to (6.18) multiplied by $h_{b,u}$, one can eliminate P_u to obtain:

$$P_1(h_{b,d_1}h_{d_2,u} - h_d h_{b,u}) > P_2\eta_2 h_{b,u},$$

which is the SIC condition (6.17) introduced to remove m_u at the level of d_2. Also, eliminating P_1 from the two PMCs by means of adding (6.16) multiplied by h_d to (6.18) multiplied by h_{b,d_1} yields (6.15). Consequently, the PMCs for the decoding of m_1 at the level of b, and m_u at the level of d_2 imply their counterpart rate conditions. Moreover, it is noted from (6.17) that the same necessary condition (6.6) that is found in HD-SIC between b and d_2 as receivers, is obtained for the application of FD-SIC between d_2 and b:

$$h_{b,d_1}h_{d_2,u} > h_d h_{b,u}. \tag{6.6}$$

Note that if (6.6) is not true, (6.17) becomes impossible to satisfy no matter P_1 and P_2; however, when (6.6) is true, (6.17) can be satisfied under an adequate power play between P_1 and P_2.

We now move to the PMCs and SIC conditions for the decoding of m_2 and m_u at the level of b and d_1 respectively, i.e. (6.11), (6.14), (6.10) and (6.13).

By adding (6.11) multiplied by h_d to (6.14) multiplied by h_{b,d_2}, P_2 is eliminated to yield:

$$P_u(h_{d_1,u}h_{b,d_2} - h_{b,u}h_d) > P_1(h_{b,d_1}h_d + \eta_1 h_{b,d_2}), \tag{6.27}$$

which can be further transformed into:

$$\begin{gathered} P_1(\eta_1 h_{b,d_2} - h_{b,d_1}h_d) + P_u(h_{d_1,u}h_{b,d_2} - h_{b,u}h_d) > 2P_1\eta_1 h_{b,d_2} \\ \Rightarrow P_1(\eta_1 h_{b,d_2} - h_{b,d_1}h_d) + P_u(h_{d_1,u}h_{b,d_2} - h_{b,u}h_d) > 0. \end{gathered}$$

Thus, PMCs (6.11) and (6.14) imply (6.10). In fact, not only do they imply the rate condition, but it is clear that the PMCs represent more restrictive constraints than rate conditions. Finally, eliminating P_u from the PMCs through the combination of (6.11) multiplied by $h_{d_1,u}$ with (6.14) multiplied by $h_{b,u}$ yields:

$$P_2(h_{b,d_2}h_{d_1,u} - h_d h_{b,u}) > P_1(h_{b,d_1}h_{d_1,u} + \eta_1 h_{b,u}), \tag{6.28}$$

which can be rearranged into:

$$\begin{gathered} P_2(h_{b,d_2}h_{d_1,u} - h_d h_{b,u}) > P_1(h_{b,d_1}h_{d_1,u} + \eta_1 h_{b,u}) \\ P_2(h_{d_1,u}h_{b,d_2} - h_{b,u}h_d) + P_1(h_{d_1,u}h_{b,d_1} - h_{b,u}\eta_1) > 2P_1 h_{b,d_1}h_{d_1,u} \Rightarrow (6.13). \end{gathered}$$

Once again, the PMCs for the decoding of m_2 and m_u at b and d_1 imply their rate condition counterparts. Note that the necessary channel condition that appears from (6.27) and (6.28) is the same as in the case of HD-SIC in the second half-time slot:

$$h_{d_1,u}h_{b,d_2} > h_{b,u}h_d. \tag{6.8}$$

Also, the combinations of (6.16) with (6.14), and (6.18) with (6.11), while eliminating P_u, give the following condition:

$$P_1(h_{b,d_1}h_{d_1,u} - \eta_1 h_{b,u}) > P_2 h_d h_{b,u},$$
$$P_2(h_{b,d_2}h_{d_2,u} - \eta_2 h_{b,u}) > P_1(h_{b,d_1}h_{d_2,u} + h_d h_{b,u}).$$

These inequalities yield two other necessary, but not sufficient, channel conditions for the application of full SIC to the system:

$$h_{b,d_1}h_{d_1,u} > \eta_1 h_{b,u}, \tag{6.29}$$
$$h_{b,d_2}h_{d_2,u} > \eta_2 h_{b,u}. \tag{6.30}$$

Repeating the same procedure for the second decoding order delivers the same results:
1) the PMCs encompass the rate conditions,
2) the same four necessary channel conditions (6.6), (6.8), (6.29), and (6.30) are obtained.

Therefore, in the FD-SIC scenario, the system checks the validity of (6.6),(6.8), (6.29), and (6.30) prior to solving the PA problem for each decoding order. If the channel conditions are not valid or no solution is obtained for (6.2), the FD-SIC algorithm reverts to the FD-NoSIC procedure described in section 6.3.1.
As a conclusion for this section, Problem (6.2) is now only equipped with the PMC set corresponding to the decoding order (i.e. eqs. (6.11), (6.14), (6.16) and (6.18), or eqs. (6.23) to (6.26)), in addition to constraints eqs. (6.2a) to (6.2d). This reduces the number of combinations of active/inactive constraints from $2^{12} - 1$ to $2^8 - 1$ which is still considerable. The aim of the next section is to workaround the need of a full search over the corresponding 255 cases for determining the optimal PA. This is done by efficiently determining the meaningful constraint combinations, based on the geometrical interpretation of the FD-SIC PA problem. Considerable complexity reductions arise from this approach as shown next.

6.8 Solution for FD-SIC Optimal Power Allocation

The proposed geometrical resolution of the FD-SIC D2D rate maximization PA problem is presented in detail for the first decoding order. First, the geometrical representation of the solution space satisfying the PMCs and power limit constraints is provided. Then, a procedure is elaborated leading to the reduction of the search space to the minimum required. Afterwards, the optimization is conducted on the resulting reduced search space. At last, a quick summary of the optimal PA procedure is presented including the required changes to obtain the optimal PA for the second decoding order.

6.8.1 3D Solution Space Representation

The four PMCs that must be satisfied for the first decoding order (eqs. (6.11), (6.14), (6.16) and (6.18)) are re-written in the following form:

$$P_u h_{b,u} < P_2 h_{b,d_2} - P_1 h_{b,d_1} \qquad (PMC_1)$$
$$P_u h_{d_1,u} > P_2 h_d + P_1 \eta_1 \qquad (PMC_2)$$
$$P_u h_{b,u} < P_1 h_{b,d_1} \qquad (PMC_3)$$
$$P_u h_{d_2,u} > P_1 h_d + P_2 \eta_2 \qquad (PMC_4)$$

In the 3D space of axes x, y, z representing variables P_1, P_2 and P_u respectively, we introduce planes $\mathcal{PL}_1, \mathcal{PL}_2, \mathcal{PL}_3$ and $\mathcal{PL}_4$ whose equations are given by PMCs 1, 2, 3 and 4 when the conditions are met with equality. In the following, we refer to $\mathcal{PL}_i$ as the plane *derived from*, or equivalently, *corresponding to*, or simply, as *the plane of* PMC_i. Each PMC restricts the search space either to the half space below its corresponding plane like for PMC_1 and PMC_3, or to the half space above its corresponding plane as for PMC_2 and PMC_4. On the other hand, the transmit power limits restrict the search space to the region within the parallelepiped defined by the sides $x = P_{1,M}$, $x = 0$, $y = P_{2,M}$, $y = 0$, $z = P_{u,M}$, $z = P_{u,m}$. To have a non-empty search space (i.e. FD-SIC is feasible), the

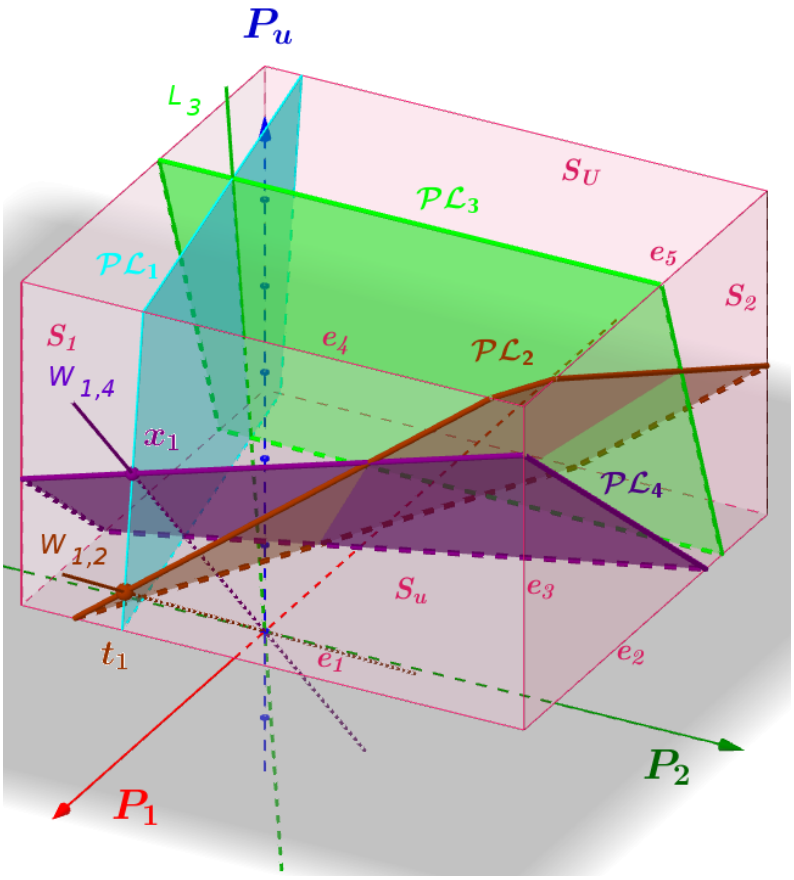


Figure 6.3 – Schematic of the search space formed inside the intersection of the PMC planes with the parallelepiped of power limits.

pentahedron defined by the space region above $\mathcal{PL}_2$ and $\mathcal{PL}_4$ and below $\mathcal{PL}_1$ and $\mathcal{PL}_3$ must be non-empty, and it must have a common region with the parallelepiped.

- Non-empty pentahedron: The pentahedron is non-empty if the planes $\mathcal{PL}_1$ and $\mathcal{PL}_3$ are *on top of* $\mathcal{PL}_2$ and $\mathcal{PL}_4$. For that to be the case, the intersection lines of $\mathcal{PL}_1$ with $\mathcal{PL}_2$ and $\mathcal{PL}_4$ ($W_{1,2}$ and $W_{1,4}$ respectively), must be below $\mathcal{PL}_3$, as shown in Fig. 6.3.
  Let $\vec{u}$ be the direction vector of $W_{1,2}$; $W_{1,2}$ is below $\mathcal{PL}_3$ if and only if the slope of $W_{1,2}$'s projection on the (P_1, P_u) plane is less steep than that of $\mathcal{PL}_3$. This translates into having $\dfrac{z(\vec{u})}{x(\vec{u})} < h_{b,d_1}/h_{b,u}$, which is shown in appendix 6.A to yield the following

channel condition:

$$h_{b,d_1}h_{d_1,u} - \eta_1 h_{b,u} > 2h_{b,u}h_d\frac{h_{b,d_1}}{h_{b,d_2}}. \tag{6.31}$$

Condition (6.31) imposes more restrictive constraints on $h_{b,d_1}h_{d_1,u} - \eta_1 h_{b,u}$ than in (6.29) which was expected, as it turns the previously necessary channel condition into a sufficient one. Moreover, (6.31) can also be equivalently rewritten as follows:

$$h_{b,d_2}h_{d_1,u} - h_{b,u}h_d > \frac{h_{b,u}}{h_{b,d_1}}(\eta_1 h_{b,d_2} + h_{b,d_1}h_d), \tag{6.32}$$

which is also an enhanced constraint on $h_{b,d_2}h_{d_1,u} - h_{b,u}h_d$ with respect to (6.8) to turn it into a sufficient constraint.
Following the same approach for $W_{1,4}$ (c.f. appendix 6.A), the necessary channel condition can be written in the two equivalent forms:

$$h_{b,d_1}h_{d_2,u} - h_d h_{b,u} > 2h_{b,u}\eta_2\frac{h_{b,d_1}}{h_{b,d_2}}, \tag{6.33}$$

$$h_{b,d_2}h_{d_2,u} - h_{b,u}\eta_2 > \frac{h_{b,u}}{h_{b,d_1}}(h_{b,d_2}h_d + \eta_2 h_{b,d_1}). \tag{6.34}$$

Again, the necessary condition expressed in (6.33) and (6.34) is more restrictive than the necessary conditions of (6.6) and (6.30).

- Pentahedron ∩ parallelepiped: For the pentahedron to have a non-empty intersection with the parallelepiped, it is sufficient to make sure that the intersection line of $\mathcal{PL}_1$ with $\mathcal{PL}_3$ (L_3) intersects the plane of equation $z = P_{u,m}$ within the $P_{1,M}$ and $P_{2,M}$ limits. These conditions on the x, y coordinates of $\mathcal{PL}_3 \cap \mathcal{PL}_1 \cap P_{u,m}$ yield the constraints:

$$P_{u,m}\frac{h_{b,u}}{h_{b,d_1}} < P_{1,M} \quad \& \quad 2P_{u,m}\frac{h_{b,u}}{h_{b,d_2}} < P_{2,M}. \tag{6.35}$$

Conditions (6.31), (6.33) and (6.35) form the necessary and sufficient constraints for the existence of a solution to the FD-SIC PA problem according to the first decoding order.

6.8.2 Search Space Reduction

We prove in this section that the optimal solution lies on the intersection line of $\mathcal{PL}_2, \mathcal{PL}_4$ or the lower side of the parallelepiped, S_u, with one of the outer sides of the parallelepiped, S_1, S_2, S_U (cf. Fig. 6.3), respectively defined by: $x = P_{1,M}$ for $(y, z) \in [0, P_{2,M}] \times [P_{u,m}, P_{u,M}]$, $y = P_{2,M}$ for $(x, z) \in [0, P_{1,M}] \times [P_{u,m}, P_{u,M}]$, and $z = P_{u,M}$ for $(x, y) \in [0, P_{1,M}] \times [0, P_{2,M}]$.

Proposition 6.1. *The optimal solution lies on one of the outer sides of the parallelepiped.*

Proof. The D2D rate is given by:

$$R_{D2D}(P_1, P_2) = B\log_2(1 + \frac{P_1 h_d}{P_2\eta_2 + \sigma^2}) + B\log_2(1 + \frac{P_2 h_d}{P_1\eta_1 + \sigma^2})$$

For any couple (P_1, P_2), and $\forall \beta > 1$, the throughput of $(\beta P_1, \beta P_2)$ is greater than $R_{D2D}(P_1, P_2)$ since:

$$\begin{aligned} R_{D2D}(\beta P_1, \beta P_2) =& B\log_2(1 + \frac{P_1 h_d}{P_2\eta_2 + \sigma^2/\beta}) + B\log_2(1 + \frac{P_2 h_d}{P_1\eta_1 + \sigma^2/\beta}) \\ >& B\log_2(1 + \frac{P_1 h_d}{P_2\eta_2 + \sigma^2}) + B\log_2(1 + \frac{P_2 h_d}{P_1\eta_1 + \sigma^2}) \\ =& R_{D2D}(P_1, P_2). \end{aligned}$$

Therefore, given an initial triplet (P_1, P_2, P_u), a higher throughput-achieving triplet can be obtained by simply multiplying the components by a factor larger than 1. The higher β, the higher the throughput, meaning that β should be increased until reaching the boundaries of the region, which can be either $P_{1,M}, P_{2,M}$ or $P_{u,M}$. □

Moreover, the D2D rate is independent of P_u. This means that when moving on a vertical line in the solution space, R_{D2D} is constant and P_u only affects the CU rate. To keep the CU rate as close as possible to $R_{u,min}$, we select the smallest P_u value from the range of admissible values for a given (P_1, P_2) couple. Since every point in the solution space must be on top of $\mathcal{PL}_2$ and $\mathcal{PL}_4$, the minimum allowed value of P_u is given by forcing the equality either on PMC_2 or on PMC_4, according to the one that delivers the higher minimum value of P_u for the considered (P_1, P_2) couple.

As a conclusion, the optimal solution lies on the intersection segment of one of the outer sides of the parallelepiped, S_1, S_2, or S_U, with one of the planes $\mathcal{PL}_2$, $\mathcal{PL}_4$, or S_u. S_u intersects S_1 and S_2 in the edges e_1 and e_2 (c.f. Fig. 6.3), whereas $\mathcal{PL}_2$ and $\mathcal{PL}_4$ can yield three intersection lines each, one with S_1, one with S_2 and one with S_U. Thus, the search space is reduced to these eight intersection segments. However, given the shape of the solution space, some of these segments are mutually exclusive. The aim of the next section is to determine which subset of segments should be accounted for in the power optimization process, depending on the channel conditions of the D2D-CU couple.

6.8.3 Selection of the Useful Intersections

As can be seen from Fig. 6.3, some of the eight intersections can be discarded. For example, the intersection of S_u with S_1 and S_2 is not relevant, since the value of P_u is decided by PMC_2 and PMC_4, whose planes are on top of S_u near sides S_1 and S_2. Fig. 6.4 shows the projection on plane (P_1, P_2) of the partition of the space into two vertical regions where PMC_4 encompasses PMC_2 for region 1, and PMC_2 encompasses PMC_4 for region 2. The plane separating the two regions is the vertical plane passing through the straight line $L_\lambda \triangleq \mathcal{PL}_4 \cap \mathcal{PL}_2$. Therefore, for the case of Figs. 6.3 and 6.4, the D2D rate optimization is to be conducted over segment $\overline{x_1 i} \cup \overline{i v_4}$ which is included in S_1, over segment $\overline{v_4 v_5}$ included in S_U, and over segment $\overline{v_5 s_2}$ included in S_2. By doing so, the optimization over segments $\overline{t_1 i}$, $\overline{i g_3}$, $\overline{g_3 g_2}$ and $\overline{g_2 j_2}$ is avoided.
Therefore, the first step in reducing the number of intersections to be considered lies in determining which of PMC_4 and PMC_2 encompasses the other, and for which region of the space. To that end, a schematic of $\mathcal{PL}_2$ and $\mathcal{PL}_4$ is presented in Figs. 6.5a and 6.5b, showing their intersection with the planes defined by $P_1 = 0$ and $P_2 = 0$. The angles of these intersection lines and their slopes are shown in Fig. 6.5.

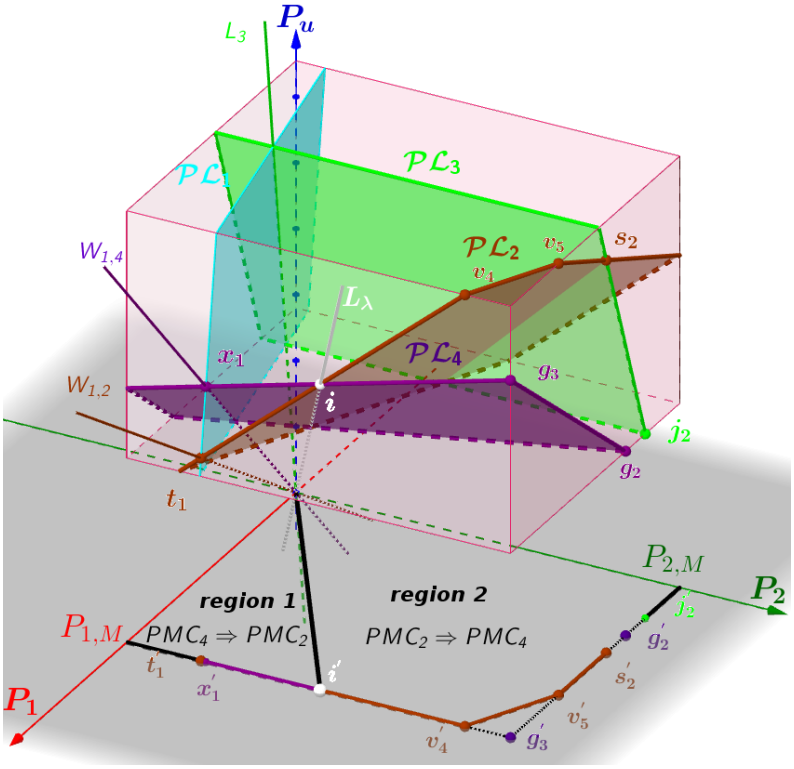


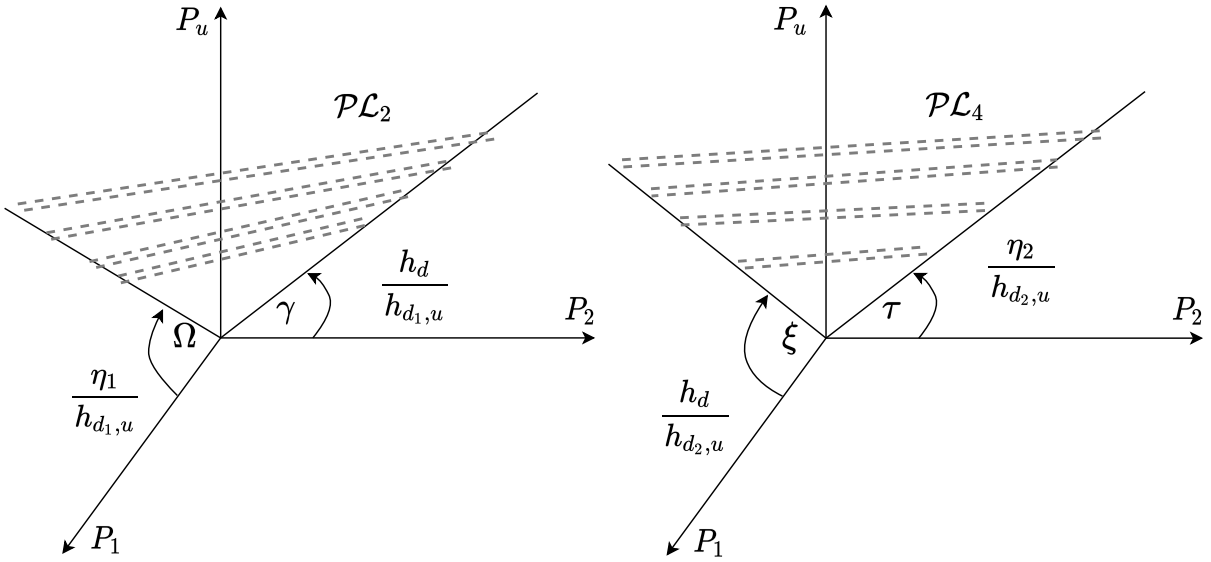
Figure 6.4 – Schematic of the solution space showing the regions of dominance of PMC_4 over PMC_2 and vice-versa.

6.8.3.1 Interplay between PMC_2 and PMC_4

Depending on the angles Ω, γ, ξ and τ, four cases are identified to determine the interplay between PMC_2 and PMC_4:

1. $\Omega > \xi$, $\gamma > \tau$: PMC_2 encompasses PMC_4 ($PMC_2 \Rightarrow PMC_4$) over all the positive (P_1, P_2) plane.
2. $\Omega < \xi$, $\gamma < \tau$: PMC_4 encompasses PMC_2 ($PMC_4 \Rightarrow PMC_2$) over all the positive (P_1, P_2) plane.
3. $\Omega < \xi$, $\gamma > \tau$: PMC_4 encompasses PMC_2 in region 1 and PMC_2 encompasses PMC_4 in region 2, (cf. Fig. 6.4).
4. $\Omega > \xi$, $\gamma < \tau$: PMC_2 encompasses PMC_4 in region 1 and PMC_4 encompasses PMC_2 in region 2.

Before proceeding, note that even for cases 3) and 4), it is still possible for a PMC to encompass the other on the entire search space if the whole search space is included either in region 1 or 2. This is depicted in the examples of Figs. 6.6 and 6.7 which take back the conditions of Fig. 6.4 with some modifications. In Fig. 6.6, $\mathcal{PL}_1$ is such that $W_{1,4}$ is at the right side of L_λ ($W_{1,4}$ is in region 2), then the search space is included in region 2 and only PMC_2 needs to be accounted for. The other scenario is represented in Fig. 6.7 where $\mathcal{PL}_3 \cap \mathcal{PL}_2$ is at the left side of L_λ (in region 1), hence PMC_4 encompasses PMC_2 over the entirety of the search space. The first scenario occurs when L_λ is on top of $\mathcal{PL}_1$, and the second one occurs when L_λ is on top of $\mathcal{PL}_3$.



(a) PMC_2 (b) PMC_4

Figure 6.5 – Isolated schematics of $\mathcal{PL}_2$ and $\mathcal{PL}_4$ in the 3D space.

To determine if the search space is totally included in region 1 or 2 for the cases 3) and 4), we introduce f_1, f_3 and f_λ, the functions of P_1, P_2 which yield the P_u value corresponding to the planes $\mathcal{PL}_1, \mathcal{PL}_3$ and to L_λ. A parametric equation of L_λ is given by:

$$L_\lambda = \begin{cases} x = \left(\frac{h_d}{h_{d_1,u}} - \frac{\eta_2}{h_{d_2,u}}\right) m = [\tan(\gamma) - \tan(\tau)]m \\ y = \left(\frac{h_d}{h_{d_2,u}} - \frac{\eta_1}{h_{d_1,u}}\right) m = [\tan(\xi) - \tan(\Omega)]m \\ z = \frac{h_d^2 - \eta_1\eta_2}{h_{d_1,u}h_{d_2,u}} m \end{cases}$$

In the case of Fig. 6.6, the search space is included in region 2 if and only if L_λ is on top of $\mathcal{PL}_1$. For the case of Fig. 6.7, the search space is included in region 1 if and only if L_λ is on top of $\mathcal{PL}_3$. To determine the conditions of each scenario, we first have to check if the conditions of case 3), where $\gamma > \tau$ and $\xi > \Omega$, or those of case 4), where $\gamma < \tau$ and $\xi < \Omega$, are met. To study the relative position of L_λ with respect to $\mathcal{PL}_1$ and $\mathcal{PL}_3$, m is chosen such that the comparison is conducted in the first octant. Since in case 3), $\gamma > \tau \Rightarrow \tan(\gamma) - \tan(\tau) > 0$, then m must be positive in case 3) and, conversely, negative in case 4).
The search space is included in region 2 if:

$$f_\lambda(P_1, P_2) > f_1(P_1, P_2)$$
$$\Rightarrow \frac{h_d^2 - \eta_1\eta_2}{h_{d_1,u}h_{d_2,u}} m > \frac{P_2 h_{b,d_2} - P_1 h_{b,d_1}}{h_{b,u}}$$

Replacing P_1 by $(h_d/h_{d_1,u} - \eta_2/h_{d_2,u})m$, and P_2 by $(h_d/h_{d_2,u} - \eta_1/h_{d_1,u})m$, we get:

$$\frac{h_{b,u}(h_d^2 - \eta_1\eta_2)}{h_{d_1,u}h_{d_2,u}} m > (\frac{h_d}{h_{d_2,u}} - \frac{\eta_1}{h_{d_1,u}})m h_{b,d_2} - (\frac{h_d}{h_{d_1,u}} - \frac{\eta_2}{h_{d_2,u}})m h_{b,d_1}$$

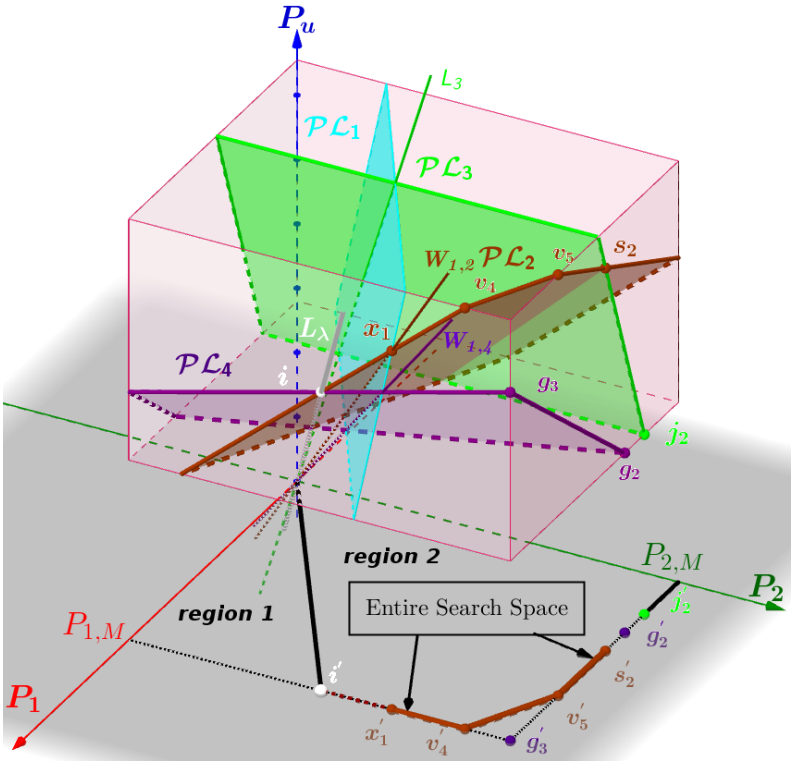


Figure 6.6 – Figure representing case 3) with the solution search space included in region 2.

Let Γ be the following proposition:

$$\frac{h_{b,u}(h_d^2 - \eta_1\eta_2)}{h_{d_1,u}h_{d_2,u}} > \frac{h_{b,d_2}h_d + h_{b,d_1}\eta_2}{h_{d_2,u}} - \frac{h_{b,d_2}\eta_1 + h_{b,d_1}h_d}{h_{d_1,u}}$$

Since m should be positive for case 3) and negative for case 4), we conclude that:

- The search space included in region 2 for case 3) is equivalent to having the proposition $\Gamma = 1$.
- The search space included in region 2 for case 4) is equivalent to having the proposition $\Gamma = 0$.

On the other hand, the search space is included in region 1 if:

$$f_\lambda(P_1, P_2) > f_3(P_1, P_2)$$
$$\Rightarrow (h_d^2 - \eta_1\eta_2)m > \frac{(h_d h_{d_2,u} - \eta_2 h_{d_1,u})h_{b,d_1}}{h_{b,u}}m$$

Let Ξ be the following proposition:

$$(h_d^2 - \eta_1\eta_2)h_{b,u} > (h_d h_{d_2,u} - \eta_2 h_{d_1,u})h_{b,d_1} \tag{6.36}$$

Therefore, the search space is included in region 1 if:

- $\Xi = 1$ for case 3),
- $\Xi = 0$ for case 4).

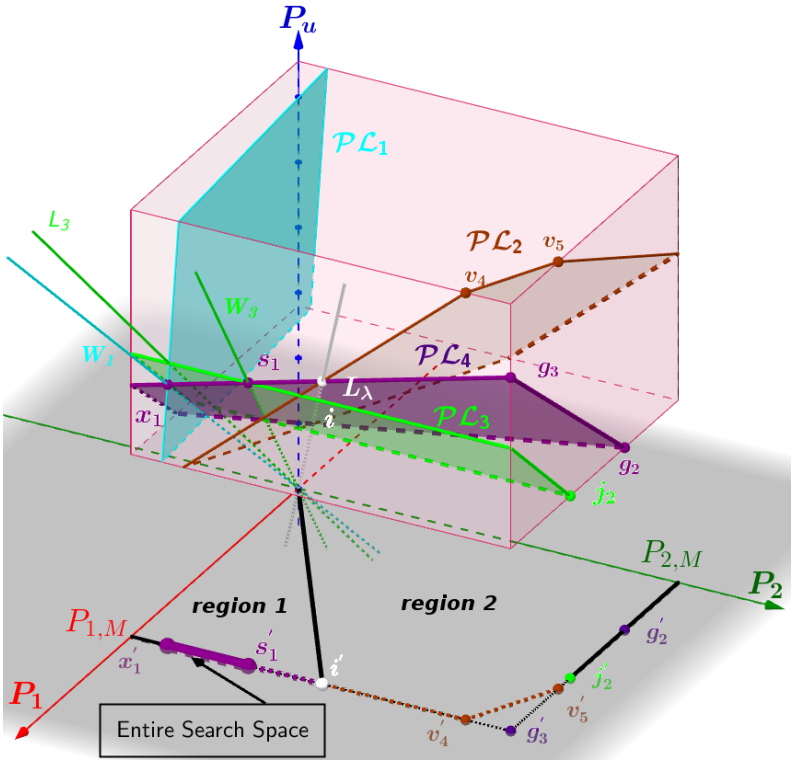


Figure 6.7 – Figure representing case 3) with the solution search space included in region 1.

Conclusion: to determine if the search space is completely included in one of the two regions, for cases 3) and 4), we simply have to test the validity of Γ and Ξ and draw the corresponding conclusion to each case.

To summarize, by comparing Ω to ξ and γ to τ, and according to the values of Ξ and Γ in cases 3) and 4), the number of intersections to be considered is reduced by selecting the appropriate PMC between PMC_2 and PMC_4 in the corresponding space region. For the sake of clarity, we introduce $PMC_{2,4}$ as the efficient combination of PMC_2 and PMC_4, given by:

$$P_u \geq \begin{cases} \dfrac{P_2 h_d + P_1 \eta_1}{h_{d_1,u}}, & \text{if } P_2\left(\dfrac{h_d}{h_{d_1,u}} - \dfrac{\eta_2}{h_{d_2,u}}\right) > P_1\left(\dfrac{h_d}{h_{d_2,u}} - \dfrac{\eta_1}{h_{d_1,u}}\right) \\ \dfrac{P_1 h_d + P_2 \eta_2}{h_{d_2,u}}, & \text{elsewhere.} \end{cases}$$

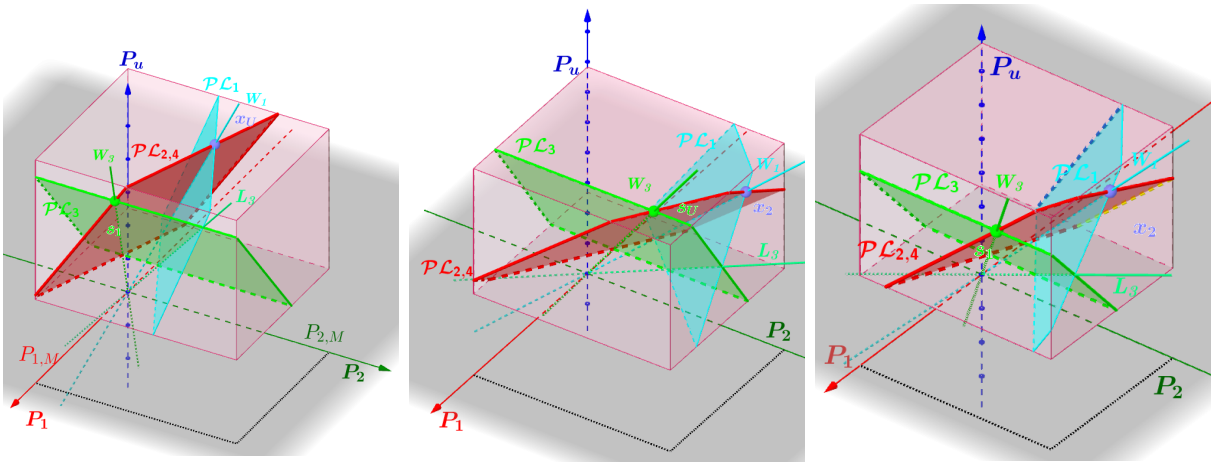
6.8.3.2 Selection of the Useful Parallelepiped Sides

With $PMC_{2,4}$ at hand, the next step is to reduce the unnecessary sides of the parallelepiped. Unnecessary sides are defined as those which do not intersect with $\mathcal{PL}_{2,4}$, or those whose intersection with $\mathcal{PL}_{2,4}$ is outside the range of allowed values between $\mathcal{PL}_1$ and $\mathcal{PL}_3$. To that end, we study PMC_1 and PMC_3 which do not affect the intersection segments (of $\mathcal{PL}_{2,4}$ with the parallelepiped sides) as such, but rather the end points of these intersection segments. A typical example is given in Fig. 6.6 where PMC_1 sets the end point x_1 from the side S_1, and PMC_3 sets the end point s_2 from the side S_2.

Let W_1 regroup the intersection lines $W_{1,2}$ and $W_{1,4}$ such that $W_1 = \mathcal{PL}_{2,4} \cap \mathcal{PL}_1$, and let W_3 be the intersection line of $\mathcal{PL}_3$ with $\mathcal{PL}_{2,4}$ (cf. Fig. 6.8). Each of W_1 and W_3 may

intercept sides S_1 or S_2 or S_U, yielding a total of nine potential combinations. Since each side S_i is a rectangular surface within the infinite plane $\mathcal{S}_i$ of equation $P_i = P_{i,M}$, then W_1 and W_3 can intercept only one side of the parallelepiped (S_u aside) for a given channel configuration. Let x_i and s_i be the intersection points of W_1 and W_3 with S_i, we have:

$$x_i = W_1 \cap S_i,\ s_i = W_3 \cap S_i, \forall i \in \{1, 2, U\}.$$



(a) W_1 intercepts S_U and W_3 intercepts S_1

(b) W_1 intercepts S_2 and W_3 intercepts S_U

(c) W_1 intercepts S_2 and W_3 intercepts S_1

Figure 6.8 – The 3 non-feasible combinations between x_i and s_i for a successful FD-SIC.

To determine which sides are intercepted by W_1 (resp. W_3), i.e. to determine if we have x_1, x_2 or x_U (resp. s_1, s_2 or s_U), we consider the points xl_i (resp. sl_i), intersections of W_1 (resp. W_3) with the planes $\mathcal{S}_i, \forall i \in \{1, 2, U\}$. The coordinates of xl_i are given by:

$$xl_1 = \begin{pmatrix} P_{1,M} \\ y(W_1 \cap \mathcal{S}_1) \\ z(W_1 \cap \mathcal{S}_1) \end{pmatrix}, xl_2 = \begin{pmatrix} x(W_1 \cap \mathcal{S}_2) \\ P_{2,M} \\ z(W_1 \cap \mathcal{S}_2) \end{pmatrix}, xl_U = \begin{pmatrix} y(W_1 \cap \mathcal{S}_U) \\ y(W_1 \cap \mathcal{S}_U) \\ z = P_{u,M} \end{pmatrix}$$

Then, two tests are needed to determine which of x_1, x_2 or x_U occurs for the given channel states as shown in Algorithm 6.1.

Note that if $y(xl_1) < P_{2,M}$ while $z(xl_1) < P_{u,m}$, even though $xl_1 \notin S_1$, we still say that xl_i *is on the side* S_1 (or on the side of P_1) and this case is associated to that of $x_i = x_1$. In this situation the face S_1 still hosts an optimization segment, however the endpoint previously given by $x_1 = \mathcal{PL}_1 \cap \mathcal{PL}_{2,4} \cap S_1$ is now given by $k_1 = \mathcal{PL}_1 \cap S_u \cap S_1$.
Similarly, on *the side of* P_2, if $y(xl_1) > P_{2,M}$ while $z(xl_1) < P_{u,m}$, then the case is associated to that of $x_i = x_2$, but the point $k_2 = \mathcal{PL}_1 \cap S_u \cap S_2$ sets the segment endpoint instead of $x_2 (= S_2 \cap W_1 = \emptyset)$. The same tests are replicated for s_i.
From the nine possibilities, only six combinations are actually viable because the pairs (x_U, s_1), (x_2, s_U) and (x_2, s_1) cannot be achieved without violating (6.31) or (6.33) as can be seen in Fig. 6.8. Indeed, the three cases shown in Fig. 6.8 lead to empty search spaces. The six viable pairs are given in Table 6.1 with the correspondence between the pairs and the parallelepiped sides hosting the useful intersection segments.

Note that if $\mathcal{PL}_1$ and $\mathcal{PL}_3$ intercept $\mathcal{PL}_{2,4}$ at the same side, then the search space can be reduced to a single segment as it is the case for the first, the fourth and the fifth

Algorithm 6.1 W_1 intersection with the parallalepiped

input : $P_{1,M}$, $P_{2,M}$, $P_{u,M}$, $P_{u,m}$, $h_{b,u}$, h_d, η_1, η_2, $h_{d1,u}$, $h_{d_2,u}$, $h_{d_1,b}$, $h_{d_2,b}$

Result: Returns $i/W_1 \cap S_i = x_i = xl_i \neq \emptyset$.

```
if y(xl_1) < P_{2,M} then
    if z(xl_1) ≤ P_{u,M} then
        i = 1, keep x_1
    else
        i = U, keep x_U
    end
else
    if z(xl_2) < P_{u,M} then
        i = 2, keep x_2
    else
        i = U, keep x_U
    end
end
```

	S_1	S_2	S_U
x_U and s_U			✓
x_1 and s_U	✓		✓
x_U and s_2		✓	✓
x_1 and s_1	✓		
x_2 and s_2		✓	
x_1 and s_2	✓	✓	Depends

Table 6.1 – Table showing the sides involved in the D2D rate optimization for each of the six (x_i, s_j) viable pairs due to PMC_1 and PMC_3.

rows in Table 6.1. For the second and third rows, two segments are involved in the D2D rate optimization. Finally, in the case where W_1 intercepts S_1 and W_3 intercepts S_2 (as in Fig. 6.4), the segment $\overline{v_4v_5}$ belonging to S_U is to be included in the D2D optimization process – in addition to the segments in S_1 and S_2 – if and only if the value of P_u obtained from $\mathcal{PL}_{2,4}$ at $P_1 = P_{1,M}$ and $P_2 = P_{2,M}$ is greater than $P_{u,M}$.

6.8.3.3 Segments Endpoints

Having determined the relevant intersection segments (a maximum of three segments) for the D2D rate optimization using PMCs 1 and 3, we detail hereafter how the endpoints of every segment are determined for each side of the parallelepiped. For the sake of clarity, let e_1, e_2, e_3, e_4, e_5 be the edges of the parallelepiped (cf. Fig. 6.3) given by:

$$e_1 = S_u \cap S_1,\ \ e_2 = S_u \cap S_2,\ \ e_4 = S_U \cap S_1,\ \ e_5 = S_U \cap S_2,\ \ e_3 = S_2 \cap S_1.$$

Also, let the three families of points v_i, g_i, and w_i be the intersections of $\mathcal{PL}_2, \mathcal{PL}_4$ and $\mathcal{PL}_{2,4}$ with e_i:

$$v_i = \mathcal{PL}_2 \cap e_i, \quad g_i = \mathcal{PL}_4 \cap e_i, \quad w_i = \mathcal{PL}_{2,4} \cap e_i.$$

Examples of such points can be seen in Fig. 6.7 for v_4, v_5, g_3 and g_2. Note that points w_i are only used to designate the points v_i or g_i depending on whether we are in region 1 or 2. We can now efficiently designate the segment endpoints on each side.

6.8.3.3.1 Side S_2 : The optimization over S_2 translates into an optimization over P_1, since P_2 is equal to $P_{2,M}$. It is clear that the minimal value of P_1 is bound to PMC_3. In Fig. 6.9a for example, the minimal value of P_1 is obtained for the point s_2, intersection of $\mathcal{PL}_3$ with $\mathcal{PL}_{2,4}$. Another alternative for the minimum P_1 value is when $\mathcal{PL}_3$ intercepts the edge $e_2 = S_2 \cap S_u$ of the prism as shown in Fig. 6.9d. Therefore, the segment endpoint over S_2 is either s_2 or j_2, and the minimum P_1 value is obtained by comparing their abscissa:

$$\min P_1 = \max\left[x(\mathcal{PL}_3 \cap \mathcal{PL}_{2,4} \cap S_2), x(\mathcal{PL}_3 \cap S_u \cap S_2)\right],$$
$$\min P_1 = \max\left[x(s_2), x(j_2)\right].$$

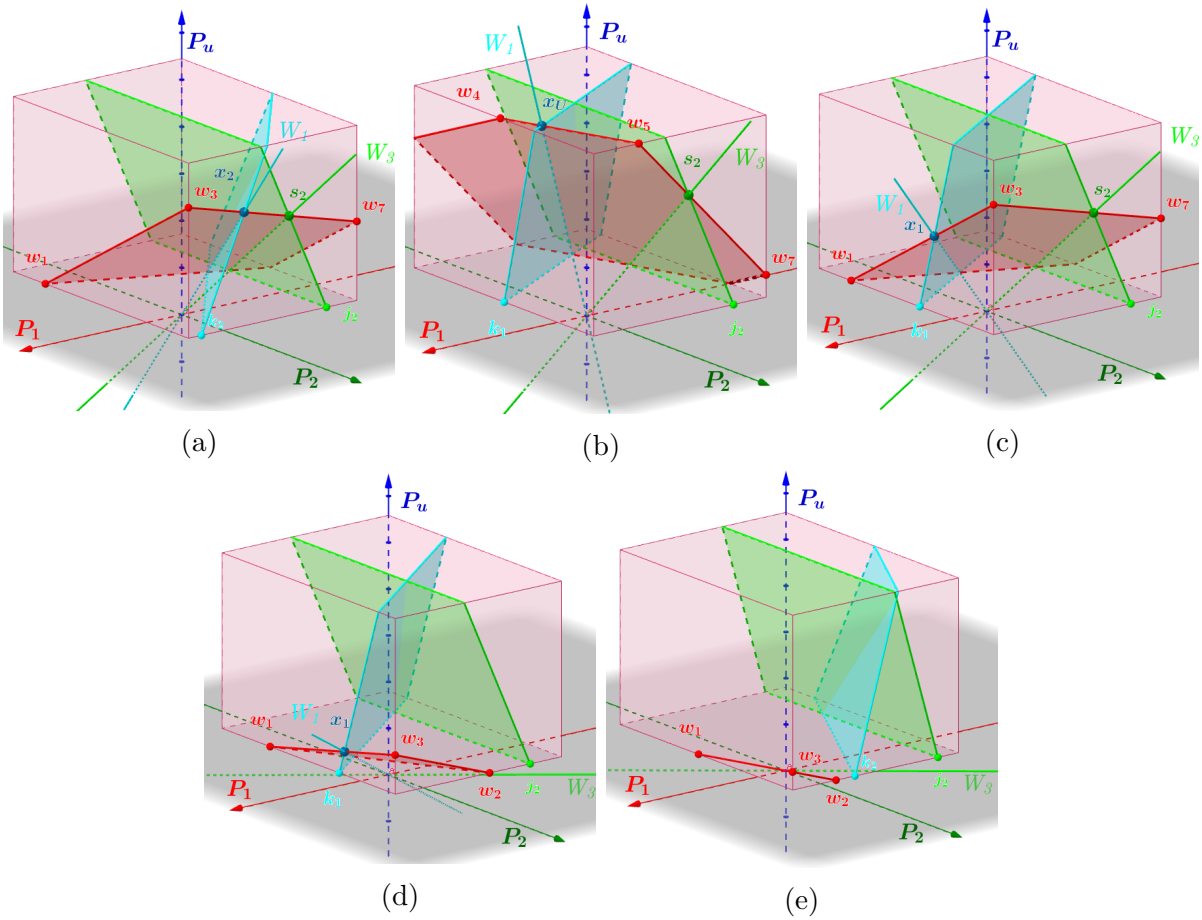


Figure 6.9 – The possible combinations of the segment endpoints (maximum P_1, minimum P_1) over S_2.

Regarding the maximum value of P_1, it can be due to the intersection of $S_2 \cap \mathcal{PL}_{2,4}$ with either S_U (like for w_5 in Fig. 6.9b), S_1 (like for w_3 in Fig. 6.9c), or with $\mathcal{PL}_1$ (like for x_2 in Fig. 6.9a, corresponding to the case of x_2 and s_2 in the fifth row of table 6.1).

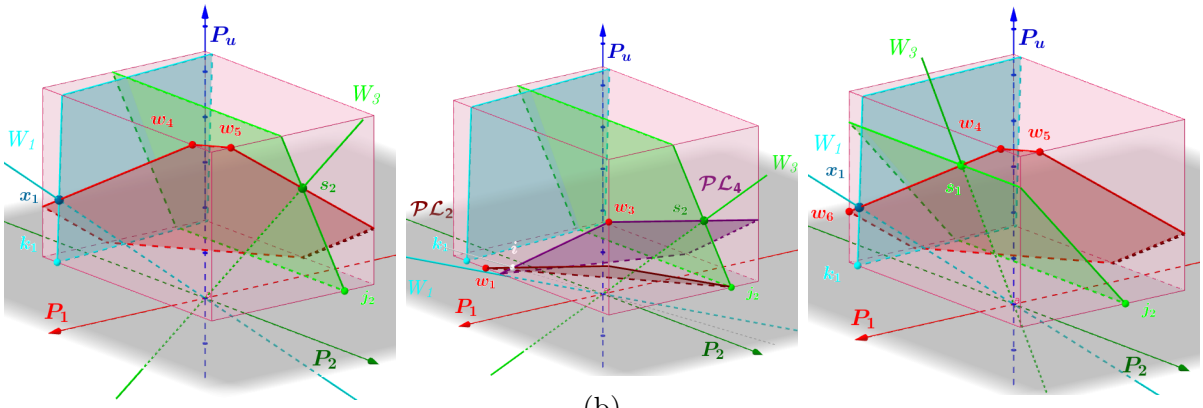
Also, the maximum P_1 value may be simply set by k_2, the intersection of $S_2 \cap \mathcal{PL}_1$ with S_u, as in Fig. 6.9e. The maximum value of P_1 is therefore given by:

$$\max P_1 = \min\left[x(\mathcal{PL}_{2,4} \cap e_2), x(\mathcal{PL}_{2,4} \cap e_3), x(\mathcal{PL}_{2,4} \cap \mathcal{PL}_1 \cap S_2), x(\mathcal{PL}_1 \cap e_2)\right],$$
$$\max P_1 = \min\left[x(w_5), x(w_3), x(x_2), x(k_2)\right].$$

Note that for the side S_2, PMC_3 is involved in the minimum P_1 value, and PMC_1 in the maximum value.

6.8.3.3.2 Side S_1 : Regarding side S_1, the minimum P_2 value is settled by PMC_1. The segment endpoint corresponding to the minimum P_2 value could be due to $\mathcal{PL}_{2,4} \cap S_1$ intercepting either $\mathcal{PL}_1$ as in Fig. 6.10a, or the edge $e_1 = S_1 \cap S_u$ as in Fig. 6.10b. Thus, the minimum P_2 value is obtained from:

$$\min P_2 = \max\left[y(\mathcal{PL}_1 \cap \mathcal{PL}_{2,4} \cap S_1), y(\mathcal{PL}_1 \cap S_u \cap S_1)\right],$$
$$\min P_2 = \max\left[y(x_1), y(k_1)\right].$$



(a) (b) (c)

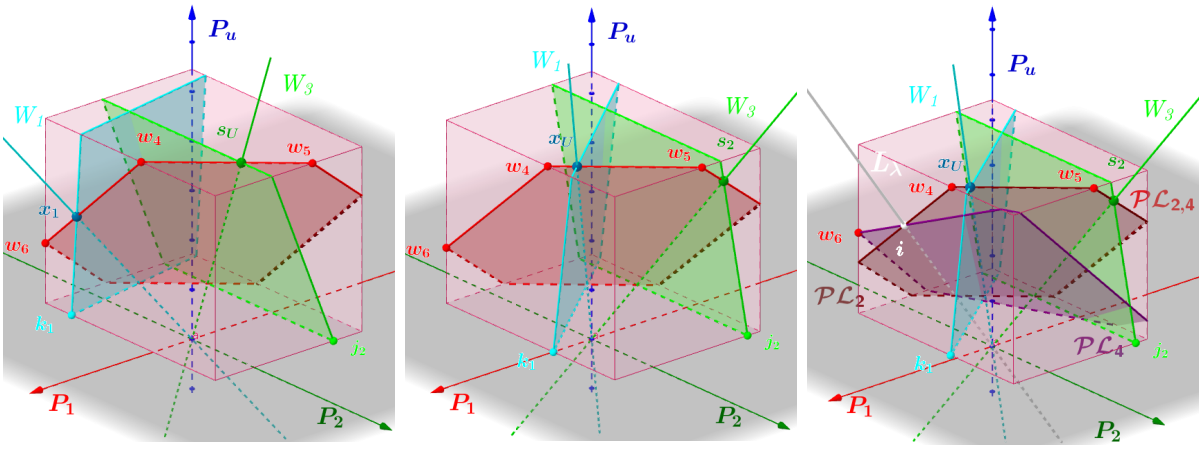
Figure 6.10 – The possible combinations of the segment endpoints (maximum P_2, minimum P_2) over S_1.

The maximum value of P_2 depends on which plane intercepts first $\mathcal{PL}_{2,4}$ among the three candidates: S_U as in Fig. 6.10a, S_2 as in Fig. 6.10b, or $\mathcal{PL}_3$ as in Fig. 6.10c (fourth row of Table 6.1). The maximum P_2 value is given by:

$$\max P_2 = \min\left[y(\mathcal{PL}_{2,4} \cap \mathcal{PL}_3 \cap S_1), y(\mathcal{PL}_{2,4} \cap S_2 \cap S_1), y(\mathcal{PL}_{2,4} \cap S_U \cap S_1)\right],$$
$$\max P_2 = \min\left[y(s_1), y(w_3) = P_{2,M}, y(w_4)\right].$$

In the example of Fig. 6.10b, the intersection segment starts at w_1 and ends at w_3 passing by i. Although $\overline{w_1 i} \cup \overline{i w_3}$ is a different segment from $\overline{w_1 w_3}$, their projections over the (P_1, P_2) plane are identical. Thus, we are only interested in segment ends over both sides S_1 and S_2. However, for the case of S_U, the intersection point of $\mathcal{PL}_2$ and $\mathcal{PL}_4$ has an impact over the segments end points since the projection of the segments is affected as it is discussed next.

6.8.3.3.3 Side S_U : Unlike for the other sides, none of P_1 or P_2 is fixed, but P_2 can be expressed in terms of P_1; therefore, we evaluate the position of the endpoints of the segments on S_U in terms of maximum P_1 and minimum P_1.
When $L_\lambda = \mathcal{PL}_2 \cap \mathcal{PL}_4$ does not intercept S_U as it is the case for Fig. 6.11c for example, the intersection of $\mathcal{PL}_{2,4}$ with S_U yields a unique segment ($\overline{w_4 w_5}$ in case of Fig. 6.11c). The endpoint corresponding to the minimum value of P_1 is due to the intersection of $\mathcal{PL}_{2,4}$ with either S_2 as in Fig. 6.11b (which is further detailed in Fig. 6.11c), or with $\mathcal{PL}_3$ as in Fig. 6.11a.



(a) (b) (c)

Figure 6.11 – Possible combinations of the segment endpoints (maximum P_1, minimum P_1) over S_U when i does not reside on S_U.

The minimum P_1 value is obtained through the comparison:

$$\min P_1 = \max\left[x(\mathcal{PL}_{2,4} \cap \mathcal{PL}_3 \cap S_U), x(\mathcal{PL}_{2,4} \cap S_2 \cap S_U)\right],$$
$$\min P_1 = \max\left[x(s_U), x(w_5)\right]. \tag{6.37}$$

The endpoint corresponding to the maximum value of P_1 is due to the intersection of $\mathcal{PL}_{2,4}$ with either $\mathcal{PL}_1$ to yield x_U as in Fig. 6.11b, or S_1 to yield w_4 as in Fig. 6.11a. The maximum P_1 value is thus given by:

$$\max P_1 = \min\left[x(\mathcal{PL}_{2,4} \cap \mathcal{PL}_1 \cap S_U), x(\mathcal{PL}_{2,4} \cap S_1 \cap S_U)\right],$$
$$\max P_1 = \min\left[x(x_U), x(w_4)\right]. \tag{6.38}$$

If i resides on S_U, then the intersection segment of $\mathcal{PL}_{2,4}$ with S_U is broken into two segments as shown in Fig. 6.12. In that case, if we let a and b be the points given by (6.37) and (6.38) in the general case ($a = s_U$ and $b = w_4$ in the case of Fig. 6.12), then the optimization over S_U has to be conducted separately over $\overline{bi}$ from the side of region 1, and over $\overline{ia}$ from the side of region 2. In this case, i corresponds to the max P_1 point in $\overline{ia}$ and to the min P_1 point in $\overline{bi}$. Assuming the conditions of the last row in Table 6.1, this is the only case where 4 segments in total have to be checked to find the optimal D2D throughput achieving point.

The coordinates of all the points mentioned in this section, i.e. $k_1, k_2, j_2, x_1x_2, x_U, s_1, s_2, s_U, v_1, v_2, v_3, v_4, v_5, g_1, g_2, g_3, g_4, g_5$ and i when it resides on S_U are given below. Note that

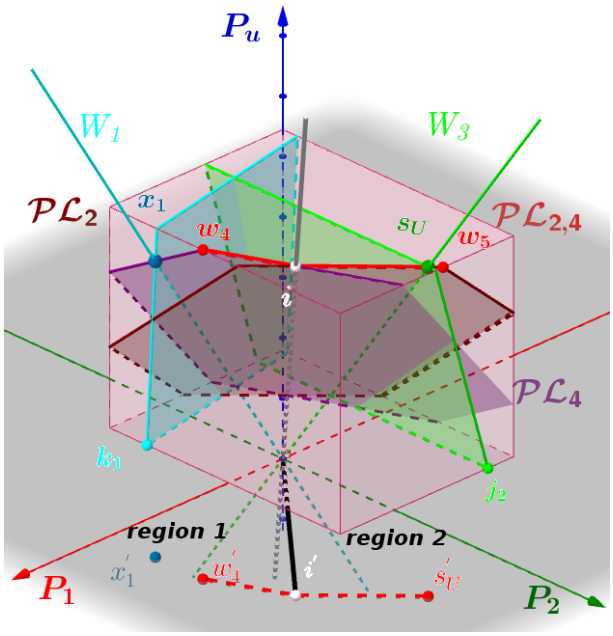


Figure 6.12 – Example of i residing over S_U: the optimization segment is broken into two: $\overline{s_U i}$ and $\overline{i w_4}$.

x_i and xl_i (resp. s_i and sl_i) have the same expressions with the difference that xl_i (resp. sl_i) is not defined outside of S_i. Moreover, x_i and s_i have strictly positive coordinates since $h_{d_1,u}h_{b,d_2} - h_{b,u}h_d > 0$ from eq. (6.8), and $h_{b,d_1}h_{d_1,u} - \eta_1 h_{b,u} > 0$ from eq. (6.29).

$$
\begin{aligned}
k_1 &= (P_{1,M}, (P_{u,m}h_{bu} + P_{1,M}h_{b,d_1})/h_{b,d_2}, P_{u,m}),\\
k_2 &= ((P_{2,M}h_{b,d_2} - P_{u,m}h_{b,u})/h_{b,d_1}, P_{2,M}, P_{u,m}),\\
j_2 &= (P_{u,m}h_{b,u}/h_{b,d_1}, P_{2,M}, P_{u,m}),\\
x_1 &= (1, \frac{h_{b,u}\eta_1 + h_{d_1,u}h_{b,d_1}}{h_{d_1,u}h_{b,d_2} - h_{b,u}h_d}, \frac{h_{b,d_2}\eta_1 + h_d h_{b,d_1}}{h_{d_1,u}h_{b,d_2} - h_{b,u}h_d})P_{1,M},\\
x_2 &= (\frac{h_{d_1,u}h_{b,d_2} - h_{b,u}h_d}{h_{b,u}\eta_1 + h_{d_1,u}h_{b,d_1}}, 1, \frac{h_{b,d_2}\eta_1 + h_d h_{b,d_1}}{h_{b,u}\eta_1 + h_{d_1,u}h_{b,d_1}})P_{2,M},\\
s_1 &= (1, \frac{h_{b,d_1}h_{d_1,u} - \eta_1 h_{b,u}}{h_{b,u}h_d}, \frac{h_{b,d_1}}{h_{b,u}})P_{1,M},\\
s_2 &= (\frac{h_{b,u}h_d}{h_{b,d_1}h_{d_1,u} - \eta_1 h_{b,u}}, 1, \frac{h_{b,d_1}h_d}{h_{b,d_1}h_{d_1,u} - \eta_1 h_{b,u}})P_{2,M},\\
s_u &= (\frac{h_{b,u}}{h_{b,d_1}}, \frac{h_{b,d_1}h_{d_1,u} - \eta_1 h_{b,u}}{h_{b,d_1}h_d}, 1)P_{u,M},\\
x_u &= (\frac{h_{d_1,u}h_{b,d_2} - h_{b,u}h_d}{h_{b,d_2}\eta_1 + h_d h_{b,d_1}}, \frac{h_{b,u}\eta_1 + h_{d_1,u}h_{b,d_1}}{h_{b,d_2}\eta_1 + h_d h_{b,d_1}}, 1)P_{u,M},\\
i &= (\frac{h_d h_{d_2,u} - \eta_2 h_{d_1,u}}{h_d^2 - \eta_1\eta_2}, \frac{h_d h_{d_1,u} - \eta_1 h_{d_2,u}}{h_d^2 - \eta_1\eta_2}, 1)P_{u,M}.
\end{aligned}
$$

The w_i family is obtained by combining v_i and g_i.

$$v_1 = (P_{1,M}, (P_{u,m}h_{d_1,u} - P_{1,M}\eta_1)\frac{1}{h_d}, P_{u,m}),$$
$$v_2 = ((P_{u,m}h_{d_1,u} - P_{2,M}h_d)\frac{1}{\eta_1}, P_{2,M}, P_{u,m}),$$
$$v_3 = (P_{1,M}, P_{2,M}, (P_{1,M}\eta_1 + P_{2,M}h_d)\frac{1}{h_{d_1,u}}),$$
$$v_4 = (P_{1,M}, (P_{u,M}h_{d_1,u} - P_{1,M}\eta_1)\frac{1}{h_d}, P_{u,M}),$$
$$v_5 = ((P_{u,M}h_{d_1,u} - P_{2,M}h_d)\frac{1}{\eta_1}, P_{2,M}, P_{u,M}),$$
$$g_1 = (P_{1,M}, (P_{u,m}h_{d_2,u} - P_{1,M}h_d)\frac{1}{\eta_2}, P_{u,m}),$$
$$g_2 = ((P_{u,m}h_{d_2,u} - P_{2,M}\eta_2)\frac{1}{h_d}, P_{2,M}, P_{u,m}),$$
$$g_3 = (P_{1,M}, P_{2,M}, (P_{1,M}h_d + P_{2,M}\eta_2)\frac{1}{h_{d_2,u}}),$$
$$g_4 = (P_{1,M}, (P_{u,M}h_{d_2,u} - P_{1,M}h_d)\frac{1}{\eta_2}, P_{u,M}),$$
$$g_5 = ((P_{u,M}h_{d_2,u} - P_{2,M}\eta_2)\frac{1}{h_d}, P_{2,M}, P_{u,M}).$$

6.8.4 D2D Throughput Optimization

At last, given the segments locations and endpoints, the analytical power optimization can be conducted. The mathematical formulation varies according to the side the segment is included in.

6.8.4.1 Side S_1

The optimization variable is P_1 and the problem formulation is the following:

$$P_1^* = \arg\max_{P_1} \left(B\log_2(1 + \frac{P_1 h_d}{P_2\eta_2 + \sigma^2}) + B\log_2(1 + \frac{P_2 h_d}{P_1\eta_1 + \sigma^2}) \right),$$

such that

$$P_1 \in [\min P_1, \max P_1]$$
$$P_2 = P_{2,M}$$

Taking the derivative of $F(P_1) = R_{D2D}(P_1, P_{2,M})$ with respect to P_1, we get:

$$\frac{\partial F}{\partial P_1}\frac{\ln 2}{B} = \frac{h_d}{P_1 h_d + P_{2,M}\eta_2 + \sigma^2} + \frac{-\eta_1 P_{2,M} h_d}{(P_1\eta_1 + \sigma^2)(P_1\eta_1 + P_{2,M}h_d + \sigma^2)}$$

The sign of $\partial F/\partial P_1$ is equal to the sign of the following second-degree polynomial of P_1:

$$P_1^2 \underbrace{\eta_1^2}_{a} + P_1 \underbrace{2\eta_1\sigma^2}_{b} + \underbrace{P_{2,M}(h_d - \eta_1)\sigma^2 - P_{2,M}^2\eta_2\eta_1 + \sigma^4}_{c}$$

If $\Delta = b^2 - 4ac < 0$, the second-degree polynomial is positive, hence the throughput is increasing with P_1, and P_1^* is obtained by setting P_1 to $\max P_1$.
If $\Delta > 0$, the polynomial is negative inside the solutions interval, and positive elsewhere. The solutions are: $sol_1 = (-b - \sqrt{\Delta})/2a, sol_2 = (-b + \sqrt{\Delta})/2a$. Therefore, the throughput is decreasing between sol_1 and sol_2, then increasing for $P_1 > sol_2$. Since $sol_1 < 0$, three cases are identified depending on the location of sol_2 with respect to $\min P_1$ and $\max P_1$:

- $sol_2 < \min P_1$: the throughput increases with $P_1 \Rightarrow P_1^* = \max P_1$.
- $sol_2 > \max P_1$: the throughput decreases with $P_1 \Rightarrow P_1^* = \min P_1$.

- $sol_2 \in [\min P_1, \max P_1]$: as shown in the variation table of Fig. 6.13, the throughput is decreasing between $\min P_1$ and sol_2, and increasing between sol_2 and $\max P_1$. Therefore, we obtain $P_1^* = \arg\max[F(\min P_1), F(\max P_1)]$.

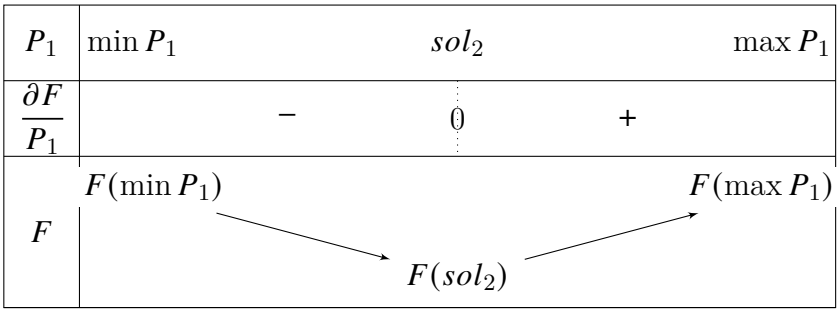


Figure 6.13 – R_{D2D} variation table when $sol_2 \in [\min P_1, \max P_1]$.

As a conclusion, no matter if Δ is positive of negative, it is sufficient to test which of $\min P_1$ or $\max P_1$ delivers the best throughput and then select the corresponding segment endpoint. The coordinates of the endpoint form the optimal triplet (P_1^*, P_2^*, P_u^*) maximizing the D2D throughput over the side S_1.

6.8.4.2 Side S_2

Following the same reasoning as for S_1 (with the only difference that the optimization variable is now P_2 instead of P_1, and $P_1 = P_{1,M}$), the same conclusion is reached, i.e. the maximum D2D throughput is delivered by the points corresponding either to $\min P_2$ or to $\max P_2$.

6.8.4.3 Side S_U

Since the intersection point i is accounted for in the maximum and minimum values of P_1 for each intersection segment, the optimization can thus be conducted over each segment independently.
The D2D throughput maximization problem over the intersection segment of $\mathcal{PL}_2$ with S_U can be written as follows:

$$P_1^* = \arg\max_{P_1} \left(F(P_1, P_2) = B\log_2(1 + \frac{P_1 h_d}{P_2\eta_2 + \sigma^2}) + B\log_2(1 + \frac{P_2 h_d}{P_1\eta_1 + \sigma^2}) \right)$$

such that

$$P_{u,M} = \frac{P_1\eta_1 + P_2 h_d}{h_{d_1,u}},$$
$$P_1 \in \mathbb{U} = [\min P_1, \max P_1],$$
$$P_u = P_{u,M}.$$

Replacing P_2 by $(P_{u,M}h_{d_1,u} - P_1\eta_1)/h_d$ in $F(P_1, P_2)$, we get:

$$F(P_1) = B\log_2\left(1 + \frac{P_1 h_d^2}{(P_{u,M}h_{d_1,u} - P_1\eta_1)\eta_2 + h_d\sigma^2}\right) + B\log_2\left(1 + \frac{P_{u,M}h_{d_1,u} - P_1\eta_1}{P_1\eta_1 + \sigma^2}\right).$$

Since $P_2 > 0$, we must have $P_1 < P_{u,M} h_{d_1,u}/\eta_1$, which adds to the constraint of $\max P_1$. In other words, the new maximum allowed value for P_1 is now given by:

$$\max P_1 = \min \left[x(x_U), x(w_4), P_{u,M} h_{d_1,u}/\eta_1\right]$$

Taking the derivative of F with respect to P_1 leads to:

$$\frac{\ln(2)}{B}\frac{\partial F}{\partial P_1} = \frac{h_d^2(h_{d_1,u}\eta_2 P_{u,M} + \sigma^2 h_d)/[P_{u,M} h_{d_1,u}\eta_2 - P_1\eta_1\eta_2 + h_d\sigma^2]}{(P_{u,M} h_{d_1,u}\eta_2 - P_1(\eta_1\eta_2 - h_d^2) + h_d\sigma^2)} - \frac{\eta_1}{(P_1\eta_1 + \sigma^2)}.$$

Since $P_1 \in \mathbb{U}$, it can be easily verified that both denominators are positive; therefore, only the numerator is needed to evaluate the sign of $\partial F/\partial P_1$:

$$\begin{aligned}\operatorname{sgn}\frac{\partial F}{\partial P_1} = \operatorname{sgn}[&h_d^2(h_{d_1,u}\eta_2 P_{u,M} + \sigma^2 h_d)(P_1\eta_1 + \sigma^2) \\ &- \eta_1(P_{u,M} h_{d_1,u}\eta_2 - P_1\eta_1\eta_2 + h_d\sigma^2)(P_{u,M} h_{d_1,u}\eta_2 - P_1(\eta_1\eta_2 - h_d^2) + h_d\sigma^2)].\end{aligned}$$

After some simplifications and re-arrangements, the sign of $\partial F/\partial P_1$ can be written as the sign of a second-degree polynomial of P_1 of the form $AP_1^2 + BP_1 + C$ with:

$$\begin{aligned}A = -(\eta_1\eta_2 - h_d^2)\eta_1^2\eta_2; \quad B = 2\eta_1^2\eta_2(P_{u,M} h_{d_1,u}\eta_2 + \sigma^2 h_d); \\ C = -P_{u,M}^2 h_{d_1,u}^2\eta_2^2\eta_1 + P_{u,M}\sigma^2 h_d h_{d_1,u}\eta_2(h_d - 2\eta_1) + \sigma^4 h_d^2(h_d - \eta_1).\end{aligned}$$

Given one of the polynomial roots $sol_1 = (-B - \sqrt{B^2 - 4AC})/2A$, we show next that P_1^* is either given by $\min P_1$, $\max P_1$, or sol_1 (when it is included in the interval $[\min P_1, \max P_1]$), according to the value delivering the highest throughput.

Proof. Consider the sign of the polynomial's discriminant $\Delta = B^2 - 4AC$. If $\Delta < 0$: $\operatorname{sgn}\partial F/\partial P_1 = \operatorname{sgn}(h_d^2 - \eta_1\eta_2)$.

- If $h_d^2 > \eta_1\eta_2, F$ is increasing with $P_1 \Rightarrow$ Set P_1^* to $\max P_1$
- If $h_d^2 < \eta_1\eta_2, F$ is decreasing with $P_1 \Rightarrow$ Set P_1^* to $\min P_1$

However, if $\Delta > 0$, then we have the two solutions $sol_1 = (-B - \sqrt{\Delta})/2A$ and $sol_2 = (-B + \sqrt{\Delta})/2A$, with the variation tables (Figs. 6.14 and 6.15) depending on the sign of $h_d^2 - \eta_1\eta_2$.

- If $h_d^2 > \eta_1\eta_2 \Rightarrow sol_1 < sol_2$ and $sol_1 < 0$. But not much can be said about the sign of sol_2 and how it compares to $\min P_1$ and $\max P_1$.

P_1	$-\infty$		sol_1		sol_2		$+\infty$
$\frac{\partial F}{\partial P_1}$		$+$	0	$-$	0	$+$	
F	$-\infty$	↗	$F(sol_1)$	↘	$F(sol_2)$	↗	$+\infty$

Figure 6.14 – Variation table for $h_d^2 > \eta_1\eta_2$.

However, we note that the right side of the variation table (where $P_1 > sol_1$) is similar to the variation table in Fig. 6.13. Therefore, we conclude that:

$$P_1^* = \arg\max[F(\min P_1), F(\max P_1)].$$

- if $h_d^2 < \eta_1\eta_2 \Rightarrow sol_2 < sol_1$ and $sol_1 > 0$.

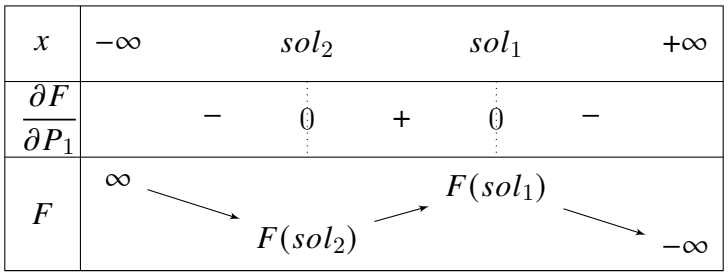


Figure 6.15 – Variation table for $h_d^2 < \eta_1\eta_2$.

Since sol_1 is a local maximum, $F(sol_1) > F(P_1), \forall P_1 > sol_2$. Then, the only values of P_1 which might give a better throughput than sol_1 are those at the left of sol_2. We can distinguish the following three cases:

- if $\max P_1 < sol_1$, set P_1^* to $\max P_1$.
- if $sol_1 < \min P_1$, set P_1^* to $\min P_1$.
- if $sol_1 \in [\min P_1, \max P_1]$, then:
  - if $\min P_1 > sol_2$, set P_1^* to sol_1.
  - if $\min P_1 < sol_2$, set $P_1^* = \arg\max[F(\min P_1), F(sol_1)]$.

To sum up, in the optimization over the intersection segment of $\mathcal{PL}_2$ with S_U, all the possible channel conditions lead at some point to choosing P_1^* from the values $\min P_1$, $\max P_1$, and sol_1 (when it is included in the interval $\mathbb{U}$) according to the one delivering the highest throughput. □

Regarding the optimization over the intersection segment of $\mathcal{PL}_4$ with S_U, the same steps are followed to determine the optimal value of P_1: we start by writing the expression of $F(P_1)$ by replacing P_2 in $F(P_1, P_2)$ with $(P_{u,M}h_{d_2,u} - P_1h_d)/\eta_2)$. Then, the study of the sign of $\partial F/\partial P_1$ turns into the study of the sign of another second-degree polynomial $A'P_1^2 + B'P_1 + C'$ with:

$$A' = (\eta_1\eta_2 - h_d^2)\eta_1; \quad B' = 2\eta_1(P_{u,M}h_{d_2,u}h_d + \sigma^2\eta_2);$$
$$C' = -P_{u,M}^2h_{d_2,u}^2\eta_1 - \sigma^2\eta_1h_{d_2,u}P_{u,M} + \sigma^4(\eta_2 - h_d).$$

Also, following the different channel conditions concerning $\text{sgn}(\eta_1\eta_2 - h_d^2)$, and considering all the possible relative positions between $\max P_1, \min P_1$, and sol_1', the same result as previously is obtained, which can be cast as:

$$P_1^* = \arg\max[F(\min P_1), F(\max P_1), F(sol_1')].$$

As a conclusion, the optimization over the sides S_1 and S_2 resides in selecting the corresponding endpoint achieving the highest throughput. On the side S_U, a maximum of three additional points (i, sol_1, sol_1') may need to be considered to get the highest D2D throughput.

6.8.5 Summary of the Power Allocation Procedure and Extension to the Second Decoding Order

In this section (sec. 6.8), the geometrical representation of the FD-SIC PA problem allowing for a drastic reduction of the search space size was described. It was shown that the initial search volume in section 6.8.1 can be reduced to a set of intersection segments (section 6.8.2) from which a subset is selected (6.8.3). These segments search spaces are then further reduced to become a finite set of points (sections 6.8.3.3, 6.8.4). In the worst case scenario, the original PA problem, which had $2^{12}-1$ variants, is converted into the search for the maximum of a throughput list of seven elements: two elements from S_1, two from S_2 and three additional elements from S_U (w_4 is a common endpoint to S_1 and S_U, and w_5 is common to S_2 and S_U). The global PA procedure to determine the optimal D2D rate for the first decoding order of FD-SIC is summarized in algorithm 6.2.
Regarding the resolution for the second decoding order, the PA procedure itself is unchanged, but the changes in $\boldsymbol{PMC_1}$ and $\boldsymbol{PMC_3}$ lead to some modifications. Here is the list:

- Modification in the expressions of $\boldsymbol{PMC_1}$ and $\boldsymbol{PMC_3}$:

$$P_u h_{b,u} < P_1 h_{b,d_1} - P_2 h_{b,d_2} \qquad (\boldsymbol{PMC_1})$$
$$P_u h_{b,u} < P_2 h_{b,d_2} \qquad (\boldsymbol{PMC_3})$$

- The necessary and sufficient conditions (6.31), (6.33) and (6.35) become:

$$h_{d_1,u} h_{b,d_2} - h_d h_{b,u} > 2\eta_1 h_{b,u} \frac{h_{b,d_2}}{h_{b,d_1}} \qquad (6.31)$$

$$h_{d_2,u} h_{b,d_2} - \eta_2 h_{b,u} > 2 h_{b,u} h_d \frac{h_{b,d_2}}{h_{b,d_1}} \qquad (6.33)$$

$$2P_{u,m} \frac{h_{b,u}}{h_{b,d_1}} < P_{1,M} \quad \&\& \quad P_{u,m} \frac{h_{b,u}}{h_{b,d_2}} < P_{2,M} \qquad (6.35)$$

- Concerning section 6.8.3.3, the roles of $\boldsymbol{PMC_1}$ and $\boldsymbol{PMC_3}$ are interchanged concerning the settlement of the segment endpoints.
- The three non-occuring (x_i, s_i) pairs of section 6.8.3.2 become: $(x_1, s_2), (x_1, s_U), (x_U, s_2)$.

Sections 6.8.3.1 and 6.8.4 are kept unchanged because building $\boldsymbol{PMC_{2,4}}$ is independent of $\boldsymbol{PMC_1}$ and $\boldsymbol{PMC_3}$, and given the endpoints of the segments subset, the optimization of section 6.8.4 is not affected by the change in $\boldsymbol{PMC_1}$ and $\boldsymbol{PMC_3}$.

6.9 Channel Allocation

In this section, the procedure for optimal channel allocation to D2D devices is conducted. Recalling that the D2D system is underlaying a pre-established CU network, D2D channel allocation is equivalently referred to as D2D-CU pairing.

Having determined the analytical PA solutions for all the transmission scenarios, their resolution cost is a constant time operation. Therefore, filling the D2D rate tables

Algorithm 6.2 Optimal PA procedure for FD-SIC

input : $P_{1,M}$, $P_{2,M}$, $P_{u,M}$, $P_{u,m}$, $h_{b,u}$, h_d, η_1, η_2, $h_{d1,u}$, $h_{d_2,u}$, $h_{d_1,b}$, $h_{d_2,b}$
Result: Optimal triplet (P_1^*, P_2^*, P_u^*).
if (6.31) $\wedge$ (6.33) $\wedge$ (6.35) **then**
  Test $\Omega, \xi, \gamma, \tau$ and build $PMC_{2,4}$
  Execute Algorithm 6.1 to determine x_i and s_j
  Follow Table 6.1 to keep the necessary segments
  Compute R_{D2D} for the edges of each segment
  Keep the point providing the highest throughput.
else
  Empty search space, no solution
end

$\mathcal{R}_{D2D}^{FD-NoSIC}, \mathcal{R}_{D2D}^{HD-NoSIC}, \mathcal{R}_{D2D}^{HD-SIC}$, and $\mathcal{R}_{D2D}^{FD-SIC}$ for every D2D-CU pair is accomplished with a complexity in $O(KD)$. In the case of FD-SIC, the channel links, required CU rate and transmit power limits of a D2D n and a CU u_i may be such that one of the conditions (6.31), (6.33), (6.35) is not valid. If this is the case for both decoding orders, then the PA of FD-SIC reverts to that of FD-NoSIC to fill the element $\mathcal{R}_{D2D}^{FD-SIC}(n,i)$ as explained in the end of section 6.7. Also, if both decoding orders are possible for this combination, $\mathcal{R}_{D2D}^{FD-SIC}(n,i)$ is filled with the highest rate among the two possible orders. When filling matrix $\mathcal{R}_{D2D}^{HD-SIC}$, and as explained in section 6.5, HD-SIC reverts to HD-NoSIC in any of the two half-time slots, when conditions (6.6) or (6.8) are not valid. Given these rate tables, the optimal channel allocation tables $O_{FD-NoSIC}^*, O_{HD-NoSIC}^*, O_{FD-SIC}^*$, and O_{HD-SIC}^* corresponding to every transmission scenario are obtained by solving the channel assignment problem in a way to maximize the total D2D throughput. This problem takes the generic formulation given by:

$$O^* = \underset{(i,j)\in[\![1,D]\!]\times[\![1,K]\!]}{\arg\max} \left(\sum_{i=1}^{D}\sum_{j=1}^{K} \mathcal{R}_{D2D}(i,j) \times o(i,j)\right)$$

such that the constraints of (6.1) are verified

This assignment problem is efficiently solved by the Kuhn-Munkres (KM) algorithm [144], also called the Hungarian method, with a complexity of $O(D^2K)$ [145]. Note that the global resource allocation complexity is now dominated by the channel assignment after the important PA complexity reduction. The Hungarian method can be directly applied in our study to yield the optimal channel assignment by rewriting the problem as a minimization of the opposite objective function $(-\mathcal{R}_{D2D})$. The required input for the KM algorithm is therefore the opposite of the rate tables of each transmission scenario. As a conclusion, the optimal PA procedures allowed for an efficient filling of the rate tables which are then fed to the KM solver. This delivers the global optimal solution of the joint channel and power allocation problem formulated in section 6.2.1.

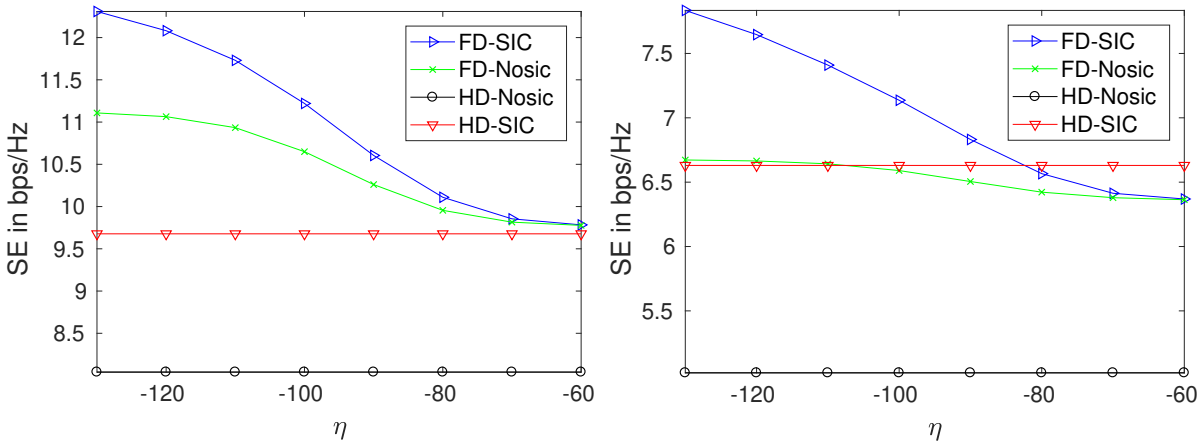
6.10 Numerical Results

In our simulation setup, the BS is positioned at the center of a hexagonal cell with an outermost radius of 300 m. The D2D users and the CUs are randomly located within the

cell. The distance between the D2D users of every pair is below a maximum value d_{max}. The propagation model includes large scale fading with a path loss exponent $\alpha = 3.76$, and an 8 dB zero mean lognormal shadowing. The maximum transmit power of the devices and CU is 24 dBm. The system bandwidth is 20 MHz, divided into $N = 64$ channels, leading to a UL bandwidth of $B = 312.5$ kHz, with a noise power of -119 dBm. The minimum required rate $R_{u,min}$ is the same for all the CU users, and the SI cancellation factor η is the same for all D2D pairs, its value being varied between -130 and -80 dB. The results are averaged over 1000 different realizations of the devices and CU positions. First, we present the simulation results for a single D2D-CU pair, in order to gain insights on the characteristics of the mutual SIC technique for a D2D application. Then, we present the results for a fully fledged cellular network with K CUs and D D2D pairs.

6.10.1 Results for a Single D2D-CU System

Hereinafter, "Global" figures present the SE results averaged over all the simulated D2D-CU triplets, including both SIC success and failure cases (in case of failure, SIC algorithms revert to their NoSIC counterparts). On the other hand, the "SIC-only" figures present the results averaged over the cases of FD-SIC success. Throughout this section, a value of $d_{max} = 100$ m is considered.

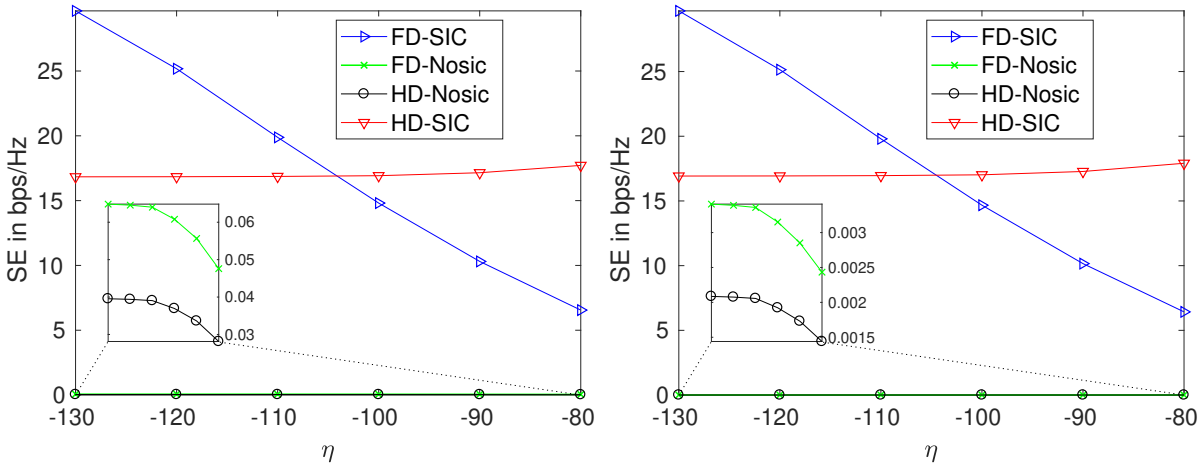


(a) $R_{u,min} = 3$ bps/Hz (b) $R_{u,min} = 7$ bps/Hz

Figure 6.16 – Global D2D spectral efficiency as a function of η.

The evolution of the average D2D SE with η is shown in Figs. 6.16a and 6.16b, for a minimum target CU rate $R_{u,min} = 3$ and 7 bps/Hz respectively. As expected, the improvement of the SI cancellation capabilities increases the performance of FD algorithms: for instance, the D2D SE of FD-NoSIC in Fig. 6.16a falls from 11.1 to 9.9 bps/Hz when η varies between -130 and -80 dB. Also, the increase of QoS requirement impacts FD and HD algorithms by limiting the achieved SE (in FD-NoSIC, for $\eta = -80$ dB and $R_{u,min} = 7$ bps/Hz, $R_{D2D} = 6.4$ bps/Hz), and also by reducing the range of variation of FD algorithms with η (for FD-NoSIC, $\Delta R_{D2D} = R_{D2D}(-130dB) - R_{D2D}(-80dB) = 0.3$ bps for $R_{u,min} = 7$ bps/Hz, compared to 1.15 bps/Hz for $R_{u,min} = 3$ bps/Hz). On the other hand, the HD curves are independent of η since they do not suffer from SI. Note that, in the case of NoSIC, FD always outperforms HD since, by shutting down the power of the adequate device, it can revert to the half time slot in HD delivering the best throughput and then

extend it to the other half. This is clearly not the case for SIC scenarios where HD-SIC may outperform FD-SIC as shown in Figs. 6.17a and 6.17b. More on the reasons behind this behavior later on.



(a) $R_{u,min} = 3$ bps/Hz (b) $R_{u,min} = 7$ bps/Hz

Figure 6.17 – SIC-only D2D rates as a function of η.

It can be remarked from Figs. 6.16a and 6.16b that the improvement of FD-SIC with respect to FD-NoSIC is virtually the same, independently of the required QoS (around 1.2 bps/Hz for $\eta = -130$ dB). This is even clearer in Fig. 6.17a and 6.17b, which are nearly identical despite the different required rates. The reason behind this behavior is quite simple: thanks to the mutual SIC, the D2D rates do not suffer from the higher CU interference levels since interference is canceled. To get further insights, attention is drawn to Fig. 6.2 where the optimal HD-SIC PA is depicted: whether $P_{1,M}$ is such that the solution is in the order of $P^1_{1,M}, P^2_{1,M}$ or $P^3_{1,M}$, increasing $R_{u,min}$ simply raises the horizontal $P_{u,m}$ line. Since the optimal PA is obtained from the intersection of the blue segments with the line $P_1 = \min(P_{1,M}, P_{u,M}/A)$, the abscissa of the optimal PA (P_1) is not affected in any ways by $P_{u,m}$, therefore the D2D rate is unchanged. This is the same for the case of FD-SIC (with the $P_{u,m}$ line becoming the horizontal $P_{u,m}$ plane), leading to the same independence of the D2D rate from $R_{u,min}$.

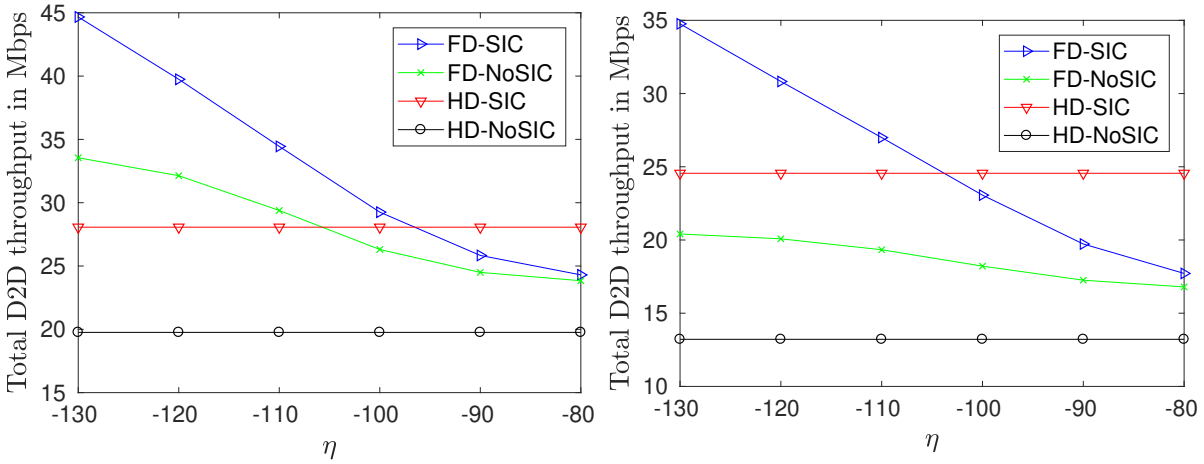
More importantly, Fig. 6.17a and 6.17b show that when the SIC procedure is applicable, legacy D2D NoSIC algorithms perform poorly. In other terms, for scenarios where CU interference would severely hinder D2D communication rendering its use futile (e.g. case of close CU interferer to the D2D pair), implementing the SIC procedure completely changes the situation by taking advantage of high CU interference levels for a better cancellation. Therefore, conducting the mutual SIC procedure expands the field of relevant D2D applications to broader channel configurations and user placement scenarios. To sum up, we say that the interference cancellation strategy of NOMA mutual SIC complements the interference avoidance approach of standard D2D applications.

6.10.2 Results for a complete cellular system with K CUs and D D2Ds

In this section, we present the results of the proposed optimal PAs and D2D-CU pairing for a complete cellular system. Unless specified otherwise, $R_{u,min}$ is set to 1.5 Mbps, K is

set to 20 CUs, d_{max} = 100 m, and D = 5 D2D pairs.

Figure 6.18 presents the total D2D throughput as a function of η, for two different values of $R_{u,min}$. It is observed that mutual SIC-enabled schemes outperform their counterpart No-SIC schemes for both HD and FD transmission scenarios. In other terms, the SINR advantages of the SIC operation outweigh the burden incurred by the additional PMCs on the solution to the PA problem. Indeed, a 41 % rate increase is observed in Fig. 6.18a between HD-SIC and HD-NoSIC (going from 19.8 Mbps to 28.1 Mbps). The throughput enhancements due to mutual SIC for the case of FD transmission vary between a 2 % increase for $\eta = -80$ dB, to 33 % increase for $\eta = -130$ dB. The performance gains of FD-SIC with respect to FD-NoSIC increase with the SI cancellation capabilities of the devices because of two reasons: on the one hand, the decrease of η relaxes the constraints (6.29) and (6.30), thereby increasing the number of D2D-CU pairs that benefit from FD-SIC (from an average of 0.36 FD-SIC D2D pairs for $\eta = -80$ dB to 1.92 pairs for $\eta = -130$ dB, with $R_{u,min}$ = 1.5 Mbps). On the other hand, the decrease of η reduces the interference terms in the D2D throughput expression, which translates into a higher achieved throughput.



(a) $R_{u,min}$ = 1.5 Mbps

(b) $R_{u,min}$ = 3 Mbps

Figure 6.18 – Total D2D throughput as a function of η for K = 20 CUs, D = 5 D2D pairs, and d_{max} = 100 m.

As expected, when comparing the performance for different required CU rates between Figs. 6.18a and 6.18b, the increase of R_u from 1.5 Mbps to 3 Mbps decreases the achieved D2D throughput for all proposed methods. However, the percentage gain in the performance of SIC procedures with respect to NoSIC increases from 41 % to 86 % for the HD case, and from 33 % to 70 % for the FD case (for $\eta = -130$ dB). The reason behind this gain increase is that NoSIC algorithms are highly affected by the value of P_u ($\geq P_{u,m}$) since they suffer from its interference, which is not the case of SIC techniques as discussed earlier. In fact, even though the total number of FD-SIC enabled D2D-CU pairs decreases with the harsher mutual SIC constraints of increasing $R_{u,min}$ (from an average of 1.6 pairs for $R_{u,min}$ = 1.5 Mbps to 1.4 pairs for $R_{u,min}$ = 3 Mbps, with $\eta = -90$ dB), the Munkres allocation yields an increasing number of selected D2D-CU pairs achieving FD-SIC (or HD-SIC) with $R_{u,min}$ (from an average of 0.8 pairs for $R_{u,min}$ = 1.5 Mbps to an average of 1.24 pairs for $R_{u,min}$ = 3 Mbps, with $\eta = -90$ dB). This corroborates the idea that the throughput decrease of No-SIC techniques with $R_{u,min}$ is more important than

that of SIC techniques, to a point where the contribution of mutual SIC techniques in maximizing the throughput is more prominent when $R_{u,min}$ increases. This is verified by comparing the percentage decrease of D2D throughput for every algorithm when moving from R_u = 1.5 Mbps to R_u = 3 Mbps: a decrease of 39 %, 33 %, 22 %, and 13 % is observed for the algorithms FD-NoSIC, HD-NoSIC, FD-SIC, HD-SIC respectively. The greater decrease of FD-NoSIC performance compared to HD-NoSIC justifies the shift of the intersection point between FD-SIC and HD-SIC to the left when $R_{u,min}$ increases. Indeed, as explained in section 6.9, FD-SIC and HD-SIC are applied when possible, on top of FD-NoSIC and HD-NoSIC respectively. If the performance gap between FD-NoSIC and HD-NoSIC diminishes, HD-SIC will outperform FD-SIC over a broader span of η values before FD-SIC eventually catches up and surpasses HD-SIC for smaller η values (i.e. for better SI cancellation capabilities of the devices).

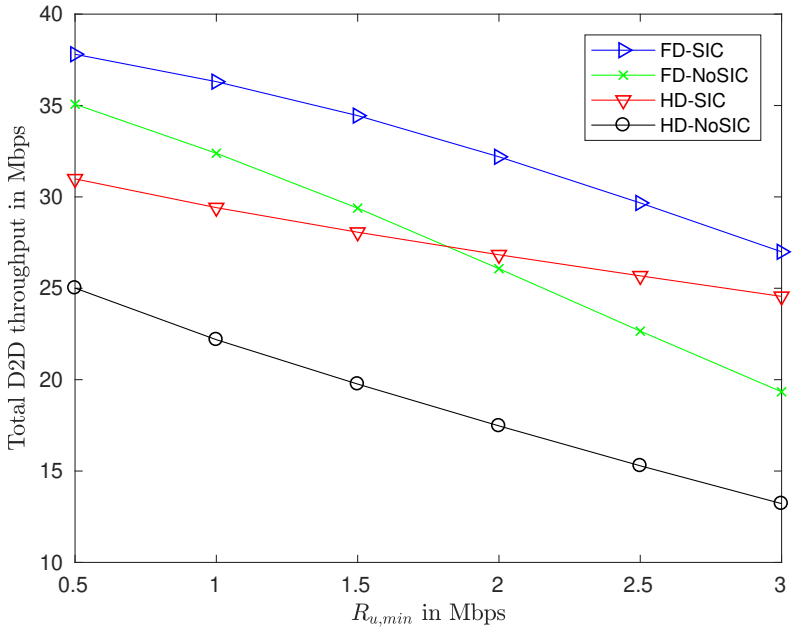


Figure 6.19 – Total D2D throughput as a function of $R_{u,min}$ for $\eta = -110$ dB.

This evolution of FD-SIC and HD-SIC can also be observed from another perspective in Fig. 6.19, where the total D2D throughput is presented as a function of the CU required rate. In the conditions of Fig. 6.19, the gap between FD-NoSIC and HD-NoSIC is large enough so that no intersection occurs between FD-SIC and HD-SIC. However, it can still be observed that the gap between FD-SIC and HD-SIC reduces as the CU required rate increases.

In Fig. 6.20, the variation of the total D2D throughput is presented as a function of the D2D maximum user distance d_{max}. The increase of d_{max} leads to a significant decrease in the performance of all proposed methods since h_d, the channel gain of the direct link between d_1 and d_2, is reduced on average. However, this increase of d_{max} is accompanied by a greater percentage increase in performance due to mutual SIC for FD and HD transmission scenarios, with respect to No-SIC scenarios. Indeed, FD-SIC achieves a D2D throughput 128 % higher than FD-NoSIC for d_{max} = 100 m, compared to the 81 % increase achieved for d_{max} = 20 m. This is due to having more FD-SIC enabled D2D-CU pairs when distancing the D2D users further apart from one another, since an average of 1.96 pairs successfully apply FD-SIC for d_{max} = 20 m as opposed

to 3.33 pairs for d_{max} = 100 m. The reason behind this increase is the decrease in h_d which relaxes the sufficient conditions (6.31) and (6.33), thereby enabling more FD-SIC cases. This highlights once again the complementarity between D2D and mutual SIC: although increasing D2D distances would usually disqualify classical D2D application, the application of mutual SIC provides a renewed interest in D2D communication.

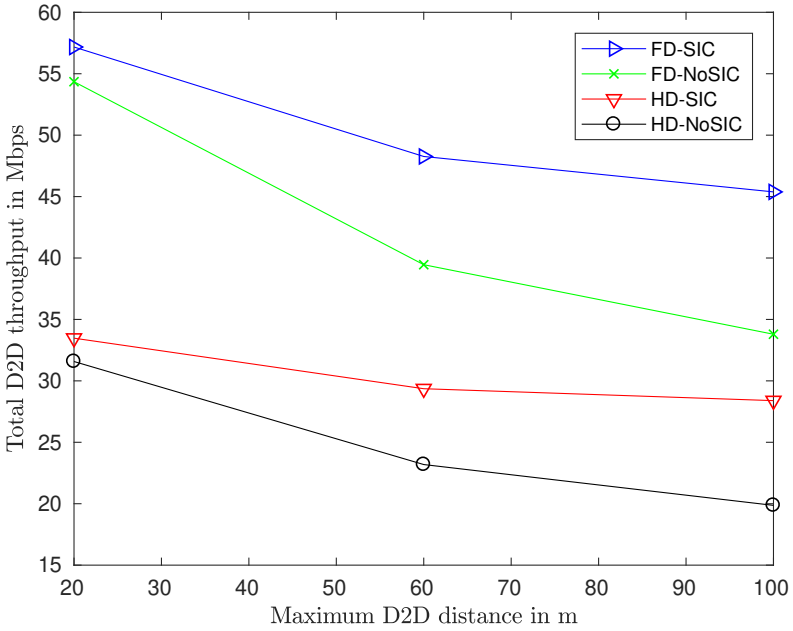


Figure 6.20 – Total D2D throughput as a function of d_{max} for $\eta = -130$ dB.

Fig. 6.21 presents the evolution of the D2D throughput as a function of the number of CUs in the cell. Although the channel properties of the D2D users (i.e. h_d, $h_{d_1,u}$ and $h_{d_2,u}$) are unchanged, the total D2D throughput of all techniques benefits from the additional diversity provided by the greater number of CU users. This also favors the FD-SIC enabled pairs, as their average number grows from 1.98 for K = 20 to 2.46 for K = 50. We can therefore conclude that the important performance gain achieved by SIC methods, with respect to No-SIC methods, can be obtained without requiring the implementation of SIC at all D2D and CU receivers. Indeed, generally only 2 or 3 triplets need to perform SIC which is enough to boost the D2D system capacity, while the others can settle for the simple classical No-SIC receivers. Therefore, the additional complexity is localised at the level of the users performing SIC for which the major throughput increase is worth the incurred SIC complexity. Finally, the total and average throughput variations are presented in Fig. 6.22 as a function of the number of D2D pairs in the system, for a fixed value of K = 50. In Fig. 6.22a, the average throughput per D2D pair is shown to slightly decrease with the increasing number of D2D pairs. In a sense, this is the dual of the behavior observed in Fig. 6.21, since the ratio K/D decreases with D and thus the system diversity – in terms of the average number of possible CU channel choices for every D2D pair to be collocated on – decreases, thus reducing the achievable throughput per D2D pair. Nonetheless, the total throughput follows a quasi linear progression with the number of D2D pairs because the additional D2D pairs are allocated on orthogonal channels, therefore each D2D pair can be associated more or less to an additional D2D rate unit. Figures 6.21 and 6.22 indicate that, for a fixed number of $D2D$ users or CUs, the effect of the proportion K/D on the average D2D throughput per D2D pair is rather

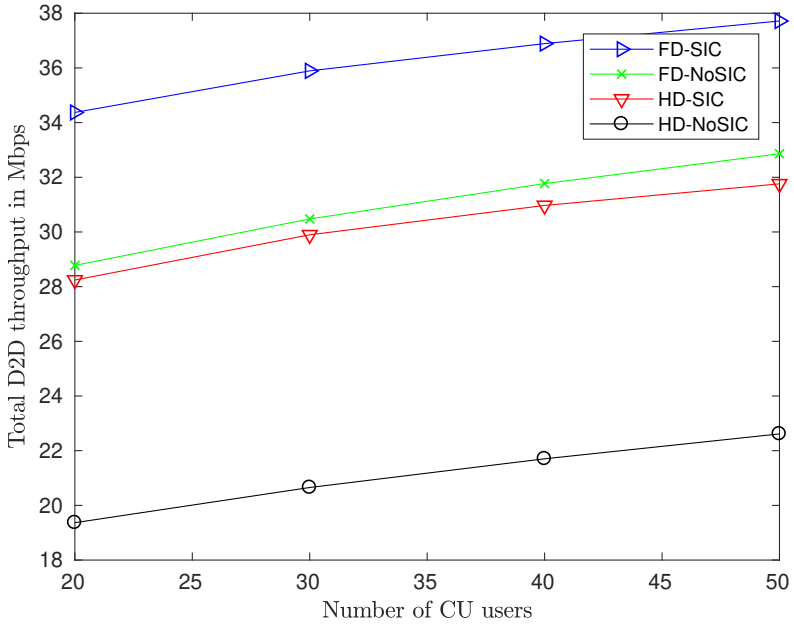
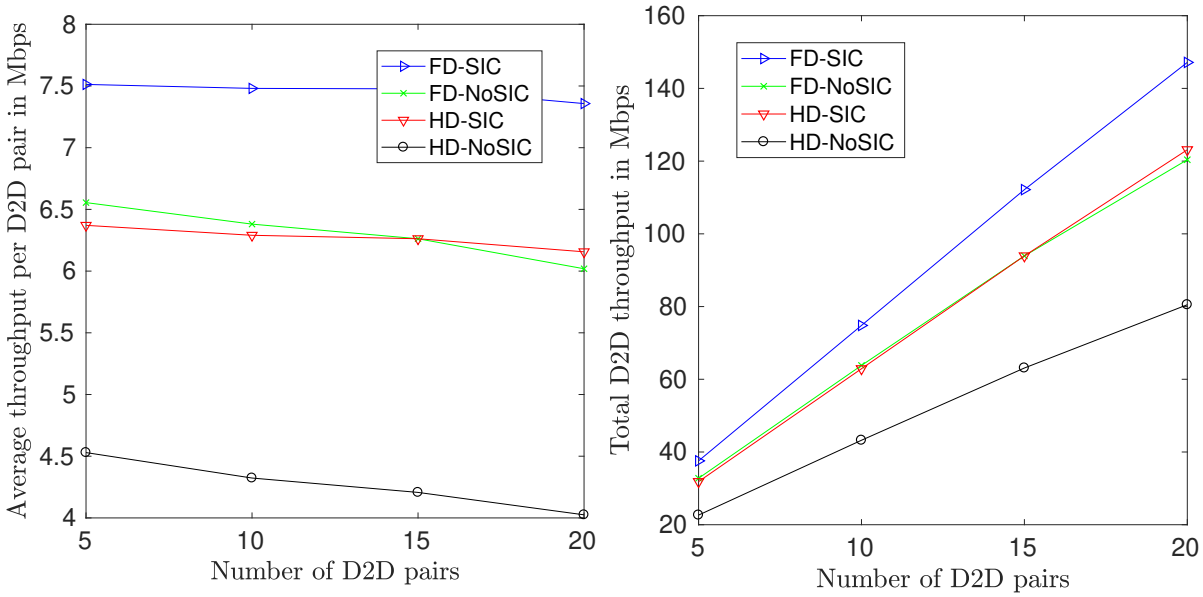


Figure 6.21 – Total D2D throughput as a function of K for $\eta = -110$ dB.



(a) Average D2D throughput

(b) Total D2D throughput

Figure 6.22 – Total and average D2D throughput as a function of the number of D2D pairs for $K = 50$ CUs and $\eta = -110$ dB.

limited. The most dominant factors remain the distance between D2D users, the SI cancellation capabilities of the receivers (for FD-SIC), and the required CU rate.

6.11 Conclusion

In this chapter, the use of NOMA with mutual SIC was proposed for the first time between cellular users and FD-D2D devices underlaying the cellular channels. The neces-

sary and sufficient conditions for applying FD-SIC were derived and a highly efficient PA procedure was elaborated to solve, in constant time operation, the throughput maximization problem of significant original complexity. The optimal, yet simple, PA resolution allowed for achieving global optimal resource allocation by conveniently combining the Kuhn-Munkres channel assignment with the proposed PA methods. The results showed important performance gains obtained by applying SIC in D2D underlay systems in both HD and FD transmission schemes, promoting thereby the use of mutual SIC NOMA for D2D systems whenever possible. When applying mutual SIC, the comparison between HD and FD transmission scenarios showed that FD-SIC is more efficient for average to high SI cancellation capabilities, moderate CU rate requirements and significant D2D distances, while HD-SIC performs better especially at low SI cancellation capabilities. The obtained results advocate for the use of NOMA mutual SIC in conjunction with D2D as its application takes advantage of the near-far effect to unlock D2D implementation for further use case scenarios.

The contributions of this chapter led to the submission of the following journal paper:

A. Kilzi, J. Farah, C. Abdel Nour and C. Douillard, "Optimal Resource Allocation for Full-Duplex IoT Systems Underlaying Cellular Networks with Mutual SIC NOMA," under revision in *IEEE Internet Things J.*,

and to the publication of the following conference paper:

A. Kilzi, J. Farah, C. Abdel Nour and C. Douillard, "Inband Full-Duplex D2D Communications Underlaying Uplink Networks with Mutual SIC NOMA," 2020 *IEEE 31st Annual Int. Symp. Pers., Indoor and Mobile Radio Commun. (PIMRC)*, London, United Kingdom, Sept. 2020.

Appendix

6.A Necessary and Sufficient Conditions for the Existence of a Power Allocation Enabling FD-SIC

To determine the director vector of $W_{1,2}$, we first derive its parametric equation:

$$W_{1,2} = \begin{cases} P_u h_{b,u} = P_2 h_{b,d_2} - P_1 h_{b,d_1} \\ P_u h_{d_1,u} = P_2 h_d + P_1 \eta_1 \end{cases} \Rightarrow \begin{cases} P_u (h_{b,u}\eta_1 + h_{d_1,u} h_{b,d_1}) = P_2 (h_{b,d_2}\eta_1 + h_d h_{b,d_1}) \\ P_u (h_{d_1,u} h_{b,d_2} - h_{b,u} h_d) = P_1 (\eta_1 h_{b,d_2} + h_{b,d_1} h_d) \end{cases}$$

By choosing the parameter t such that $P_u = (h_{b,d_2}\eta_1 + h_d h_{b,d_1})t$, we get the director vector $\vec{u} = \begin{pmatrix} h_{d_1,u} h_{b,d_2} - h_{b,u} h_d \\ h_{b,u}\eta_1 + h_{d_1,u} h_{b,d_1} \\ \eta_1 h_{b,d_2} + h_{b,d_1} h_d \end{pmatrix}$. Then, writing down $\frac{z(\vec{u})}{x(\vec{u})} < \frac{h_{b,d_1}}{h_{b,u}}$ we get:

$$\frac{\eta_1 h_{b,d_2} + h_{b,d_1} h_d}{h_{d_1,u} h_{b,d_2} - h_{b,u} h_d} < \frac{h_{b,d_1}}{h_{b,u}}.$$

Since one of the necessary conditions for enabling FD-SIC is to have (6.8) valid (i.e. $h_{d_1,u}h_{b,d_2} - h_{b,u}h_d > 0$), a straightforward rearrangement of the above condition yields (6.32):

$$h_{d_1,u}h_{b,d_2} - h_{b,u}h_d > \frac{h_{b,u}}{h_{b,d_1}}(\eta_1 h_{b,d_2} + h_{b,d_1}h_d). \tag{6.32}$$

On the other hand, regrouping the terms in the following fashion yields (6.31) through:

$$\begin{aligned} h_{b,d_1}(h_{d_1,u}h_{b,d_2} - h_{b,u}h_d) &> (\eta_1 h_{b,d_2} + h_{b,d_1}h_d)h_{b,u}, \\ h_{b,d_1}h_{d_1,u}h_{b,d_2} - \eta_1 h_{b,d_2}h_{b,u} &> 2h_{b,u}h_d h_{b,d_1}, \\ h_{b,d_1}h_{d_1,u} - \eta_1 h_{b,u} &> 2h_{b,u}h_d\frac{h_{b,d_1}}{h_{b,d_2}}. \end{aligned} \tag{6.31}$$

Following the same steps for $W_{1,4}$, the parametric equations are derived first:

$$W_{1,4} = \begin{cases} P_u h_{b,u} = P_2 h_{b,d_2} - P_1 h_{b,d_1} \\ P_u h_{d_2,u} = P_1 h_d + P_2\eta_2 \end{cases} \Rightarrow \begin{cases} P_u(h_{b,u}h_d + h_{d_2,u}h_{b,d_1}) = P_2(h_{b,d_2}h_d + \eta_2 h_{b,d_1}) \\ P_u(h_{d_2,u}h_{b,d_2} - h_{b,u}\eta_2) = P_1(h_d h_{b,d_2} + h_{b,d_1}\eta_2) \end{cases}$$

by choosing the parameter t such that $P_u = (h_{b,d_2}h_d + \eta_2 h_{b,d_1})t$, we get the director vector $\vec{v} = \begin{pmatrix} h_{d_2,u}h_{b,d_2} - h_{b,u}\eta_2 \\ h_{b,u}h_d + h_{d_2,u}h_{b,d_1} \\ h_d h_{b,d_2} + h_{b,d_1}\eta_2 \end{pmatrix}$. Then writing down $\frac{z(\vec{v})}{x(\vec{v})} < \frac{h_{b,d_1}}{h_{b,u}}$ we get:

$$\frac{h_d h_{b,d_2} + h_{b,d_1}\eta_2}{h_{d_2,u}h_{b,d_2} - h_{bu}\eta_2} < \frac{h_{b,d_1}}{h_{b,u}}.$$

Since one of necessary condition for enabling FD-SIC is to have (6.30) valid (i.e. $h_{d_2,u}h_{b,d_2} - h_{b,u}\eta_2 > 0$), a straightforward rearrangement of the above condition yields (6.34):

$$h_{d_2,u}h_{b,d_2} - h_{b,u}\eta_2 > \frac{h_{b,u}}{h_{b,d_1}}(h_d h_{b,d_2} + h_{b,d_1}\eta_2). \tag{6.34}$$

On the other hand, regrouping the terms in the following fashion yields to (6.33) through:

$$\begin{aligned} h_{b,d_1}(h_{d_2,u}h_{b,d_2} - h_{b,u}\eta_2) &> (h_d h_{b,d_2} + h_{b,d_1}\eta_2)h_{b,u}, \\ h_{b,d_1}h_{d_2,u}h_{b,d_2} - h_d h_{b,d_2}h_{b,u} &> 2h_{b,u}\eta_2 h_{b,d_1}, \\ h_{b,d_1}h_{d_2,u} - h_d h_{b,u} &> 2h_{b,u}\eta_2\frac{h_{b,d_1}}{h_{b,d_2}}. \end{aligned} \tag{6.33}$$

Conclusion and Future Works

In this thesis, we have studied the combination of NOMA with multiple communication technologies such as D2D and FD, and network paradigms like DAS, CoMP, and UAVs in order to propose novel solutions for future generation networks relying on efficient interference management.

First, we began by addressing the problem of downlink power minimization in a DAS cell with user rate requirements. The review of the waterfilling concept for power allocation enabled important complexity simplifications, resulting in an efficient joint channel and power assignment schemes for the classical single-antenna NOMA serving. Then, we explored the possibilities provided by DAS to power-multiplexed signals from different RRHs. This led to the definition of the new concept of mutual SIC which unveiled the hidden potentials of DAS spatial diversity and enabled a complete inter-user interference cancellation. The obtained results showed the superiority of mutual SIC NOMA compared to standard single SIC.

Moving forward, the practical case of power-limited antennas was explored in the HDAS context. The presence of power constraints on serving antennas could potentially cause a failure in meeting user QoS requirements. Thus, the channel allocation conditions allowing for successful user serving were derived. The understanding of those constraints helped shaping the resource allocation strategies that meet the user demands for various system conditions. Two separate approaches were proposed to account for the antenna power limits during the power minimization process: one carrying great results for mild system conditions, and the other presenting robust performance for harsh system conditions.

Afterwards, we were interested in applying the principles of the mutual SIC procedure in a more general case encompassing multi-cell environments with enabled coordination/cooperation. Therefore, the mutual SIC concept was extended to account for JT-CoMP transmission and an arbitrary number of NOMA users. Then, the case studies of DMSIC and TMSIC were carried out, showing considerable performance improvement over previous OMA JT-CoMP techniques, or uncoordinated NOMA single SIC techniques. Furthermore, an interesting result was highlighted in the DMSIC case, where it was shown that favoring *cancellable* interference through unconventional choices of user-antenna association can be more beneficial than the traditional RSS-based antenna-user association.

The potential paradigm changes due to DMSIC and TMSIC motivated the proposal for positioning procedures of UAV-assisted networks that enable TMSIC application, and thereby, inherit all its advantages in terms of fairness and throughput. A probabilistic framework was proposed to account for the random nature of the air-to-ground links between the UAV and the users, while also seeking a TMSIC application. Several optimization metrics were proposed, providing a wide panel of selection for the network planner with a multitude of answers to face the variations in time of the users traffic

requirements.

Finally, the ecosystem of D2D enabled communications was addressed in conjunction with FD communication and NOMA between CUs and D2D devices. The mutual SIC conditions specific for FD-D2D where thoroughly investigated, and necessary and sufficient channel conditions were identified. Furthermore, a geometrical representation of the solution space allowed for an efficient optimal PA resolution, enabling subsequent optimal D2D-CU assignments. Moreover, the application of the mutual SIC procedure in the D2D context proved to be particularly beneficial from many view points. On the one hand, significant performance gains where achieved thanks to the interference cancellation, compared to the classical No-SIC strategy between CUs and D2Ds. On the other hand, implementing mutual SIC showed great complementarity with D2D applications: when classic D2D fails in bringing the additional capacity boost to a wireless system, due to too high D2D distances, mutual SIC can be applied to take advantage of the near-far effect.

Future Works

The work presented in this thesis showed how the key concept of mutual SIC can be adapted to various network scenarios and use cases such as DAS, CoMP, UAV-assisted networks and D2D communications. This is to be expected since any new asset for combating interference is valuable for tomorrow's future generation networks which are seriously interference-limited. Yet, several aspects of these studies are far from unveiling their fullest potentials.

First of all, the derived resource allocation schemes assumed perfect channel information knowledge. In practice, this is hardly feasible, and more research is required to determine the outcome of the proposed RA techniques for the context of statistical CSI knowledge and/or imperfect instantaneous CSI. Consequently, a possible work direction could be to design robust RA schemes mitigating the performance gap between perfect and noisy CSI, where different CSI noise models could be assumed depending on the context [37–39].

A direct sequel to that study would be the analysis of the impact of imperfect SIC implementation on the performance of the proposed procedures. One the one hand, the erroneous CSI could mislead the designer into the application of mutual SIC in inadequate scenarios, which might fire back in terms of the incurred interference. On the other hand, residual interference could still remain following an imperfect SIC procedure due to channel quantization and estimation errors resulting in imperfect equalization. The induced performance degradation would require further testing and possibly mitigation through robust RA schemes taking into account the mentioned imperfection in their design.

Although we proposed a generalized mutual SIC procedure in the CoMP scenario, the exponential complexity of decoding orders to be considered keeps NOMA cluster sizes to a maximum of three users. A possible future work direction could be to combine experimental and theoretical analysis to determine the most likely decoding orders for achieving mutual SIC. This can provide the linear capacity gains with every new added user while jeopardizing the scheduling complexity. A direct follow up to this study can be to design user grouping strategies enabling maximum number of mutual SIC applications. In that regard, state-of-the-art user-centric clustering techniques in CoMP can be envisioned to

include multiple users at once. Additionally, the study can be extended to explore the implementation of mutual SIC into Multiple-Input Multiple-Output systems.

In Chapter 6, the proposed geometrical procedure could inspire the resolution of higher dimensionality PA problems where more than a single CU is accessing the same resource as the D2D pair, or conversely, more than two devices are in D2D communication. Furthermore, it could be interesting to derive patterns for D2D-CU pairing that would be purely based on the knowledge of the channel conditions, or even further, on their relative geographic positioning. This could be done through various tools (e.g. machine learning techniques) and would simplify the channel assignment step and facilitate the integration of the proposed methodologies to DASs.

Bibliography

[1] J. B. Say, *A treatise on political economy.* Lippincott, Grambo & Company, 1851.

[2] Cisco, "Cisco Visual Networking Index: Forecast and Trends, 2018-2023," Mar. 2020.

[3] Al-Zahrani, Ali Y and Yu, F Richard and Huang, Minyi, "A joint cross-layer and co-layer interference management scheme in hyperdense heterogeneous networks using mean-field game theory," *IEEE Trans. Veh. Technol.*, vol. 65, no. 3, pp. 1522–1535, 2015.

[4] Soret, Beatriz and Pedersen, Klaus I and Jørgensen, Niels TK and Fernández-López, Víctor, "Interference coordination for dense wireless networks," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 102–109, 2015.

[5] W. Choi and J. G. Andrews, "Downlink Performance and Capacity of Distributed Antenna Systems in a Multicell Environment," *IEEE Trans. Wireless Commun.*, vol. 6, no. 1, pp. 69–73, Jan 2007.

[6] R. W. H. Jr, T. Wu, Y. H. Kwon, and A. C. K. Soong, "Multiuser MIMO in Distributed Antenna Systems With Out-of-Cell Interference," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4885–4899, Oct 2011.

[7] M. Ding, P. Wang, D. López-Pérez, G. Mao, and Z. Lin, "Performance impact of los and nlos transmissions in dense cellular networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, pp. 2365–2380, 2016.

[8] 3GPP, "Mobile Broadband Innovation path to 4G: Release 9,10 and Beyond," *3rd Generation Partnership Project (3GPP), TR*, Feb. 2010.

[9] 3GPP, "3GPP TR 36.819 Technical Specification Group Radio Access Network; Coordinated Multi-Point Operation for LTE Physical Layer Aspects," 2011.

[10] N. Bhushan, J. Li, D. Malladi, R. Gilmore, D. Brenner, A. Damnjanovic, R. T. Sukhavasi, C. Patel, and S. Geirhofer, "Network densification: the dominant theme for wireless evolution into 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 82–89, 2014.

[11] B. Romanous, N. Bitar, A. Imran, and H. Refai, "Network densification: Challenges and opportunities in enabling 5G," in *2015 IEEE 20th Int. Workshop Comput. Aided Modelling and Design of Commun. Links and Netw. (CAMAD)*, 2015, pp. 129–134.

[12] A. Asadi, Q. Wang, and V. Mancuso, "A survey on device-to-device communication in cellular networks," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 4, pp. 1801–1819, 2014.

[13] Alliance, Wi-Fi, "Wi-Fi peer-to-peer (P2P) technical specification," *www. wi-fi. org/Wi-Fi_Direct. php*, 2010.

[14] Mumtaz, Shahid and Rodriguez, Jonathan, *Smart device to smart device communication*. Springer, 2014.

[15] Doppler, Klaus and Rinne, Mika and Wijting, Carl and Ribeiro, Cássio B and Hugl, Klaus, "Device-to-device communication as an underlay to LTE-advanced networks," *IEEE Commun. Mag.*, vol. 47, no. 12, pp. 42–49, 2009.

[16] Peng, Tao and Lu, Qianxi and Wang, Haiming and Xu, Shaoyi and Wang, Wenbo, "Interference avoidance mechanisms in the hybrid cellular and device-to-device systems," in *2009 IEEE 20th Int. Symp. Pers., Indoor and Mobile Radio Commun.* IEEE, 2009, pp. 617–621.

[17] Yu, Chia-Hao and Doppler, Klaus and Ribeiro, Cássio and Tirkkonen, Olav, "Performance impact of fading interference to device-to-device communication underlaying cellular networks," in *2009 IEEE 20th Int. Symp. Pers., Indoor and Mobile Radio Commun.* IEEE, 2009, pp. 858–862.

[18] Min, Hyunkee and Lee, Jemin and Park, Sungsoo and Hong, Daesik, "Capacity enhancement using an interference limited area for device-to-device uplink underlaying cellular networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 12, pp. 3995–4000, 2011.

[19] Xu, Wei and Liang, Le and Zhang, Hua and Jin, Shi and Li, James CF and Lei, Ming, "Performance enhanced transmission in device-to-device communications: Beamforming or interference cancellation?" in *2012 IEEE Global Commun. Conf. (GLOBECOM)*. IEEE, 2012, pp. 4296–4301.

[20] S. Ali, N. Rajatheva, and M. Latva-aho, "Full duplex device-to-device communication in cellular networks," in *2014 European Conference on Networks and Communications (EuCNC)*, 2014, pp. 1–5.

[21] K. E. Kolodziej, J. G. McMichael, and B. T. Perry, "Multitap RF Canceller for In-Band Full-Duplex Wireless Communications," *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, pp. 4321–4334, June 2016.

[22] L. Laughlin, C. Zhang, M. A. Beach, K. A. Morris, and J. Haine, "A Widely Tunable Full Duplex Transceiver Combining Electrical Balance Isolation and Active Analog Cancellation," in *2015 IEEE 81st Veh. Technol. Conf. (VTC Spring)*, May 2015, pp. 1–5.

[23] 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, 2013.

[24] Shende, Nirmal V and Gürbüz, Özgür and Erkip, Elza, "Half-Duplex or Full-Duplex Communications: Degrees of Freedom Analysis Under Self-Interference," *IEEE Trans. Wireless Commun.*, vol. 17, no. 2, pp. 1081–1093, 2018.

[25] S. Goyal, C. Galiotto, N. Marchetti, and S. Panwar, "Throughput and coverage for a mixed full and half duplex small cell network," in *2016 IEEE Int. Conf. Commun. (ICC)*, 2016, pp. 1–7.

[26] S. Goyal, P. Liu, S. S. Panwar, R. A. Difazio, R. Yang, and E. Bala, "Full duplex cellular systems: will doubling interference prevent doubling capacity" *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 121–127, 2015.

[27] Y. Saito and A. Benjebbour and Y. Kishiyama and T. Nakamura, "System-Level Performance Evaluation of Downlink Non-Orthogonal Multiple Access (NOMA)," in *2013 IEEE 24th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Sep 2013, pp. 611–615.

[28] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-Orthogonal Multiple Access (NOMA) for Cellular Future Radio Access," in *2013 IEEE 77th Veh. Tech. Conf. (VTC Spring)*, June 2013, pp. 1–5.

[29] 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.

[30] J. Choi, "Non-Orthogonal Multiple Access in Downlink Coordinated Two-Point Systems," *IEEE Commun. Lett.*, vol. 18, no. 2, pp. 313–316, Feb. 2014.

[31] 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.

[32] L. Lei and Z. Zhong and C. Lin and X. Shen, "Operator controlled device-to-device communications in LTE-advanced networks," *IEEE Wireless Commun.*, vol. 19, no. 3, pp. 96–104, 2012.

[33] Jameel, Furqan and Hamid, Zara and Jabeen, Farhana and Zeadally, Sherali and Javed, Muhammad Awais, "A survey of device-to-device communications: Research issues and challenges," *IEEE Commun. Surveys Tut.*, vol. 20, no. 3, pp. 2133–2168, 2018.

[34] Corson, M Scott and Laroia, Rajiv and Li, Junyi and Park, Vincent and Richardson, Tom and Tsirtsis, George, "Toward proximity-aware internetworking," *IEEE Wireless Commun.*, vol. 17, no. 6, pp. 26–33, 2010.

[35] Doppler, Klaus and Rinne, Mika and Wijting, Carl and Ribeiro, Cássio B and Hugl, Klaus, "Device-to-device communication as an underlay to LTE-advanced networks," *IEEE Commun. Mag.*, vol. 47, no. 12, pp. 42–49, 2009.

[36] Viswanathan, Harish and Lucent, A, "Expanding the role of the mobile network operator in M2M," in *1st ETSI TC M2M Wksp.*, 2010.

[37] S. Narmatha, R. Jeyanthi, and N. Malmurugan, "Amplify and forward relay network optimization with imperfect CSI," in *2015 2nd Int. Conf. Electron. Commun. Syst. (ICECS)*, 2015, pp. 469–473.

[38] Nguyen, Thanh-Luan and Do, Dinh-Thuan, "Power allocation schemes for wireless powered NOMA systems with imperfect CSI: An application in multiple antenna–based relay," *Int. J. Commun. Syst.*, vol. 31, no. 15, p. e3789, 2018.

[39] Li, Xingwang and Li, Jingjing and Li, Lihua, "Performance analysis of impaired SWIPT NOMA relaying networks over imperfect weibull channels," *IEEE Syst. J.*, vol. 14, no. 1, pp. 669–672, 2019.

[40] Webb, William, *Wireless communications: The future*. John Wiley & Sons, 2007.

[41] Chu, Xiaoli and Lopez-Perez, David and Yang, Yang and Gunnarsson, Fredrik, *Heterogeneous cellular networks: theory, simulation and deployment*. Cambridge University Press, 2013.

[42] Claussen, Holger and Ho, Lester TW and Samuel, Louis G, "An overview of the femtocell concept," *Bell Labs Technical Journal*, vol. 13, no. 1, pp. 221–245, 2008.

[43] H. Claussen, D. Lopez-Perez, L. Ho, R. Razavi, and S. Kucera, *Small cell networks: deployment, management, and optimization*. John Wiley & Sons, 2017.

[44] R. Heath, S. Peters, Y. Wang, and J. Zhang, "A Current Perspective on Distributed Antenna Systems for the Downlink of Cellular Systems," *IEEE Commun. Mag.*, vol. 51, no. 4, pp. 161–167, Apr 2013.

[45] Wu, Jun and Zhang, Zhifeng and Hong, Yu and Wen, Yonggang, "Cloud radio access network (C-RAN): a primer," *IEEE Network*, vol. 29, no. 1, pp. 35–41, 2015.

[46] Checko, Aleksandra and Christiansen, Henrik L and Yan, Ying and Scolari, Lara and Kardaras, Georgios and Berger, Michael S and Dittmann, Lars, "Cloud RAN for mobile networks—A technology overview," *IEEE Commun. Surveys & Tutorials*, vol. 17, no. 1, pp. 405–426, 2014.

[47] Andrews, Jeffrey G, "Seven ways that HetNets are a cellular paradigm shift," *IEEE Commun. Mag.*, vol. 51, no. 3, pp. 136–144, 2013.

[48] Z. Sattar, J. V. De Carvalho Evangelista, G. Kaddoum, and N. Batani, "Full-Duplex Two-Tier Heterogeneous Network With Decoupled Access: Cell Association, Coverage, and Spectral Efficiency Analysis," *IEEE Access*, vol. 8, pp. 172 982–172 995, 2020.

[49] E. A. Beyazit, B. Özbek, and D. L. Ruyet, "On Stream Selection for Interference Alignment in Heterogeneous Networks," *EURASIP J. Wirel. Commun. Netw.*, vol. 2016, p. 80, 2016.

[50] Tehrani, Mohsen Nader and Uysal, Murat and Yanikomeroglu, Halim, "Device-to-device communication in 5G cellular networks: challenges, solutions, and future directions," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 86–92, 2014.

[51] H. Dang, M. Van Nguyen, D. Do, H. Pham, B. Selim, and G. Kaddoum, "Joint Relay Selection, Full-Duplex and Device-to-Device Transmission in Wireless Powered NOMA Networks," *IEEE Access*, vol. 8, pp. 82 442–82 460, 2020.

[52] B. Özbek, M. Pischella, and D. Le Ruyet, "Energy Efficient Resource Allocation for Underlaying Multi-D2D Enabled Multiple-Antennas Communications," *IEEE Trans. Veh. Technol.*, vol. 69, no. 6, pp. 6189–6199, 2020.

[53] N. Docomo, "Evaluation methodologies for downlink multiuser superposition transmissions," in *3GPP TSG RAN WG1 Meeting*, vol. 81, 2015, pp. R1–153 332.

[54] Meredith, John M, "Study on downlink multiuser superposition transmission for LTE," in *TSG RAN Meeting*, vol. 67, 2015.

[55] Yuan, Zhifeng and Yu, Guanghui and Li, Weimin and Yuan, Yifei and Wang, Xinhui and Xu, Jun, "Multi-user shared access for internet of things," in *2016 IEEE 83rd Veh. Technol. Conf. (VTC Spring)*. IEEE, 2016, pp. 1–5.

[56] Hoshyar, Reza and Wathan, Ferry P and Tafazolli, Rahim, "Novel Low-Density Signature for Synchronous CDMA Systems Over AWGN Channel," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1616–1626, 2008.

[57] Nikopour, Hosein and Baligh, Hadi, "Sparse code multiple ccess," in *2013 IEEE 24th Annu. Int. Symp. Pers., Indoor, and Mobile Radio Commun. (PIMRC)*. IEEE, 2013, pp. 332–336.

[58] Zeng, Jie and Kong, Dan and Su, Xin and Rong, Liping and Xu, Xibin, "On the performance of pattern division multiple access in 5G systems," in *2016 8th Int. Conf. Wireless Commun. & Signal Process. (WCSP)*. IEEE, 2016, pp. 1–5.

[59] Huang, Jiachen and Peng, Kewu and Pan, Changyong and Yang, Fang and Jin, Huangping, "Scalable Video Broadcasting Using Bit Division Multiplexing," *IEEE Trans. Broadcast.*, vol. 60, no. 4, pp. 701–706, 2014.

[60] Z. Ding, M. Peng, and H. V. Poor, "Cooperative Non-Orthogonal Multiple Access in 5G Systems," *IEEE Commun. Lett.*, vol. 19, no. 8, pp. 1462–1465, 2015.

[61] Wei, Zhiqiang and Yuan, Jinhong and Ng, Derrick Wing Kwan and Elkashlan, Maged and Ding, Zhiguo, "A survey of downlink non-orthogonal multiple access for 5G wireless communication networks," *arXiv preprint arXiv:1609.01856*, 2016.

[62] Kimy, Beomju and Lim, Sungmook and Kim, Hyungjong and Suh, Sangwook and Kwun, Jonghyung and Choi, Sooyong and Lee, Chungyong and Lee, Sanghoon and Hong, Daesik, "Non-orthogonal multiple access in a downlink multiuser beamforming system," in *MILCOM 2013-2013 IEEE Military Commun. Conf.* IEEE, 2013, pp. 1278–1283.

[63] Z. Ding, F. Adachi, and H. V. Poor, "Performance of MIMO-NOMA Downlink Transmissions," in *2015 IEEE Global Commun. Conf. (GLOBECOM)*, Dec 2015, pp. 1–6.

[64] 3GPP, "Study on Network-Assisted Interference Cancellation and Suppression (NAIC) for LTE (Release 12)," in *TR 36.866*, Mar. 2014.

[65] Kardaras, Georgios and Lanzani, Christian, "Advanced multimode radio for wireless & mobile broadband communication," in *2009 European Wireless Technology Conference*. IEEE, 2009, pp. 132–135.

[66] Lin, Yonghua and Shao, Ling and Zhu, Zhenbo and Wang, Qing and Sabhikhi, Ravie K, "Wireless network cloud: Architecture and system requirements," *IBM Journal of Research and Development*, vol. 54, no. 1, pp. 4–1, 2010.

[67] Kuilin, C and Ran, D, "C-RAN the road towards green ran," *China Mobile Research Institute, White Paper*, 2011.

[68] Werthmann, Thomas and Grob-Lipski, Heidrun and Proebster, Magnus, "Multiplexing gains achieved in pools of baseband computation units in 4G cellular networks," in *2013 IEEE 24th Annu. Int. Symp. Pers., Indoor, and Mobile Radio Commun. (PIMRC)*. IEEE, 2013, pp. 3328–3333.

[69] Bhaumik, Sourjya and Chandrabose, Shoban Preeth and Jataprolu, Manjunath Kashyap and Kumar, Gautam and Muralidhar, Anand and Polakos, Paul and Srinivasan, Vikram and Woo, Thomas, "CloudIQ: A Framework for Processing Base Stations in a Data Center," in *Proceedings of the 18th Annu. Int. Conf. Mobile Comput. Netw.*, ser. Mobicom '12. New York, NY, USA: Association for Computing Machinery, 2012, p. 125–136. [Online]. Available: https://doi.org/10.1145/2348543.2348561

[70] Chen, C, "C-RAN: the Road Towards Green Radio Access Network." Presentation, Aug. 2012.

[71] Ashraf, Imran and Ho, Lester TW and Claussen, Holger, "Improving energy efficiency of femtocell base stations via user activity detection," in *2010 IEEE wireless Commun. Netw. Conf.* IEEE, 2010, pp. 1–5.

[72] Claussen, Holger and Ashraf, Imran and Ho, Lester TW, "Dynamic idle mode procedures for femtocells," *Bell Labs Technical Journal*, vol. 15, no. 2, pp. 95–116, 2010.

[73] S. Stefanatos and A. Alexiou, "Access point density and bandwidth partitioning in ultra dense wireless networks," *IEEE Trans. Commun.*, vol. 62, no. 9, pp. 3376–3384, 2014.

[74] I. Hwang, B. Song, and S. S. Soliman, "A holistic view on hyper-dense heterogeneous and small cell networks," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 20–27, 2013.

[75] M. I. Kamel, W. Hamouda, and A. M. Youssef, "Multiple association in ultra-dense networks," in *2016 IEEE Int. Conf. Commun. (ICC)*. IEEE, 2016, pp. 1–6.

[76] Wang, Xiaoyi and Visotsky, Eugene and Ghosh, Amitava, "Dynamic cell muting for ultra dense indoor small cell deployment scenario," in *2015 IEEE Int. Conf. on Commun. Workshop (ICCW)*. IEEE, 2015, pp. 148–153.

[77] Liu, Ling and Garcia, Virgile and Tian, Lin and Pan, Zhengang and Shi, Jinglin, "Joint clustering and inter-cell resource allocation for CoMP in ultra dense cellular networks," in *2015 IEEE Int. Conf. on Commun. (ICC)*. IEEE, 2015, pp. 2560–2564.

[78] Cho, Moon-Je and Ban, Tae-Won and Jung, Bang Chul and Yang, Hyun Jong, "A distributed scheduling with interference-aware power control for ultra-dense networks," in *2015 IEEE Int. Conf. Commun. (ICC)*. IEEE, 2015, pp. 1661–1666.

[79] Foschini, GJ and Karakayali, K and Valenzuela, RA, "Coordinating multiple antenna cellular networks to achieve enormous spectral efficiency," *IEEE Proc. Commun.*, vol. 153, no. 4, pp. 548–555, 2006.

[80] Rahul, Hariharan Shankar and Kumar, Swarun and Katabi, Dina, "JMB: scaling wireless capacity with user demands," *ACM SIGCOMM Computer Communication Review*, vol. 42, no. 4, pp. 235–246, 2012.

[81] Jungnickel, Volker and Manolakis, Konstantinos and Zirwas, Wolfgang and Panzner, Berthold and Braun, Volker and Lossow, Moritz and Sternad, Mikael and Apelfröjd, Rikke and Svensson, Tommy, "The role of small cells, coordinated multipoint, and massive MIMO in 5G," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 44–51, 2014.

[82] Xu, Wei and Huang, Jian and Yang, Fei and Zhang, Hua, "Dynamic cell clustering for CoMP in LTE-A and its calibrated system level performance evaluation," in *2013 5th IEEE Internat. Symp. Microw., Antenna, Propag. EMC Technol. for Wireless Commun.* IEEE, 2013, pp. 71–77.

[83] Liu, Mingqian and Li, Bingbing and Tong, Zhao, "Cell clustering and resource allocation scheme for CoMP SU-MIMO," *Wireless Pers. Commun.*, vol. 69, no. 4, pp. 1399–1411, 2013.

[84] S, Research, "ProSe (Proximity Services) for LTE & 5G Networks: 2017 – 2030 – Opportunities, Challenges, Strategies & Forecasts," SNS Telecom & IT, 2017.

[85] 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 *2017 IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2017, pp. 618–624.

[86] F. Yuan, S. Jin, Y. Huang, K. k. Wong, Q. T. Zhang, and H. Zhu, "Joint Wireless Information and Energy Transfer in Massive Distributed Antenna Systems," *IEEE Commun. Mag.*, vol. 53, no. 6, pp. 109–116, Jun 2015.

[87] H. Li, G. P. Koudouridis, and J. Zhang, "Antenna Selection Schemes for Energy Efficiency in Distributed Antenna Systems," in *2012 IEEE Int. Conf. Commun. (ICC)*, Jun 2012, pp. 5619–5623.

[88] C. He, G. Y. Li, F. C. 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, Mar 2014.

[89] 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.

[90] 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.

[91] 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, Aug 2017.

[92] X. Li, C. Li, and Y. Jin, "Dynamic Resource Allocation for Transmit Power Minimization in OFDM-Based NOMA Systems," *IEEE Commun. Lett.*, vol. 20, no. 12, pp. 2558–2561, Dec 2016.

[93] J. Cui, Z. Ding, and P. Fan, "Power Minimization Strategies in Downlink MIMO-NOMA Systems," in *2017 IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–6.

[94] Y. Liu, "Complexity Analysis of Joint Subcarrier and Power Allocation for the Cellular Downlink OFDMA System," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 661–664, 2014.

[95] D. Yuan, J. Joung, C. K. Ho, and S. Sun, "On Tractability Aspects of Optimal Resource Allocation in OFDMA Systems," *IEEE Trans. Veh. Technol.*, vol. 62, no. 2, pp. 863–873, 2013.

[96] Chaohuang Zeng, L. M. C. Hoo, and J. M. Cioff, "Efficient water-filling algorithms for a Gaussian multiaccess channel with ISI," in *Veh. Technol. Conf. Fall 2000. IEEE VTS Fall VTC2000. 52nd Vehicular Technology Conference (Cat. No.00CH37152)*, vol. 3, 2000, pp. 1072–1077 vol.3.

[97] D. D. Yu and J. M. Cioffi, "SPC10-2: Iterative Water-filling for Optimal Resource Allocation in OFDM Multiple-Access and Broadcast Channels," in *IEEE Globecom 2006*, Nov 2006, pp. 1–5.

[98] 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.

[99] H. Tabassum, E. Hossain, and J. Hossain, "Modeling and analysis of uplink non-orthogonal multiple access in large-scale cellular networks using poisson cluster processes," *IEEE Trans. Commun.*, vol. 65, no. 8, pp. 3555–3570, Aug. 2017.

[100] 3GPP, TR25-814 (V7.1.0), Physical Layer Aspects for Evolved UTRA (2006).

[101] J. Farah, E. Sfeir, C. A. Nour, and C. Douillard, "New Efficient Energy-Saving Techniques for Resource Allocation in Downlink OFDMA Transmission Systems," in *2017 IEEE Symp. Comput. Commun. (ISCC)*, July 2017, pp. 1056–1062.

[102] H. Ren, N. Liu, C. Pan, and C. He, "Energy Efficiency Optimization for MIMO Distributed Antenna Systems," *IEEE Trans. Veh. Technol.*, vol. 66, no. 3, pp. 2276–2288, March 2017.

[103] M. Tham, S. F. Chien, D. W. Holtby, and S. Alimov, "Energy-Efficient Power Allocation for Distributed Antenna Systems With Proportional Fairness," *IEEE Trans. Green Commun. Netw.*, vol. 1, no. 2, pp. 145–157, June 2017.

[104] 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.

[105] 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.

[106] M.-R. Hojeij, J. Farah, C. A. Nour, and C. Douillard, "New Optimal and Suboptimal Resource Allocation Techniques for Downlink Non-Orthogonal Multiple Access," *Wireless Pers. Commun.*, vol. 87, no. 3, pp. 837–867, Apr 2016.

[107] A. Beylerian and T. Ohtsuki, "Coordinated Non-Orthogonal Multiple Access (CO-NOMA)," in *2016 IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2016, pp. 1–5.

[108] M. S. Ali, E. Hossain, and D. I. Kim, "Coordinated Multipoint Transmission in Downlink Multi-Cell NOMA Systems: Models and Spectral Efficiency Performance," *IEEE Wireless Commun.*, vol. 25, no. 2, pp. 24–31, Apr. 2018.

[109] M. S. Ali, E. Hossain, A. Al-Dweik, and D. I. Kim, "Downlink Power Allocation for CoMP-NOMA in Multi-Cell Networks," *IEEE Trans. Commun.*, vol. 66, no. 9, pp. 3982–3998, Sept. 2018.

[110] M. S. Ali, H. Tabassum, and E. Hossain, "Dynamic User Clustering and Power Allocation for Uplink and Downlink Non-Orthogonal Multiple Access (NOMA) Systems," *IEEE Access*, vol. 4, pp. 6325–6343, 2016.

[111] S. M. Alamouti, "A Simple Transmit Diversity Technique for Wireless Communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[112] E. Pateromichelakis, M. Shariat, A. u. Quddus, and R. Tafazolli, "On the Evolution of Multi-Cell Scheduling in 3GPP LTE / LTE-A," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 2, pp. 701–717, Second 2013.

[113] 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. Veh. Technol.*, vol. 67, no. 12, pp. 11 873–11 885, Dec. 2018.

[114] M. J. Youssef, J. Farah, C. Abdel Nour, and C. Douillard, "Resource Allocation for Mixed Traffic Types in Distributed Antenna Systems Using NOMA," in *2018 IEEE 77th Veh. Technol. Conf. (VTC fall)*, Aug. 2018, pp. 1–5.

[115] J. Ren, Z. Wang, M. Xu, F. Fang, and Z. Ding, "Unsupervised user clustering in non-orthogonal multiple access," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2019, pp. 3332–3336.

[116] 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.

[117] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Unmanned Aerial Vehicle With Underlaid Device-to-Device Communication: Performance and Tradeoffs," *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, pp. 3949–3963, June 2016.

[118] R. I. Bor-Yaliniz, A. El-Keyi, and H. Yanikomeroglu, "Efficient 3-D placement of an aerial base station in next generation cellular networks," in *2016 IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–5.

[119] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Drone Small Cells in the Clouds: Design, Deployment and Performance Analysis," in *2015 IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2015, pp. 1–6.

[120] 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.

[121] E. Kalantari, H. Yanikomeroglu, and A. Yongacoglu, "On the Number and 3D Placement of Drone Base Stations in Wireless Cellular Networks," in *2016 IEEE 84th Veh. Technol. Conf. (VTC-Fall)*, Sept. 2016, pp. 1–6.

[122] B. Li, Z. Fei, and Y. Zhang, "UAV Communications for 5G and Beyond: Recent Advances and Future Trends," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2241–2263, 2019.

[123] 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, July 2019.

[124] M. T. Nguyen and L. B. Le, "NOMA User Pairing and UAV Placement in UAV-Based Wireless Networks," in *ICC 2019 - 2019 IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–6.

[125] 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, May 2019.

[126] X. Mu, Y. Liu, L. Guo, and J. Lin, "Non-Orthogonal Multiple Access for Air-to-Ground Communication," *IEEE Trans. Commun.*, vol. 68, no. 5, pp. 2934–2949, 2020.

[127] G. Geraci, A. Garcia-Rodriguez, L. Galati Giordano, D. López-Pérez, and E. Björnson, "Understanding UAV Cellular Communications: From Existing Networks to Massive MIMO," *IEEE Access*, vol. 6, pp. 67 853–67 865, 2018.

[128] M. Youssef, C. A. Nour, J. Farah, and C. Douillard, "Backhaul-Constrained Resource Allocation and 3D Placement for UAV-Enabled Networks," in *2019 IEEE 90th Veh. Tech. Conf. (VTC2019-Fall)*, Sept. 2019, pp. 1–7.

[129] J. Qin, Z. Wei, C. Qiu, and Z. Feng, "Edge-Prior Placement Algorithm for UAV-Mounted Base Stations," in *2019 IEEE Wireless Commun. and Netw. Conf. (WCNC)*, Apr. 2019, pp. 1–6.

[130] A. Al-Hourani, S. Kandeepan, and A. Jamalipour, "Modeling air-to-ground path loss for low altitude platforms in urban environments," in *2014 IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2014, pp. 2898–2904.

[131] 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, 2014.

[132] Garajová, Elif and Mečiar, Martin, "Solving and Visualizing Nonlinear Set Inversion Problems*." [Online]. Available: https://interval.louisiana.edu/reliable-computing-journal/volume-22/reliable-computing-22-pp-104-115.pdf

[133] Jaulin, Luc and Walter, Eric, "Set inversion via interval analysis for nonlinear bounded-error estimation," *Automatica*, vol. 29, no. 4, pp. 1053–1064, 1993.

[134] W. Shi, J. Li, W. Xu, H. Zhou, N. Zhang, S. Zhang, and X. Shen, "Multiple Drone-Cell Deployment Analyses and Optimization in Drone Assisted Radio Access Networks," *IEEE Access*, vol. 6, pp. 12 518–12 529, 2018.

[135] J. Chen, J. Jia, Y. Liu, X. Wang, and A. H. Aghvami, "Optimal Resource Block Assignment and Power Allocation for D2D-Enabled NOMA Communication," *IEEE Access*, vol. 7, pp. 90 023–90 035, 2019.

[136] Y. Pan, C. Pan, Z. Yang, and M. Chen, "Resource Allocation for D2D Communications Underlaying a NOMA-Based Cellular Network," *IEEE Wireless Commun. Lett.*, vol. 7, no. 1, pp. 130–133, Feb 2018.

[137] J. Zhao, Y. Liu, K. K. Chai, Y. Chen, and M. Elkashlan, "Joint Subchannel and Power Allocation for NOMA Enhanced D2D Communications," *IEEE Trans. on Commun.*, vol. 65, no. 11, pp. 5081–5094, Nov 2017.

[138] S. Alemaishat, O. A. Saraereh, I. Khan, and B. J. Choi, "An Efficient Resource Allocation Algorithm for D2D Communications Based on NOMA," *IEEE Access*, vol. 7, pp. 120 238–120 247, 2019.

[139] Y. Wang, D. Zhai, R. Zhang, and Z. Zhang, "Sum-Rate Maximization for D2D and Cellular Hybrid Networks Enhanced by NOMA," in *2019 IEEE 20th International Conf. on High Performance Switching and Routing (HPSR)*, 2019, pp. 1–5.

[140] Y. Dai, M. Sheng, J. Liu, N. Cheng, X. Shen, and Q. Yang, "Joint Mode Selection and Resource Allocation for D2D-Enabled NOMA Cellular Networks," *IEEE Trans. on Veh. Technol.*, vol. 68, no. 7, pp. 6721–6733, 2019.

[141] D. Zhai, R. Zhang, Y. Wang, H. Sun, L. Cai, and Z. Ding, "Joint User Pairing, Mode Selection, and Power Control for D2D-Capable Cellular Networks Enhanced by Nonorthogonal Multiple Access," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 8919–8932, 2019.

[142] P. Mach, Z. Becvar, and T. Vanek, "In-Band Device-to-Device Communication in OFDMA Cellular Networks: A Survey and Challenges," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 1885–1922, 2015.

[143] H. Chour, F. Bader, Y. Nasser, and O. Bazzi, "GALEN: A Geometric Framework for Global Optimal Power Allocation in a Full Duplex D2D Network," in *2019 IEEE Wireless Commun. and Netw. Conf. (WCNC)*, April 2019, pp. 1–7.

[144] J. Munkres, "Algorithms for the assignment and transportation problems," *Journal of the Society for Industrial and Applied Mathematics*, vol. 5, no. 1, pp. 32–38, 1957. [Online]. Available: https://doi.org/10.1137/0105003

[145] Cui, H and Zhang, J and Cui, C and Chen, Q, "Solving large-scale assignment problems by Kuhn-Munkres algorithm," in *international conference on advances in mechanical engineering and industrial informatics. Hangzhou, Zhejiang*, 2016, pp. 822–827.

UNIVERSITE BRETAGNE LOIRE / MATHSTIC

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

Titre : Nouvelles approches de gestions des interférences pour les réseaux de communication 5G et futurs utilisant la NOMA

Mots clés : Accès multiple non orthogonal, annulation mutuelle successive des interférences, multipoint coordonné , communication entre appareils, duplex intégral.

Résumé : La demande pour les systèmes de communications 5G et au-delà vont dans le sens de toujours plus de débit, plus de connectivité, moins de latence et plus de fiabilité. Pour répondre à cette demande en constante croissance, différentes propositions sont sur la table, allant des méthodes d'accès multiple non-orthogonales (NOMA), aux systèmes device-to-device (D2D) munis de fonctionalité de duplex intégral (FD), en passant par des architectures de réseaux plus denses tels que les petites cellules, les systèmes d'antennes distribués (DAS) et le cloud RAN (CRAN), et employant des méthodes de coopérations inter-cellules sophitisquées telles que le coordinated multipoint (CoMP). De nouveaux éléments tels que les drones (UAV) sont également envisagés pour servir des utilisateurs. Bien que les techniques proposées ci-dessus soient de natures très variées, le dénominateur commun qui sous-tend ces technologies se rapporte à la problématique de gestion d'interférences au sens large : interférences entre utilisateurs pour le NOMA, interférences entre cellules pour les DAS et le CoMP, et interférences entre systèmes hétérogènes pour le D2D et les UAVs. Dans cette thèse, nous proposons un nouveau schéma d'annulation d'interférences basé sur les récepteurs à annulation successive d'interférence (SIC) du NOMA que nous baptisons mutual SIC. Nous montrons le grand intérêt que représente cette technique quand elle est adéquatement intégrée aux techonologies mentionnées précédemment, tant dans des scénarios de minimisation de puissance de transmission que dans des scenarios de maximization de débit total et de l'équité entre utilisateurs.

Title : New Approaches for Interference Management in Future Generation Networks for 5G and Beyond using NOMA

Keywords : Non-Orthogonal Multiple Access, mutual Successive Interference Cancellation, Coordinated Multipoint, Device to Device, Full Duplex.

Abstract: The demands for 5G systems and beyond are pushing for more throughput, more connectivity, less latency and more reliability. To meet this ever-growing demand, various proposals are on the table ranging from non-orthogonal multiple access (NOMA), device-to-device (D2D) systems with full duplex (FD) functionality, to denser network architectures such as small cells, distributed antenna systems (DAS) and cloud RAN (CRAN), and employing sophisticated inter-cell cooperation methods such as coordinated multipoint (CoMP). New elements such as unmanned aerial vehicles (UAVs) are also being considered for current and next generation networks. Although the techniques proposed above are diverse in nature, the common denominator underlying these technologies comes back to tackling the broad problem of interference management: user-to-user interference management for NOMA, cell-to-cell interference management for DAS and CoMP, and interference management between heterogeneous systems for D2Ds and UAVs. In this thesis, we propose a new interference cancellation scheme allowing for a complete interference cancellation based on the NOMA successive interference cancelation (SIC) receivers that we call mutual SIC. We show the great interest that this technique represents when it is adequately integrated with the above-mentioned technologies, both in transmit power minimization scenarios and in rate craving scenarios of total throughput maximization with a consideration to user fairness.