Allocation de ressources distribuée dans les réseaux OFDMA multi-cellulaires

Mylène Pischella

TELECOM ParisTech

École Doctorale d'Informatique, Télécommunications et Électronique de Paris

Thèse

présentée pour obtenir le grade de Docteur
de l'École Nationale Supérieure des Télécommunications

Spécialité : **Électronique et Communications**

MYLÈNE PISCHELLA

Allocation de ressources distribuée dans les réseaux OFDMA multi-cellulaires

Soutenue le 23 mars 2009 devant le jury composé de

Ezio Biglieri	Président
Mérouane Debbah	Rapporteurs
David Gesbert	
Mischa Dohler	Examinateurs
Jean-Claude Imbeaux	
Jean-Claude Belfiore	Directeur de thèse

TELECOM ParisTech THESIS

In Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
from Ecole Nationale Supérieure
des Télécommunications

Specialization: Communications and Electronics

Mylène Pischella

**Distributed resource allocation
in multi-cell OFDMA networks**

Defended on the 23rd March 2009. The committee in charge is formed of:

President	Prof. Ezio Biglieri, DTCI-UPF (Barcelona, Spain)
Reviewers	Prof. Mérouane Debbah, Supélec (Gif-sur-Yvette, France)
	Prof. David Gesbert, Institut Eurécom (Sophia Antipolis, France)
Examiners	Dr. Mischa Dohler, CTTC (Barcelona, Spain)
	Dr. Jean-Claude Imbeaux, France Télécom R&D (Issy les Moulineaux, France)
Thesis supervisor	Prof. Jean-Claude Belfiore, TELECOM ParisTech (Paris, France)

Remerciements

Je tiens tout d'abord à exprimer ma plus vive gratitude à mon directeur de thèse, Jean-Claude Belfiore, Professeur à TELECOM ParisTech, pour sa disponibilité, son enthousiasme, et l'orientation pertinente qu'il a donné à mes travaux. Je lui suis particulièrement reconnaissante pour la confiance qu'il m'a accordée en acceptant de diriger ces travaux, et pour les bonnes conditions dans lesquelles ma thèse s'est déroulée durant ces trois années.
Mes remerciements chaleureux s'adressent également à Mérouane Debbah, Professeur à Supélec, et David Gesbert, Professeur à Eurécom, pour avoir accepté d'être les rapporteurs de ma thèse. Merci pour leurs conseils et le recul qu'ils ont apporté sur ces travaux. Je tiens aussi à exprimer ma reconnaissance à Ezio Biglieri, Professeur à DTCI-UPF, pour m'avoir fait l'honneur de présider mon jury de thèse. Enfin, mes remerciements vont à Mischa Dohler, enseignant-chercheur à CTTC, et Jean-Claude Imbeaux, ingénieur de recherche à France Télécom Recherche et Développement, pour avoir accepté de faire partie de mon jury, et pour la bienveillance et la pertinence de leurs remarques.
Cette thèse a été initiée sous l'impulsion de l'ancien responsable de l'équipe PPR du laboratoire RESA/MNA de France Télécom R&D à Issy-les-Moulineaux, Franck Lebeugle. Je lui suis reconnaissante de la confiance témoignée et de sa volonté de développer de la recherche de qualité dans l'industrie. Merci aussi à mes collègues et anciens collègues qui ont soutenu ma démarche pendant ces trois années. Je tiens également à remercier les membres du laboratoire COMELEC de TELECOM ParisTech, qui m'ont fait un bon accueil lors de mes visites hebdomadaires, et qui ne se sont pas lassés de me voir attendre dans le couloir de bon matin.
Enfin, toute ma gratitude va à ma famille et mes amis pour leurs encouragements, et surtout à mon compagnon, Benjamin, dont la patience et le soutien sans faille ont été d'un grand réconfort pendant ces trois années.

Paris, le 23 janvier 2009
Mylène Pischella

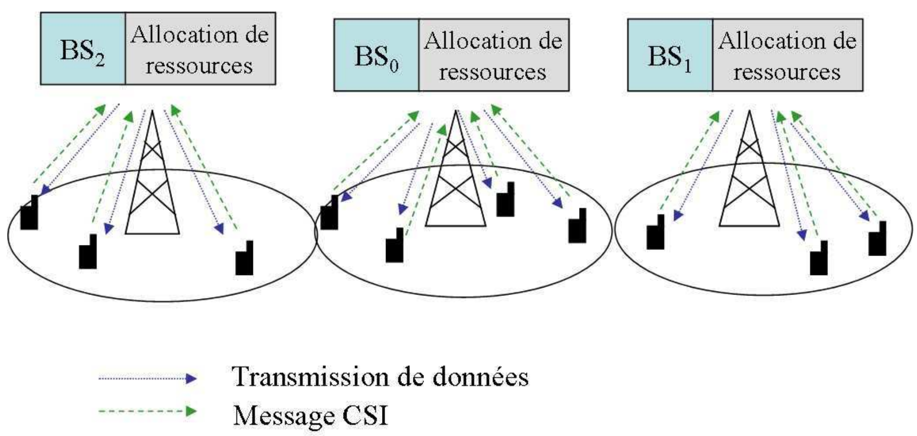
Abstract

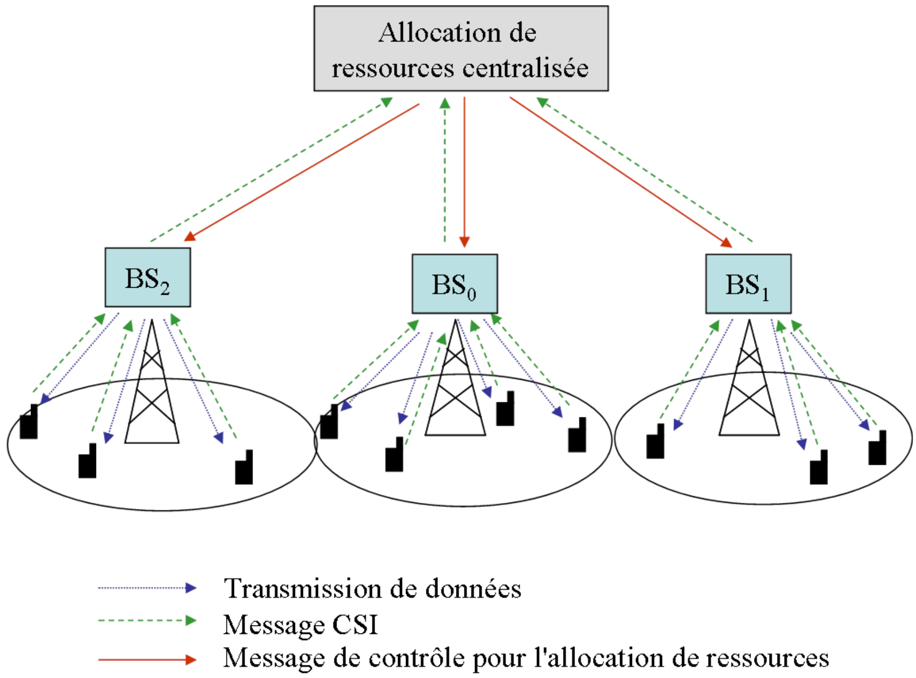
The thesis studies resource allocation methods, distributed per base station (BS) in multi-cellular OFDMA networks. The objective is to provide the Quality of Service (QoS) requested by each user, whatever its location in the cell. First, it investigates causal network coordination in distributed networks. Two BSs form a virtual Multiple-Input Multiple-Output (MIMO) array for the users located at the border of cells. These users thus benefit from a diversity gain, and from inter-cell interference mitigation. The efficiency of the associated resource allocation method depends on the fairness of the power control objective. Thus, network coordination is used for Rate Constrained (RC) users, but not for Best Effort (BE) users, in a proposed algorithm that jointly manages both QoS objectives. The thesis next considers the more general perspective of fully distributed networks. For RC users, a resource allocation process with iterative interference-based power allocation is determined to solve the Margin Adaptive problem. It includes a distributed constraint that guarantees power control convergence. The proposed method is extended to RC users in MIMO, both when full Channel State Information is available at transmission, and when only the statistical properties of the channel are available at transmission. Finally, for BE users, the objective is to maximize the weighted sum throughput, where the weight of each user is proportional to its queue length. A subcarrier allocation method, deduced from a network-wide interference graph, and a distributed power control method are proposed for that optimization problem.

Résumé des travaux de thèse

Introduction

Les réseaux cellulaires émergents et futurs reposent sur les hypothèses d'architecture plate et de transmission en mode tout IP. Ces nouvelles contraintes imposent la fonctionnalité d'allocation de ressources comme le point central entre la couche MAC (Medium Access Control) et la couche physique. En effet, l'allocation de ressources doit gérer l'accès aux ressources radio des différents utilisateurs, en fonction de leurs demandes de Qualité de Service (QdS) et de l'état du canal. Les réseaux cellulaires émergents, le WiMAX et le 3GPP LTE, utilisent tous deux l'OFDMA (Orthogonal Frequency Division Multiple Access) comme couche physique et méthode d'accès multiple. Dans les systèmes basés sur l'OFDMA, la bande passante est séparée en sous-porteuses orthogonales, chacune correspondant à un canal à bande étroite. De nombreux travaux ont été effectués dans le passé sur l'allocation opportuniste de ressources, qui nécessite la connaissance du canal dans chaque sous-porteuse. Ceci implique qu'une information sur l'état du canal, CSI (Channel State Information), doit être remontée de chaque terminal mobile à l'instance fonctionnelle responsable de l'allocation de ressources, avec le délai le plus court possible. Afin de minimiser ce délai, il a été décidé par les instances de normalisation du 3GPP LTE que l'allocation de ressources serait effectuée par chaque station de base. Ceci correspond à une contrainte dite d'architecture plate, dans laquelle l'allocation de ressources est distribuée par station de base.





Une telle allocation de ressources distribuée s'oppose à l'allocation plus classique, centralisée, dans laquelle un contrôleur global type RNC est responsable de l'allocation de ressources pour un ensemble de cellules adjacentes. L'allocation centralisée est utilisée dans les réseaux UMTS R99. Son problème principal est le délai induit dans l'allocation de ressources. Elle permet néanmoins d'optimiser globalement sur les cellules adjacentes, et de gérer ainsi les problématiques d'interférence inter-cellulaires pour les utilisateurs en bordure de cellules.
Une autre contrainte des réseaux émergents est la transmission en mode tout IP. Elle impose que l'ensemble des utilisateurs partagent les même ressources radio, et que l'allocation dynamique des ressources soit responsable de gérer la QdS de tous les utilisateurs. En OFDMA dans le sens descendant, l'allocation de ressources se compose de l'allocation de sous-porteuses, qui doit être orthogonale entre les utilisateurs d'une même cellule, et de la distribution de la puissance totale sur les différentes sous-porteuses.
Le principal désavantage de l'allocation de ressources distribuée consiste en l'absence de coordination de l'interférence inter-cellulaire. Les utilisateurs localisés en bordure de cellule sont particulièrement exposés à cette interférence, qui risque de fortement dégrader leurs performances. En conséquence, les réseaux à architecture plate ne pourront être considérés comme cellulaires que s'ils permettent à chaque utilisateur d'atteindre sa QdS, et ce quelle que soit sa localisation dans le réseau. Dans les réseaux tout IP, les utilisateurs les plus impactés par l'interférence sont les utilisateurs demandant un service temps réel qui sont positionnés en bordure de cellule. Néanmoins, comme les ressources radio sont partagées entre les différents utilisateurs, l'allocation de ressources ne doit pas se focaliser sur une seule catégorie de QdS. De même, l'interférence inter-cellulaire que chaque cellule engendre ou subit doit être considérée dans l'allocation de ressources, et ce bien qu'une allocation totalement centralisée ne soit pas réalisable. De nouvelles méthodes d'allocation de ressources doivent être déterminées, afin de tenir compte de ces contraintes dans les réseaux distribués.

Objectifs de la thèse

La thèse cherche à définir de telles méthodes, qui permettent de diminuer l'interférence inter-cellulaire de façon efficace, tout en maintenant un faible coût en termes de complexité et de signalisation entre les stations de base. Les réseaux OFDMA sont considérés, de sorte que les méthodes proposées peuvent s'appliquer au WiMAX, au 3GPP LTE, et aux futurs réseaux 4G qui seront basés sur l'OFDMA. Certains sujets traités dans la thèse peuvent aussi s'appliquer aux réseaux de type ad hoc. Les objectifs principaux de la thèse sont :

- Evaluer la faisabilité et la pertinence de la coordination de réseaux dans les réseaux distribués. La coordination de réseaux étudiée dans la thèse considère deux contraintes imposées par l'architecture distribuée : une contrainte de causalité sur la transmission entre stations de base, et une contrainte d'allocation de ressources distribuée.

- Déterminer une méthode d'allocation de ressources distribuée pour les utilisateurs demandant un service temps réel, qui doivent atteindre un débit cible. Cette méthode doit tenir compte de l'interférence inter-cellulaire en l'absence d'informations globales. La faisabilité de ce problème a été étudiée dans la littérature avec un seul canal et en SISO (Single-Input Single-Output). Nous l'étudions ici en multi-canal, en SISO et en MIMO (Multiple-Input Multiple-Output).

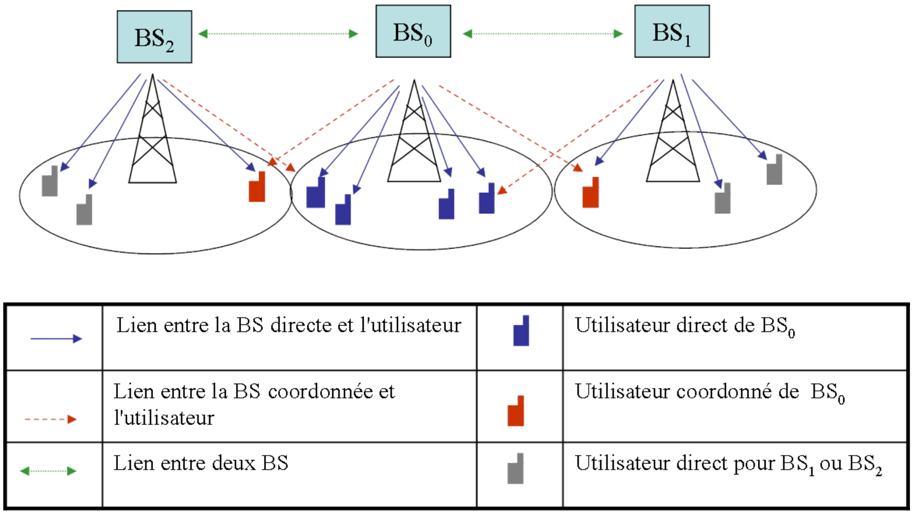
- Etudier l'allocation de ressources pour les utilisateurs Best Effort, en tenant compte de la longueur de leurs files d'attente. Une allocation de ressources utilisant la variable temporelle est permise pour ces utilisateurs. Nous nous proposons de caractériser si la transmission conjointe de deux utilisateurs est plus efficace que leur transmission séparée pour atteindre un certain objectif de QdS. L'allocation de sous-porteuses temporelle peut alors être utilisée, en plus du contrôle de puissance.

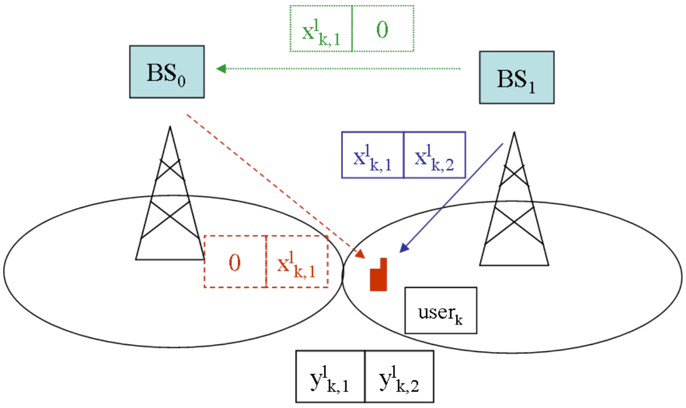
Dans tous les cas étudiés, l'allocation de ressources correspond à un objectif d'optimisation global sur le réseau, qui est décomposé par station de base, en raison de la contrainte d'architecture distribuée.
Les hypothèses principales prises dans la thèse sont les suivantes : l'allocation de ressource a lieu dans chaque intervalle de temps (TTI), et il est supposé que le canal reste inchangé durant le TTI; toutes les stations de base sont parfaitement synchrones; la constellation du signal est Gaussienne, avec un niveau de modulation suffisamment élevé pour que l'information mutuelle atteigne quasiment la capacité du canal; enfin, l'information de canal est parfaitement connue en réception et en transmission, sauf dans la deuxième partie du chapitre 6.
On se place dans des réseaux multi-cellulaires OFDMA, dans lesquels le paramétrage de la modulation OFDM est tel que chaque sous-porteuse est un canal à bande étroite sur lequel le canal rapide est plat et quasi-statique. Dans l'ensemble de la thèse, l'interférence inter-cellulaire est traitée comme du bruit. Il est à noter que des traitements sur le canal à interférence permettent de parvenir à une région de capacité plus grande, mais ces traitement complexes supposent que l'interférence soit parfois utilisée de façon constructive. Nous n'étudions pas ces méthodes dans la thèse.
Tous les problèmes d'allocation de ressources considérés s'écrivent comme des problèmes d'optimisation. A cause des contraintes d'architecture, de la complexité, et de la nécessité de tenir compte de l'interférence inter-cellulaire dans les allocations distribuées, les problèmes d'optimisation doivent être décomposés en plusieurs sous-problèmes. Dans la plupart

des cas, on cherchera à obtenir des sous-problèmes convexes, qui pourront être résolus avec des méthodes classiques d'optimisation convexes. Dans la thèse, on utilisera à la fois des méthodes numériques, comme la méthode de Newton, et des méthodes analytiques, comme les conditions de Karush-Kuhn-Tucker (KKT).
Dans la suite, nous détaillons les travaux menés dans chaque chapitre. Deux axes de recherche principaux sont considérés pour l'allocation de ressources : tout d'abord, lorsque la coordination de réseaux est possible, et ensuite, lorsqu'une allocation totalement distribuée est imposée. Le premier axe de recherche correspond aux chapitres 3 et 4. Le second est étudié dans les chapitres suivants. Dans l'ensemble de la thèse, on s'intéresse à l'allocation de ressources dans le sens descendant dans un intervalle de temps TTI, et le canal est supposé parfaitement connu en transmission, sauf dans la seconde partie du chapitre 6. Deux types d'utilisateurs sont considérés : les utilisateurs à Débit Contraint (DC), qui demandent une application de type temps réel, et pour lesquels un certain débit cible doit être atteint dans chaque intervalle de temps, et les utilisateurs Best Effort (BE), qui demandent une application n'ayant aucune contrainte de QdS spécifique par TTI, lorsque la longueur des files d'attente n'est pas prise en compte.

Chapitre 3 - Augmentation de l'équité par la coordination de réseaux

Le premier axe d'étude porte sur la coordination dans les réseaux distribués. On suppose que deux stations de base (BS) peuvent coordonner leur transmission afin de servir le même utilisateur. Cette coordination est contrainte par l'architecture qui impose que la transmission de données entre les deux BSs soit causale, et que chacune des BSs impliquées effectue l'allocation de ressources séparément, en fonction de l'information de canal remontée par le terminal mobile concernant les deux liens. Par ailleurs, on suppose que Rapport-Signal-à-Bruit (RSB) du canal entre deux BSs est parfait. Les deux BSs forment alors un canal MIMO virtuel. Afin de diminuer le coût de la coordination, seuls les utilisateurs localisés en bordure de cellule sont coordonnés. La coordination de réseaux permet d'augmenter la diversité et de diminuer l'interférence inter-cellulaire des utilisateurs concernés, car la BS voisine qui serait potentiellement la plus interférente est choisie pour la coordination. Un exemple détaillé de procédure de coordination pour un utilisateur est donné par la figure suivante, où BS_1 est la station de base directe, et BS_0 est la station de base coordonnée de l'utilisateur.





Avec cette méthode de coordination de réseaux causale sur deux TTI, le débit de l'utilisateur k dans la sous-porteuse l ayant une bande passante B_{SC} est égal à

$$
\begin{aligned}
\mathrm{R}_k^l &= \frac{B_{\mathrm{SC}}}{2}\log_2\left(\det\left(\mathbf{I}_2 + \mathbf{H}_k^l(\mathbf{H}_k^l)^H\right)\right) \\
&= \frac{B_{\mathrm{SC}}}{2}\log_2\left(\left(1+\frac{G_{d,k}^l P_{d,k}^l}{I_k^l}\right)^2 + \frac{G_{c,k}^l P_{c,k}^l}{I_k^l}\right)
\end{aligned}
$$

où $\mathbf{H}_k^l$ est le canal MIMO virtuel équivalent, $P_{d,k}^l$ (resp., $P_{c,k}^l$) est la puissance transmise par sa BS directe (resp., coordonnée) vers l'utilisateur k dans la sous-porteuse l, $G_{d,k}^l$ (resp., $G_{c,k}^l$) est le gain entre la BS directe (resp., coordonnée) de l'utilisateur k et lui-même dans la sous-porteuse l, incluant les pertes de propagations moyenne, les effets de masques, et le fast fading, et I_k^l est la somme du bruit et des interférences reçues par l'utilisateur k dans la sous-porteuse l. Le débit total de l'utilisateur k, R_k, est obtenu en sommant R_k^l sur l'ensemble des sous-porteuses qui lui sont allouées, à la fois sur le lien direct et sur le lien coordonné, dans l'ensemble Θ_k.

La méthode proposée consiste tout d'abord à identifier les utilisateurs qui doivent être coordonnés par chaque BS, puis à effectuer l'allocation de ressources. Dans ce chapitre, on s'intéresse principalement au contrôle de puissance. Les utilisateurs coordonnés sont les utilisateurs dont la distance entre le gain moyen avec sa BS directe et sa BS voisine la plus proche est inférieure à Δ dB. Cette valeur doit être paramétrée afin de ne coordonner que les utilisateurs localisés en bordure de cellule. La même méthode d'allocation de sous-porteuses, ayant pour objectif d'allouer le même nombre de sous-porteuses à tous les utilisateurs d'une cellule, est utilisée quel que soit l'objectif du contrôle de puissance. Elle alloue les mêmes sous-porteuses à un utilisateur coordonné à la fois sur son lien direct, et sur son lien coordonné.

Quatre objectifs, correspondant à des degrés d'équité variables entre utilisateurs, sont testés: Globally Optimal, Max-Min Fair, Proportional Fair et Harmonic Mean Fair. Le contrôle de puissance utilise un processus itératif, qui a lieu par BS de façon indépendante. Dans chaque BS, l'allocation de puissance est décomposée en deux processus parallèles : d'une part, l'allocation de puissance pour les utilisateurs direct, et d'autre part, l'allocation de puissance pour les utilisateurs coordonnés. Chaque sous-problème d'allocation de puissance est localement convexe, car l'interférence inter-cellulaire est fixée à une valeur donnée qui a été déterminée à l'itération précédente. Il s'écrit de la façon suivante (par exemple ici pour l'ensemble des utilisateurs directs d'une même cellule, $\mathcal{S}_d$) :

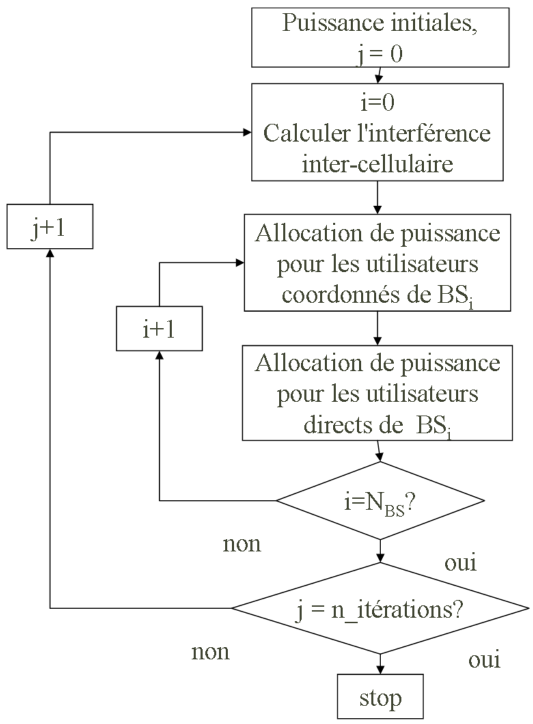
$$\begin{aligned} &\max_{\mathbf{P}_d} \sum_{k \in \mathcal{S}_d} \frac{(R_k)^{1-\alpha}}{1-\alpha} \\ &\text{s.t. } \sum_{k \in \mathcal{S}_d} \sum_{l \in \Theta_k} P_{d,k}^l = P_{\max,d} \\ &\text{s.t. } P_{d,k}^l \geq 0, \ \forall k \in \mathcal{S}_d, \ l \in \Theta_k \end{aligned} \tag{0.1}$$

α est le coefficient d'équité du problème, qui vaut:

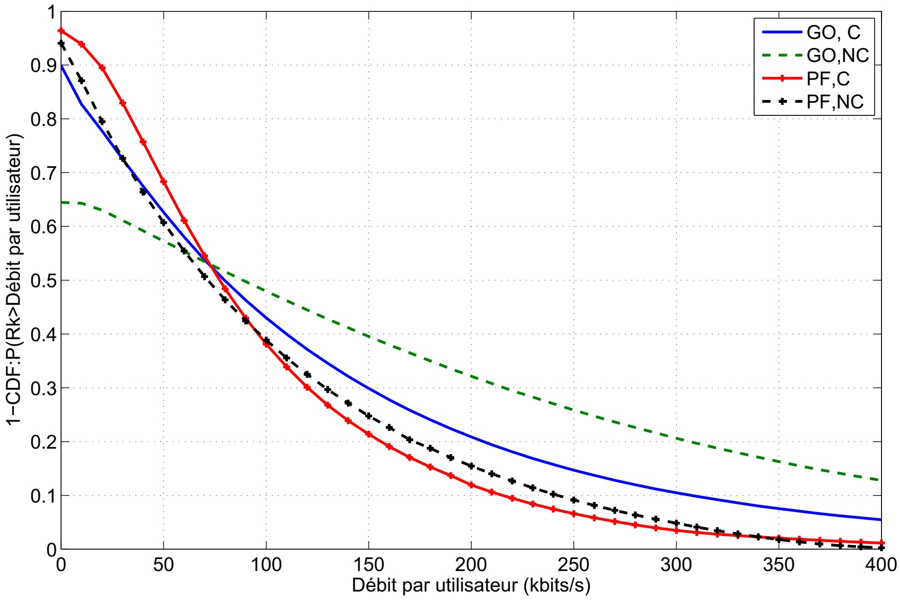
- 0 pour Globally Optimal (GO),
- 2 pour Harmonic Mean Fair (HMF),
- qui tend vers l'infini pour Max-Min Fair (MM),
- qui tend vers 1 pour Proportional Fair (PF).

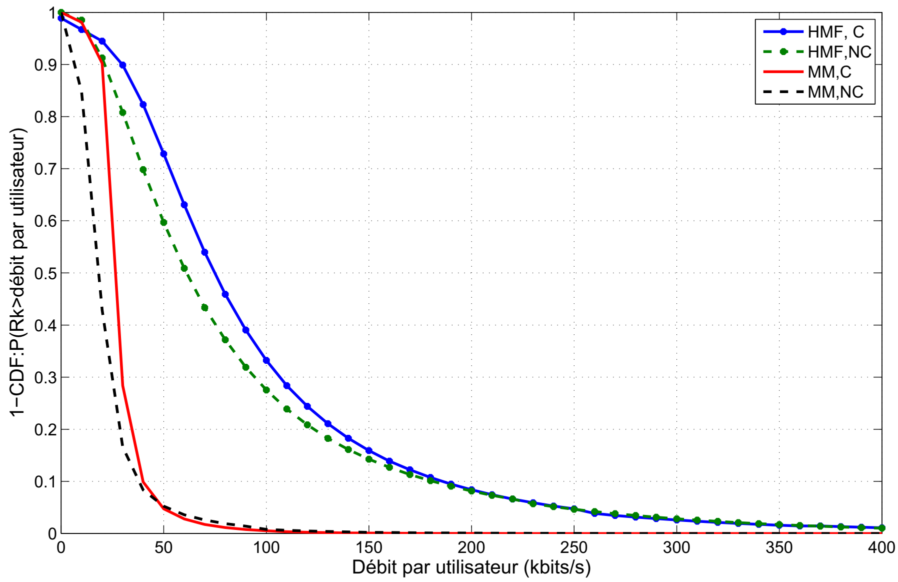
Chaque problème d'optimisation convexe possède un unique optimum global. Il est obtenu numériquement avec la méthode de Newton.

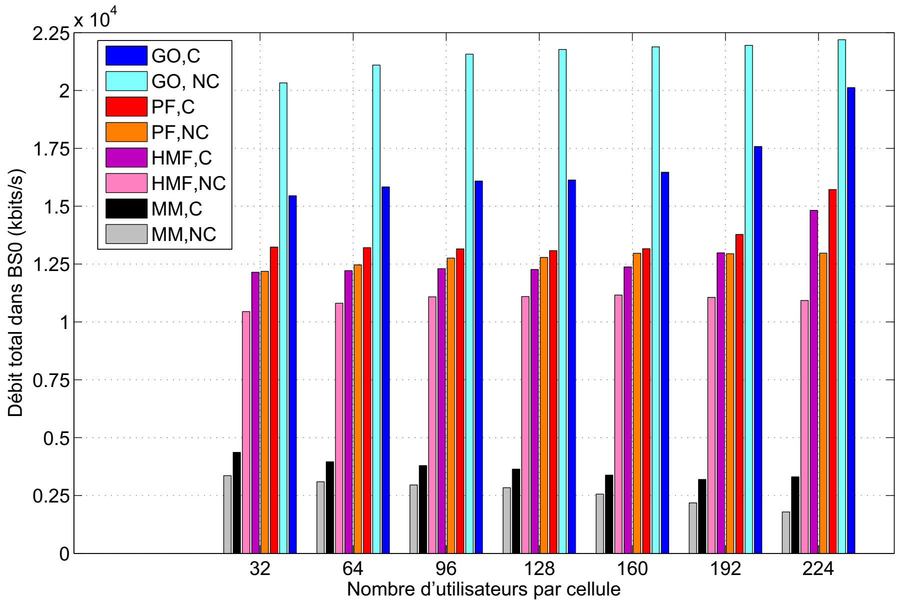
Le contrôle de puissance itératif est illustrée par la figure suivante.



Dans ce chapitre, l'influence de la coordination de réseaux sur les performances, en fonction du niveau d'équité de l'objectif de contrôle de puissance, est évaluée numériquement. Les résultats de simulations montrent que la coordination de réseaux permet d'améliorer l'équité, quel que soit le niveau de charge, et quel que soit l'objectif de contrôle de puissance. Ceci est obtenu par une augmentation du débit des utilisateurs localisés en bordure de cellule qui sont coordonnés. En conséquence, la somme des débits dans le réseau est augmentée lorsque l'objectif de contrôle de puissance est équitable, car dans ce cas, les débits de tous les utilisateurs dépendent des débits des utilisateurs localisés en bordure de cellule. Cependant, lorsque l'objectif de contrôle de puissance est inéquitable, le débit-crête est diminué par la coordination de réseaux, car le gain de performances fourni aux utilisateurs en bordure de cellules est obtenu au détriment des utilisateurs qui sont dans les meilleures conditions radio. En conséquence, la somme des débits par cellule est diminuée par la coordination causale de réseaux avec l'objectif Globally Optimal, contrairement aux autres objectifs d'allocation de puissance. Nous pouvons en déduire que la coordination de réseaux doit être réservée aux utilisateurs en bordure de cellule dont l'objectif d'allocation de ressources est équitable.


GO, C
GO,NC
PF,C
PF,NC
1−CDF:P(Rk>Débit par utilisateur)
Débit par utilisateur (kbits/s)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
50
100
150
200
250
300
350
400


HMF, C
HMF,NC
MM,C
MM,NC
1−CDF:P(Rk>débit par utilisateur)
Débit par utilisateur (kbits/s)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
50
100
150
200
250
300
350
400



Chapitre 4 - Allocation de ressources dépendant de la QdS avec coordination de réseaux distribués

Les conclusions du chapitre précédent montrent que la coordination de réseaux est efficace pour les utilisateurs à Débit Contraint (DC), dont l'objectif d'allocation de ressources est d'atteindre individuellement un débit donné par intervalle de temps, mais qu'elle est inefficace pour les utilisateurs Best Effort (BE), dont l'objectif d'allocation, Globally Optimal, est inéquitable. En conséquence, dans ce chapitre, une méthode pour satisfaire un mélange d'utilisateurs ayant l'un ou l'autre des objectifs de QdS est déterminée. La coordination de réseaux est restreinte aux utilisateurs DC localisés en bordure de cellule. Le chapitre propose tout d'abord un état de l'art sur l'allocation de ressources en OFDMA. Pour les utilisateurs DC, l'allocation de ressources peut s'effectuer selon un objectif dit 'Rate Adaptive', qui est équivalent à l'objectif Max-Min Fair, ou selon un objectif 'Margin Adaptive' (MA). Ce dernier consiste à minimiser la somme de puissances requises pour atteindre le débit cible de chaque utilisateur.

La plupart des articles de la litérature portent sur l'allocation de ressource en OFDMA dans le cas d'une cellule isolée. Le problème de la gestion de l'interférence inter-cellulaire dans l'allocation de ressources a été peu étudié, et uniquement résolu par des heuristiques. Dans le cas d'une cellule isolée, il a été démontré que, lorsque l'objectif de l'allocation de ressources est Globally Optimal, séparer l'allocation de sous-porteuses de l'allocation de puissance est optimal. De plus, l'allocation de sous-porteuses en OFDMA (c'est à dire, en autorisant la transmission d'un unique utilisateur par sous-porteuse) est elle aussi optimale, comparé à une allocation en OFDM, qui autoriserait la transmission de plusieurs utilisateurs dans une même sous-porteuse. Néanmoins, pour les problèmes Rate Adaptive et Margin Adaptive en OFDMA, l'allocation de ressources ne peut pas se décomposer en deux étapes successives (allocation de sous-porteuses puis de puissance) comme pour Globally Optimal.

L'allocation conjointe portant à la fois sur les sous-porteuses et sur la puissance est un problème qui n'est pas convexe, car l'allocation de sous-porteuses et un problème discret, tandis que le contrôle de puissance est un problème continu. Elle peut être résolue par relâchement de la contrainte d'allocation discrète portant sur l'allocation de sous-porteuses, qui transforme le problème en un problème convexe. Cependant, une étape finale est nécessaire pour déterminer quel utilisateur unique doit transmettre dans chaque sous-porteuse, la méthode conjointe pouvant mener à un partage des sous-porteuses par utilisateur. Cette dernière étape est sous-optimale, et mène à des performances de l'algorithme qui ne sont pas significativement meilleures que ce qui est obtenu en séparant l'allocation de sous-porteuses de l'allocation de puissance. Ces conclusions nous indiquent qu'il est préférable, pour une raison de compromis entre la complexité des algorithmes et le gain obtenu, d'effectuer l'allocation de sous-porteuses et l'allocation de puissance en série. Ceci est d'autant plus nécessaire en multi-cellulaire, lorsque l'interférence inter-cellulaire doit être considérée.
Dans ce chapitre, les utilisateurs DC sont caractérisés par un objectif MA, tandis que les utilisateurs BE sont caractérisés par un objectif Globally Optimal. L'allocation de ressources ayant lieu dans le sens descendant, les utilisateurs DC et BE sont liés par une contrainte de puissance totale par cellule. Le problème d'allocation de ressources par cellule est le suivant:

$$\begin{aligned}
&\min_{\{\mathbf{P},\mathbf{\Theta}\}} \sum_{k\in\mathcal{S}_{\mathrm{DC},d}} \sum_{l\in\Theta_k} \left(P_{d,k}^l + P_{c,k}^l\right)\\
&\text{s.t. } \sum_{l\in\Theta_k} R_k^l \geq R_{k,\mathrm{cible}},\ \forall k\in\mathcal{S}_{\mathrm{DC},d}\\
&\max \sum_{k\in\mathcal{S}_{\mathrm{BE}}} R_k\\
&\text{s.t. } \sum_{l=1}^{L_{\mathrm{SC}}} P_{n_{\mathrm{BS}}}^l \leq P_{\max}\\
&\text{s.t. } P_{n_{\mathrm{BS}}}^l \geq 0, \forall l\in\{1,...,L_{\mathrm{SC}}\}\\
&\text{s.t. } \Theta_k\cap\Theta_{k'} = \emptyset, \forall (k,k')\in\{\mathcal{S}_{\mathrm{DC},d}\cup\mathcal{S}_{\mathrm{BE}}\}^2, k\neq k'
\end{aligned} \tag{0.2}$$

où $\mathcal{S}_{\mathrm{DC},d}$ est l'ensemble des utilisateurs DC directs servis par la station de base n_{BS}, $\mathcal{S}_{\mathrm{DC},c}$ est l'ensemble des utilisateurs DC coordonnés par n_{BS}, and $\mathcal{S}_{\mathrm{BE}}$ est l'ensemble des utilisateurs BE directs servis par n_{BS}.

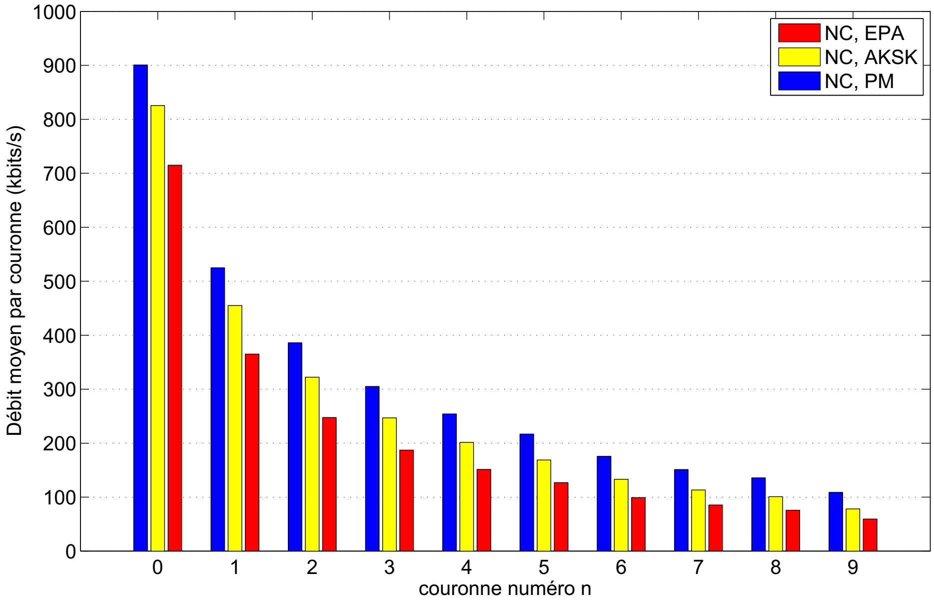
L'allocation de sous-porteuses et le contrôle de puissance sont utilisés pour optimiser l'allocation de ressources. La méthode proposée donne la priorité aux utilisateurs DC par rapport aux utilisateurs BE, à la fois pour l'allocation de sous-porteuse et pour le contrôle de puissance. Ces deux étapes sont effectuées de façon séparée, en séquence, pour des raisons de complexité. L'allocation de sous-porteuses pour les utilisateurs DC a pour objectif de minimiser la somme des puissances requises pour atteindre les débits cibles, sous les hypothèses simplificatrices que la puissance et l'interférence par sous-porteuses sont équitablement réparties sur toutes les sous-porteuses. S'il reste des sous-porteuses libres à l'issue de cette allocation, celles-ci sont distribuées aux utilisateurs BE, avec pour objectif de maximiser le gain direct. Le contrôle de puissance pour les utilisateurs DC suit lui aussi un objectif MA, dans lequel la minimisation de la somme des puissances porte à la fois sur la puissance transmise par la BS directe de chaque utilisateur, mais aussi par sa BS coordonnée. La contrainte de puissance totale par cellule n'est pas considérée dans le problème MA, afin de simplifier le problème. Il en résulte que le contrôle de puissance peut s'effectuer de façon distribuée, indépendamment pour chaque utilisateur. Le problème d'optimisation correspondant,

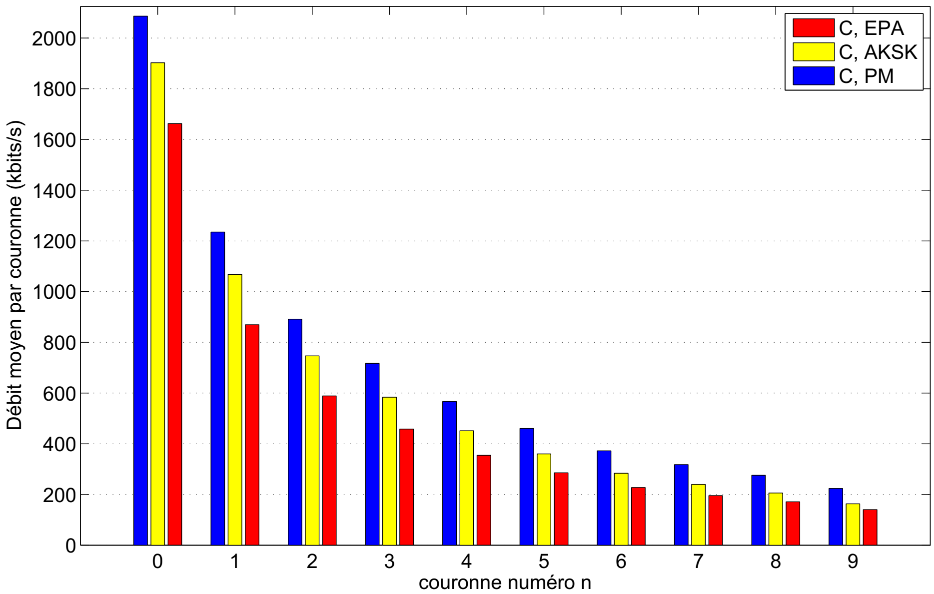
$$\begin{aligned}
&\min_{\{\mathbf{P}_{d,k},\mathbf{P}_{c,k}\}} \sum_{l\in\Theta_k} \left(P_{d,k}^l + P_{c,k}^l\right) \\
&\text{s.t. } \sum_{l\in\Theta_k} R_k^l \geq R_{k,\text{cible}} \\
&\text{s.t. } P_{c,k}^l \geq 0 \text{ and } P_{d,k}^l \geq 0, l \in \Theta_k
\end{aligned} \tag{0.3}$$

est convexe sous certaines conditions, et peut dans tous les cas se résoudre avec les conditions KKT. A l'issue du contrôle de puissance pour les utilisateurs DC, s'il reste de la puissance disponible sur la cellule, celle-ci est distribuée aux utilisateurs BE suivant un objectif d'allocation Globally Optimal. Si la somme des puissance excède le maximum autorisé, alors le contrôle d'admission rejette les utilisateurs DC qui nécessitent les niveaux de puissance les plus élevés.

	Avec coordination	Sans coordination
PM (puissance requise)	1.95 W	5.19 W
AKSK (évalué par l'algorithme)	11.07 W	16.15 W
EPA (puissance requise)	10.99 W	10.31 W

Il est montré par des résultats numériques que la coordination de réseaux pour les utilisateurs DC localisés en bordure de cellule permet de diminuer la somme des puissances requises pour satisfaire ces utilisateurs, grâce au gain en diversité et à la diminution de l'interférence inter-cellulaire. Ceci est représenté par la table suivante, dans laquelle la méthode proposée, appelée 'PM', est comparée à deux autres algorithmes: 'AKSK', pour lequel la puissance allouée aux utilisateurs DC est fixée en fonction de la proportion de sous-porteuses qui sont allouées à ces utilisateurs, et pour lequel l'allocation de puissance suit le même objectif MA qu'avec notre méthode, et enfin 'EPA', dans lequel l'allocation de puissance est équirépartie sur l'ensemble des sous-porteuses. Le scénario testé correspond à 16 utilisateurs DC avec un débit cible de 64 kbits/s, et 16 utilisateurs BE par cellule. De plus, même lorsque la coordination de réseaux n'est pas utilisée, la méthode proposée de priorisation des utilisateurs entraîne une augmentation du pourcentage d'utilisateurs DC qui atteignent leur débit cible, et une augmentation de la somme des débits pour les utilisateurs BE. La méthode de priorisation est encore plus efficace lorsque la coordination de réseaux est utilisée, grâce à la diminution de puissance obtenue pour les utilisateurs DC.





Chapitre 5 - Allocation de ressources distribuée pour les utilisateurs à Débit Contraint

Dans la suite de la thèse, on s'intéresse aux réseaux totalement distribués, dans lesquels aucune information sur les données à transmettre n'est échangée entre stations de base, et dans lesquels l'allocation de ressources est distribuée par cellule. On étudie tout d'abord le problème d'allocations de ressources MA pour les utilisateurs DC. Ce problème consiste à minimiser la somme des puissances nécessaires pour atteindre le débit cible de chaque utilisateur :

$$\begin{aligned}
&\min_{\{\mathbf{P},\mathbf{\Theta}\}} \sum_{k=1}^{K_{\mathcal{N}}} \sum_{l\in\Theta_k} P_k^l \\
&\text{s. t. } \sum_{l\in\Theta_k} R_k^l \geq R_{k,\text{cible}}, \forall k \in \{1,...,K_{\mathcal{N}}\} \\
&\text{s. t. } \sum_{l=1}^{L_{\text{SC}}} P_{n_{\text{BS}}}^l \leq P_{\max}, \forall n_{\text{BS}} \in \{1,...,N_{\text{BS}}\} \\
&\text{s. t. } \sum_{l=1}^{L_{\text{SC}}} P_{n_{\text{BS}}}^l \geq 0, \forall (n_{\text{BS}}, l) \in \{1,...,N_{\text{BS}}\} \times \{1,...,L_{\text{SC}}\} \\
&\text{s.t. } \Theta_k \cap \Theta_{k'} = \emptyset, \forall (k,k'), k \neq k' \text{ servis par la même BS}
\end{aligned} \tag{0.4}$$

L'objectif DC peut ne pas être réalisable pour un ensemble d'utilisateurs interférents, ce qui entraîne une divergence du contrôle de puissance itératif. Les conditions de convergence du contrôle de puissance ont été caractérisées dans la littérature pour le canal à interférences, avec une seule sous-porteuse, lorsque le Rapport-Signal-à-Bruit-plus-Interférence (RSBI) cible de chaque utilisateur est connu. Dans ce chapitre, un critère de convergence distribué pour la canal multi-porteuses est déterminé. Il permet au contrôle de puissance distribué de converger dans les réseaux multi-cellulaires OFDMA. Il correspond à une borne supérieure par sous-porteuse sur le RSBI cible, que l'on appelera 'E', et qui dépend des gains de canal avec les stations de base interférentes :

$$E_k^l = \frac{G_{k,k}^l}{\sum_{\{n=1,n\neq k\}}^{N_{\text{BS}}} G_{n,k}^l}$$

On montre que lorsque le RSBI cible par utilisateur et par sous-porteuse, γ_k^l, est strictement inférieur à E_k^l, le problème d'allocation de puissance par sous-porteuse possède une solution optimale unique, qui est atteinte lorsque chaque utilisateur met à jour sa puissance en fonction de l'interférence qu'il reçoit et de son RSBI cible, lors d'un contrôle de puissance itératif et totalement distribué sur le réseau.

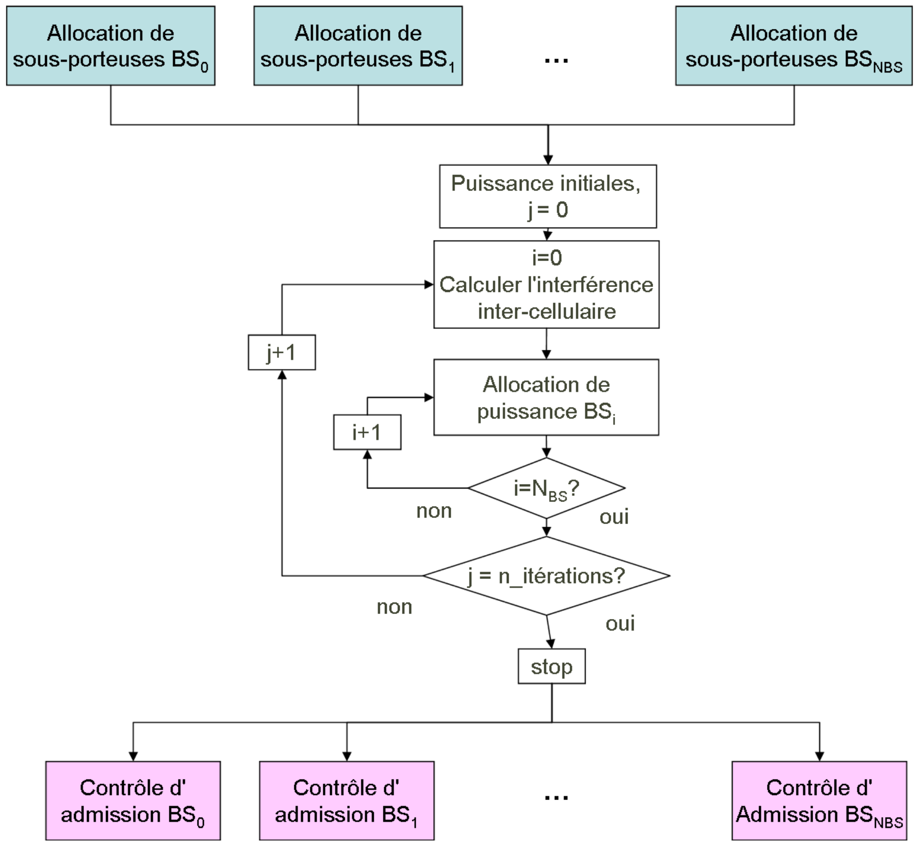
Le critère $\gamma_k^l < E_k^l$ est valide lorsque l'interférence inter-cellulaire est suffisamment élevée par rapport au niveau de bruit. En conséquence, il est utilisé de façon adaptative dans l'allocation de ressources pour résoudre le problème MA. Le problème d'allocation de puissance étant résolu de façon itérative, à chaque itération, et dans chaque sous-porteuse, le niveau d'interférence reçu lors de l'itération précédente est évalué. S'il est négligeable, alors le critère E n'est pas appliqué dans cette sous-porteuse.
La méthode d'allocation de ressources proposée est composée de trois étapes :

1. Allocation de sous-porteuses, distribuée par cellule.

2. Contrôle de puissance : composé d'allocations de puissances locales par cellule, et itéré sur l'ensemble des cellules afin de tenir compte de l'interférence inter-cellulaire.

3. Contrôle d'admission, distribué par cellule.

Elle est illustrée par la figure suivante.



L'allocation de sous-porteuses est tout d'abord effectuée sur chaque cellule de façon indépendante avec une méthode itérative, qui a pour but de minimiser la somme des puissances requises pour atteindre les débits cibles de chaque utilisateur, sous les hypothèses de puissance également répartie, et de niveau d'interférence équivalent sur chaque sous-porteuse. Les sous-porteuses sélectionnées pour chaque utilisateur maximisent le critère E proposé, ce qui permettra par la suite d'obtenir un intervalle de variations plus large pour le RSBI cible. L'allocation de puissance consiste ensuite à déterminer l'ensemble de RSBIs cibles par utilisateur, et à lancer le contrôle de puissance itératif sur chaque sous-porteuse en parallèle. On ne tient pas compte de la contrainte de puissance maximale par cellule pendant le contrôle de puissance, ce qui permet de décomposer l'allocation de puissance par utilisateur. A l'issue du contrôle de puissance, si la puissance totale d'une cellule dépasse le maximum autorisé, alors les utilisateurs qui demandent le plus de puissance sont rejetés par le module de contrôle d'admission. Le problème MA par utilisateur est équivalent au problème de minimisation suivant, qui porte sur les RSBI cibles :

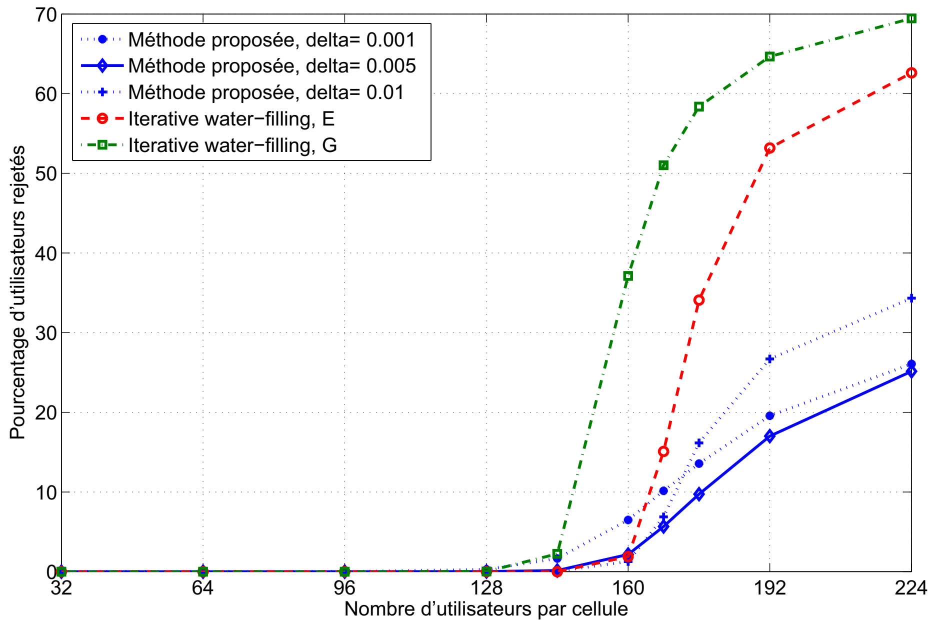
$$
\begin{aligned}
&\min_{\gamma_k} \sum_{l\in\Theta_k} \gamma_k^l \left(\frac{I_k^l}{G_{k,k}^l}\right)\\
&\text{s. t. } B_{\mathrm{SC}} \sum_{l\in\Theta_k} \log_2(1+\gamma_k^l) \geq R_{k,\text{cible}}\\
&\text{s. t. } \gamma_k^l \geq 0, \forall l \in \Theta_k\\
&\text{s. t. } E_k^l - \epsilon \geq \gamma_k^l \text{ if } E_k^l \tau_k^l \geq \delta, \forall l \in \Theta_k
\end{aligned}
\tag{0.5}
$$

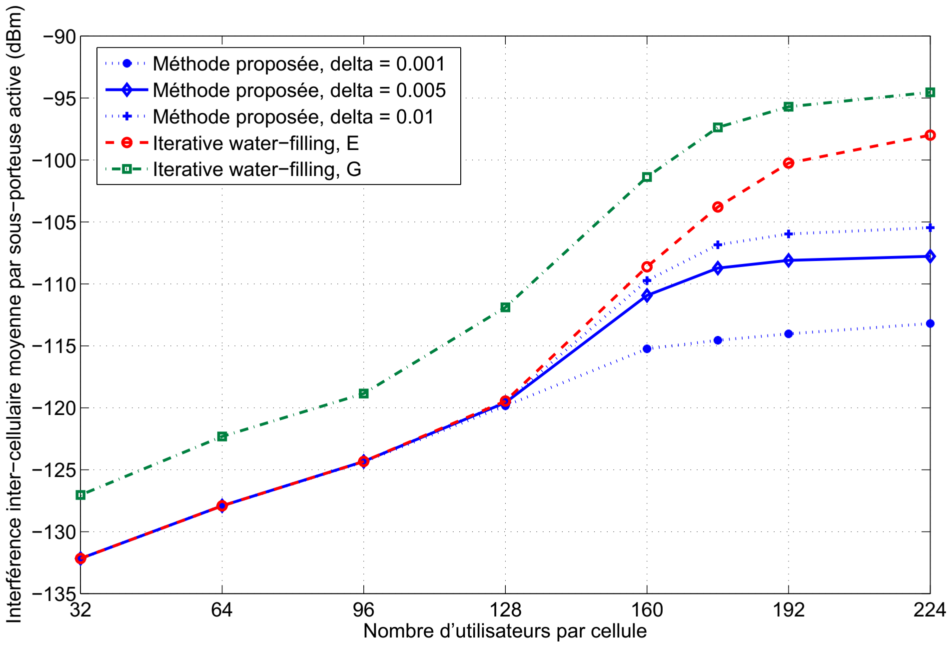
Où $E_k^l \tau_k^l \geq \delta$ est le critère pour évaluer si le niveau d'interférence est suffisamment élevé pour être pris en compte, et ϵ est une petite valeur strictement positive. Ce problème d'optimisation est convexe en γ_k et se résoud avec les KKT. Le contrôle de puissance itératif est donc composé des quatre étapes suivantes :

1. Initialisation : toutes les valeurs sont mises à 0.

2. Itération pour déterminer les RSBI cibles : pour chaque utilisateur, en fonction des valeurs de puissances reçues à l'itération précédente, calcul de γ_k afin de résoudre le problème MA.

3. Mise à jour des valeurs d'interférence inter-cellulaire, avant d'aller à l'itération suivante pour les calcul des RSBI cibles.





La méthode proposée est comparée à l'iterative water-filling. Les résultats numériques montrent qu'elle permet d'éviter les situations de divergence du contrôle de puissance qui ont lieu avec l'iterative water-filling lorsque le niveau de charge est moyen ou élevé. Elle permet aussi de diminuer le nombre de sous-porteuses utilisées et la puissance requise, et s'adapte de façon efficace au niveau d'interférence. Cette méthode est, en conséquence, une alternative possible à l'iterative water-filling pour satisfaire les utilisateurs DC dans les réseaux distribués.

Chapitre 6 - Allocation de ressources distribuée pour les réseaux MIMO

Ce chapitre étend les résultats du chapitre 5 sur l'allocation de ressources distribuées pour les utilisateurs DC au cas des réseaux MIMO. Chaque station de base et chaque terminal mobile sont équipés de plusieurs antennes. Le nombre d'antennes en transmission est n_t, et le nombre d'antennes en réception est n_r, et $n_{\min} = \min\{n_t, n_r\}$. Le problème d'optimisation MA est étudié dans deux cas : lorsque toute l'information de canal est connue en transmission (CSIT parfait), et lorsque seules les caractéristiques statistiques du canal sont connues en transmission (CSIT statistique).
Dans le premier cas, un critère E distribué permettant d'assurer la convergence du contrôle de puissance est appliqué sur le débit par sous-porteuse. Cette borne supérieure sur le débit est obtenue sous les hypothèses d'interférence inter-cellulaire moyenne et d'allocation de puissance équirépartie sur tous les flux. Elle est égale à $R_k^l < R_{k,\max}^l$ où $R_{k,\max}^l = B_{\mathrm{SC}} n_{\min} \log_2\left(1 + E_k^l\right)$ et

$$E_k^l = \frac{\left(\beta_{k,n_{\min}}^l\right)^2 g_{k,k}}{\sum_{\{n=1, n\neq k\}}^{N_{\mathrm{BS}}} g_{n,k}}$$

($\beta^l_{k,n_{\min}}$) est valeur singulière la plus élevée du canal direct. Le contrôle de puissance distribué consiste alors à déterminer, de façon itérative, la matrice de précodage qui résoud le problème MA, l'interférence inter-cellulaire étant fixée à la valeur de la précédente itération. Il est décomposé sur chaque utilisateur, car la contrainte de puissance totale n'est pas prise en compte dans le contrôle de puissance, mais dans une étape de contrôle d'admission qui a lieu ensuite. Les variables d'optimisation de ce problème sont la puissance par utilisateur et par sous-porteuse, et la matrice de précodage sur chaque sous-porteuse. Le problème d'optimisation par utilisateur est alors :

$$\begin{aligned}
&\min_{\left\{\mathbf{P_k},\boldsymbol{\Phi}_k^{\theta_k(1)},..,\boldsymbol{\Phi}_k^{\theta_k(l_{\mathrm{SC},k})}\right\}} \sum_{l\in\Theta_k} P_k^l \\
&\text{s.t. } \sum_{l\in\Theta_k} R_k^l \geq R_{k,\mathrm{cible}} \\
&\text{s.t. } P_k^l \geq 0, \forall l \ \in \Theta_k \\
&R_k^l \leq R_{k,\max}^l - \epsilon \ \text{ if } \ E_k^l \tau_k^l \geq \delta, \forall l \in \Theta_k
\end{aligned} \tag{0.6}$$

Un récepteur MMSE étant utilisé, le débit par sous-porteuse est égal à

$$R_k^l = B_{\mathrm{SC}} \sum_{j=1}^{n_{\min}} \log_2 \left(1 + \rho_k d_{k,j}^l (\lambda_{k,j}^l)^2\right)$$

où ρ_k est le RSB de l'utilisateur k tenant compte de la somme des puissances sur toutes les sous-porteuses, $d^l_{k,j}$ est la puissance normalisée de l'utilisateur k dans la sous-porteuse l et le flux j, et $\left\{\lambda^l_{k,1}, ..., \lambda^l_{k,n_{\min}}\right\}$ sont les valeurs singulières du canal équivalent $(\mathbf{Q}^l_k)^{-1/2}\mathbf{H}^l_{k,k}$, $\mathbf{Q}^l_k$ étant la matrice de covariance de l'interférence et du bruit, et $\mathbf{H}^l_{k,k}$ étant la matrice de canal direct de l'utilisateur k.
Le problème d'optimisation est convexe et se résoud avec les KKT, de façon équivalente sur l'ensemble des RSB par sous-porteuse et par flux, $\left\{\rho_k \mathbf{d}^{\theta_k(1)}, ..., \rho_k \mathbf{d}^{\theta_k(l_{\mathrm{SC},k})}\right\}$. La matrice de précodage optimale est ainsi obtenue par water-filling sur les sous-porteuses et sur les flux, en tenant compte de la contrainte de convergence du contrôle de puissance.

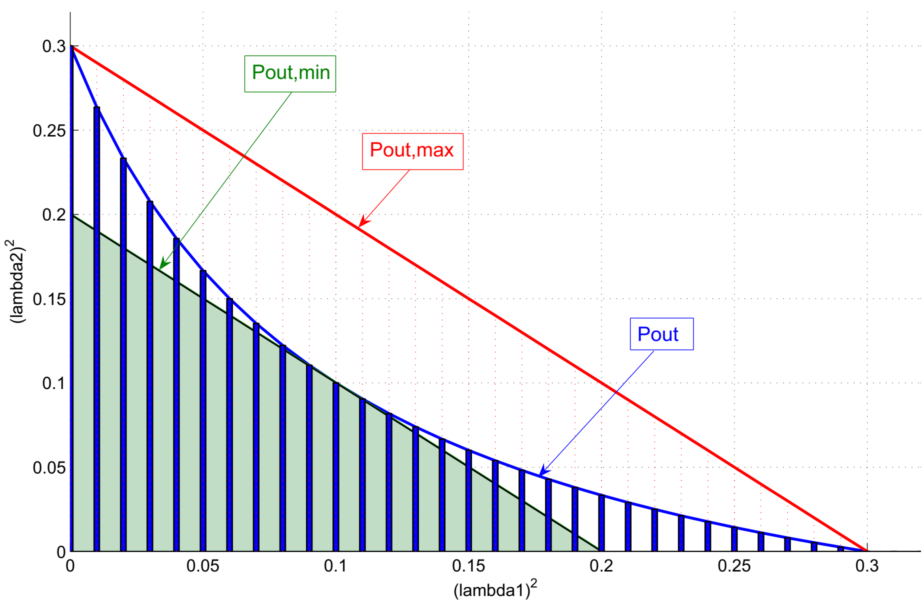
Dans le deuxième cas (CSIT statistique), on détermine une expression analytique approchée de la capacité d'outage, qui est fonction du RSB et de la probabilité d'outage. La probabilité d'outage du canal MIMO Rayleigh $\mathbf{y} = \mathbf{Hx} + \mathbf{n}$ est :

$$P_{\mathrm{out}} = \mathcal{P}\left(\prod_{i=1}^{n_{\min}} \left(1 + \frac{\rho}{n_t}{\lambda_i}^2\right) < 2^C\right)$$

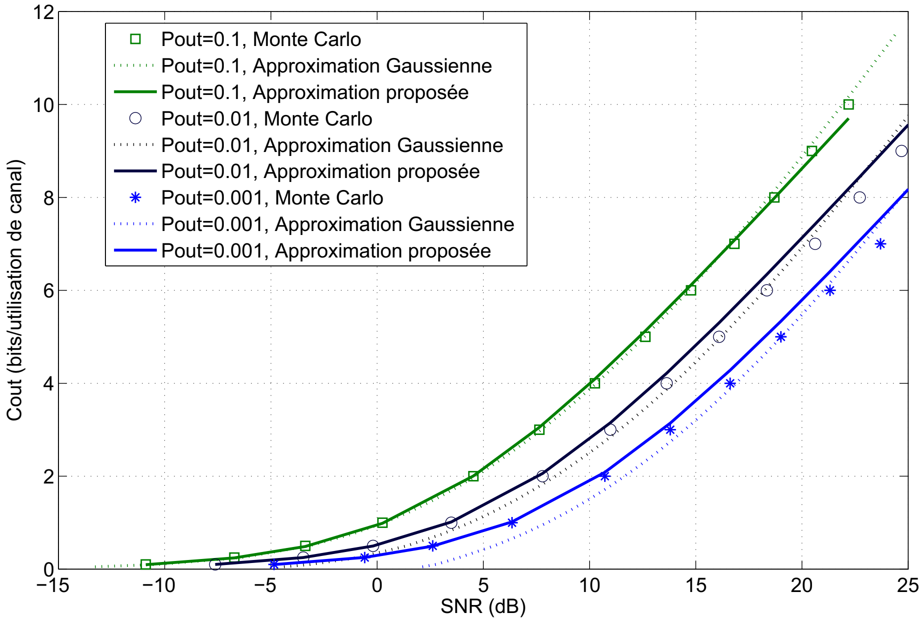
où ρ est le RSB, et $\{\lambda_i\}_{\{1\leq i\leq n_{\min}\}}$ sont les valeurs singulières du canal $\mathbf{H}$, et C est le débit, en bits par utilisation de canal. La capacité d'outage pour une probabilité d'outage cible $P_{\text{out, cible}}$ est définie comme le débit C maximum permettant d'atteindre une probabilité d'outage $P_{\text{out}} \leq P_{\text{out, cible}}$. La formule analytique recherchée doit être concave en le RSB, afin d'exprimer le problème d'allocation de puissances Margin Adaptive comme un problème d'optimisation convexe. Une borne supérieure sur la capacité d'outage est obtenue à partir de l'inégalité entre les moyennes arithmétique et géométrique. Une borne inférieure sur la capacité d'outage est obtenue en ne tenant compte que de la trace dans la formule de la capacité. Dans les deux cas, les bornes sur $\prod_{i=1}^{n_{\min}}\left(1+\frac{\rho}{n_t}{\lambda_i}^2\right)$ sont des fonctions linéaires en $\left\{{\lambda_i}^2\right\}_{\{1\leq i\leq n_{\min}\}}$ qui s'expriment sous la forme d'une somme des ${\lambda_i}^2$. Comme cette somme suit une loi du chi-2, on peut exprimer les probabilités d'outage qui en découle comme des fonctions Gamma, puis extraire les bornes sur la capacité d'outage par inversion de ces fonctions pour des valeurs de P_{out} proches de zéro. L'expression analytique de la capacité d'outage est alors égale à la moyenne arithmétique des bornes supérieure et inférieure :

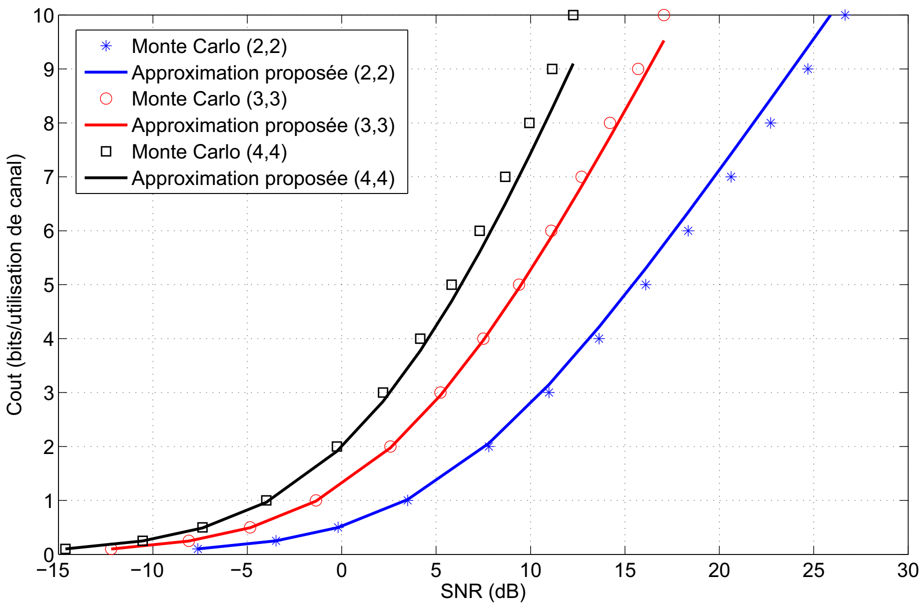
$$\tilde{C}=\frac{1}{2}\left(n_{\min}\log_2\left(1+\frac{f(P_{\text{out}})}{n_{\min}n_t}\rho\right)+\log_2\left(1+\frac{f(P_{\text{out}})}{n_t}\rho\right)\right)$$

où $f(P_{\text{out}})$ dépend du nombre d'antennes et de la probabilité d'outage cible. Les bornes sur la probabilité d'outage utilisées pour obtenir cette expression analytique sont représentées sur la figure suivante.



Il est montré que cette formule est très proche des résultats obtenus par tirages de Monte-Carlo pour des valeurs pratiques de RSBs et de probabilités d'outage, et ce pour plusieurs configurations antennaires.



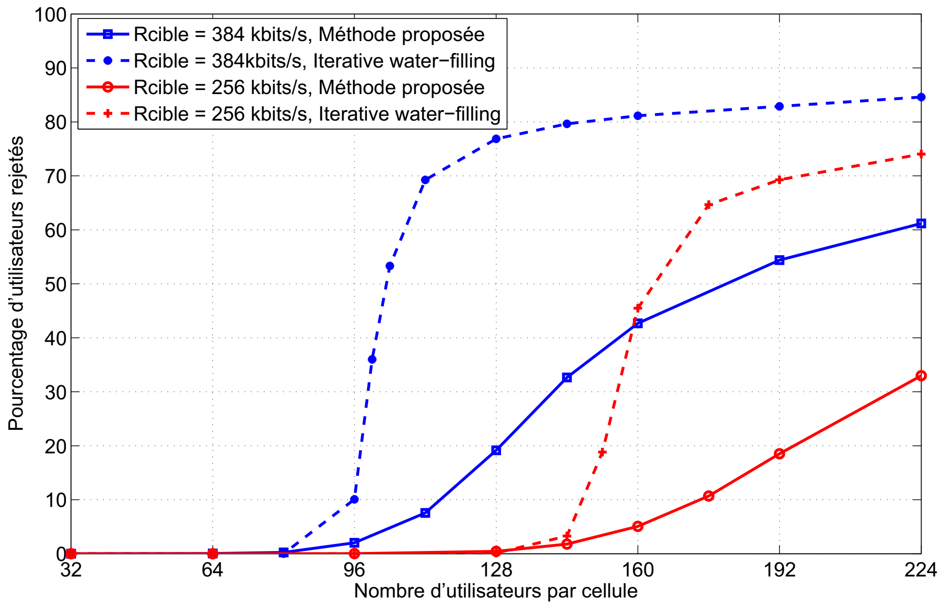


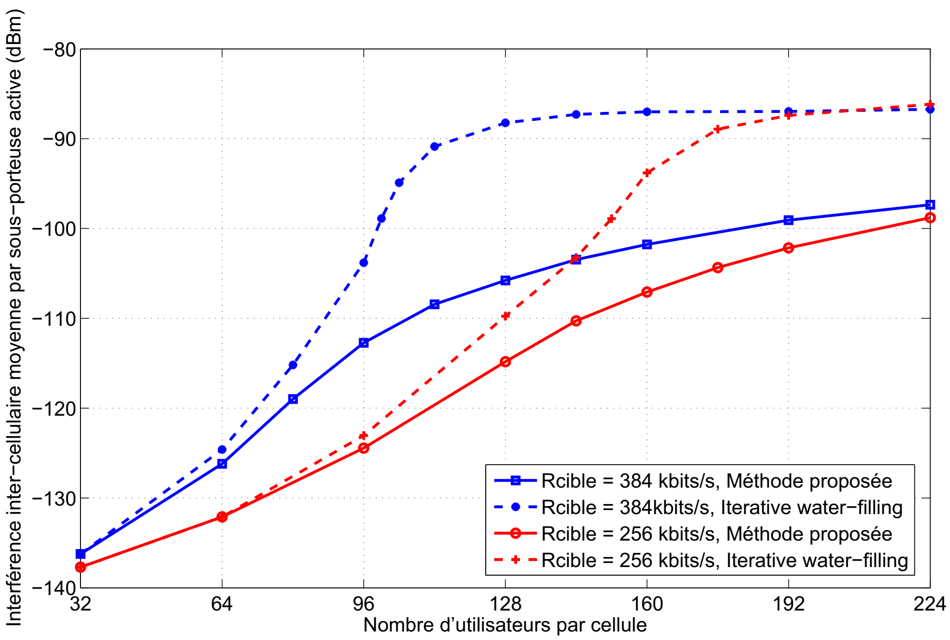
De plus, cette formule permet d'exprimer la capacité d'outage pour une probabilité d'outage donnée comme une fonction concave du SNR, et le problème MA devient alors un problème d'optimisation convexe. On en déduit une méthode de contrôle de puissance itérative qui inclut le critère de convergence E distribué suivant :

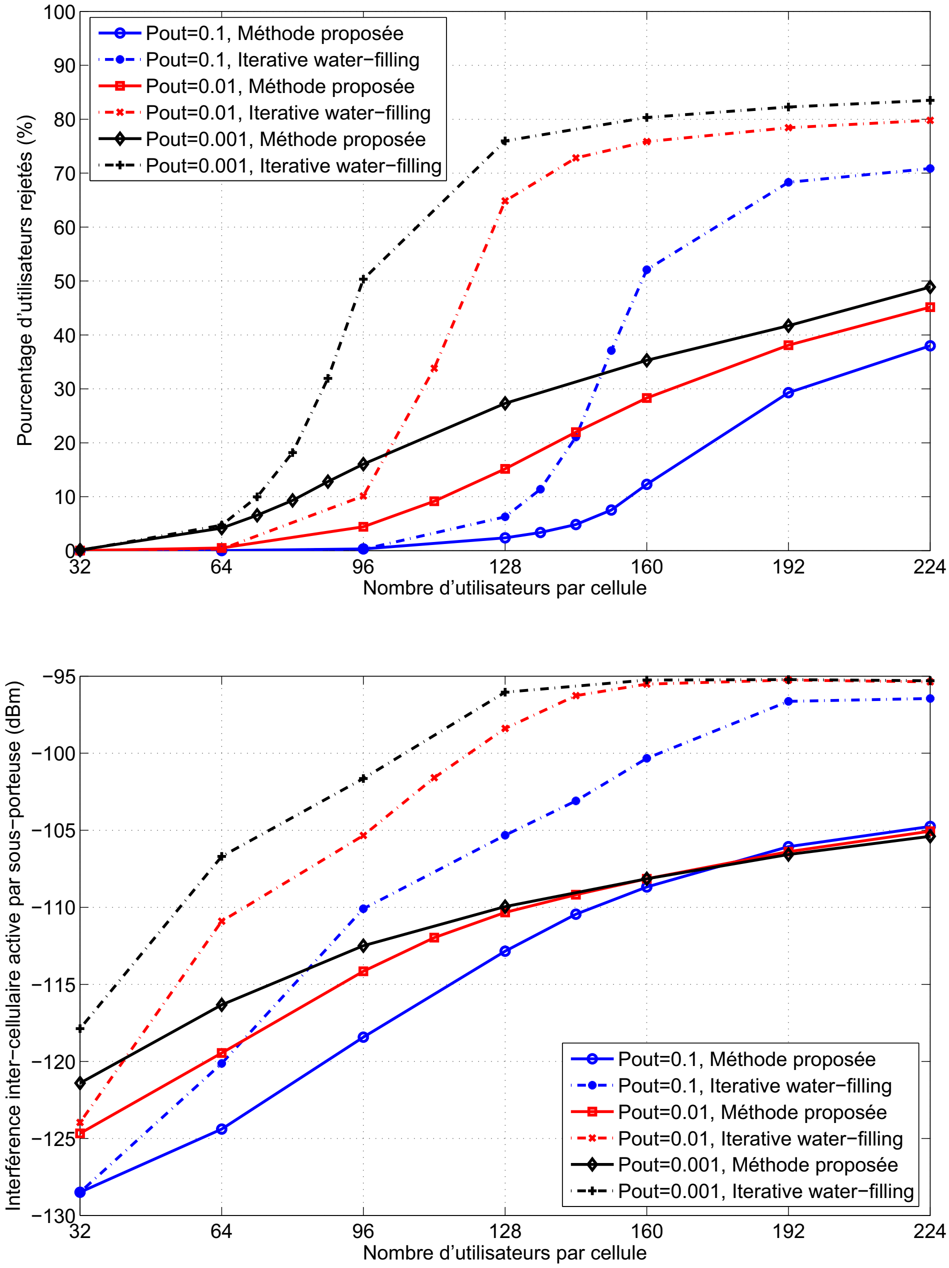
$$E_k = \frac{f(P_{\text{out}})g_{k,k}}{n_{\min}\sum_{\{n=1,n\neq k\}}^{N_{\text{BS}}} g_{n,k}} \tag{0.7}$$

Ce critère est obtenu en ne tenant compte que la borne supérieure sur la capacité d'outage.

Le problème MA par utilisateur est donc un problème d'optimisation convexe portant sur la puissance par sous-porteuse. Il est résolu avec les conditions KKT.
Les cas CSIT parfait et CSIT statistique sont testés par des simulations numériques, et comparés à l'iterative water-filling. Il est montré que les méthodes proposées évitent les situations de divergence du contrôle de puissance qui ont lieu avec l'iterative water-filling à charges moyenne et forte. En conséquence, ces méthodes permettent à plus d'utilisateurs DC d'atteindre leur débit cible, et elles diminuent les niveaux d'interférence et de puissance moyens du réseau.





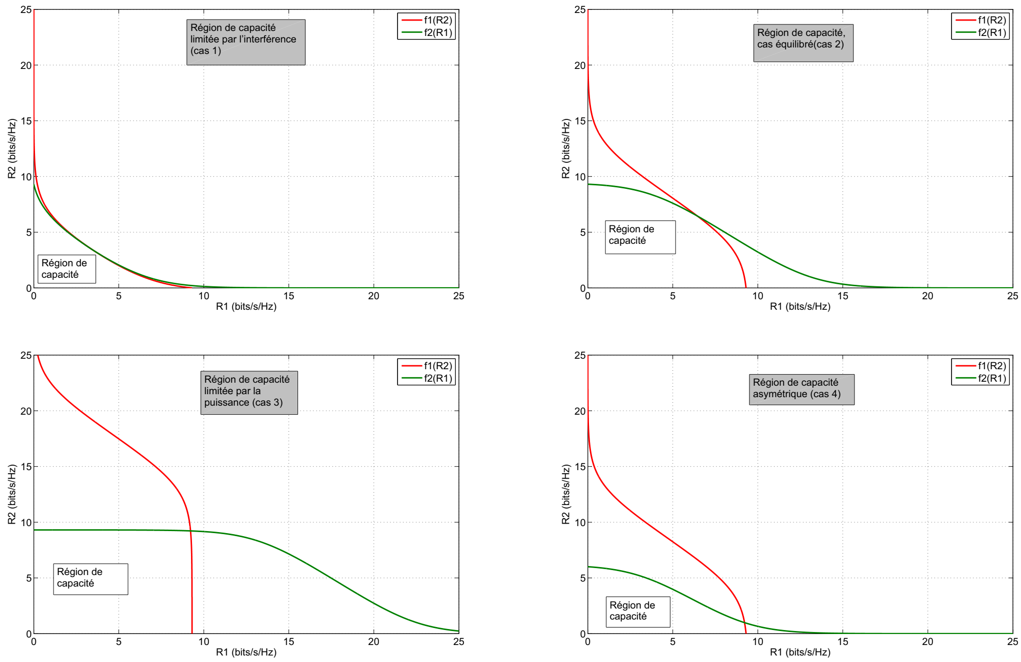


Chapitre 7 - Maximisation de la somme des débits pondérés dans les réseaux multi-cellulaires

Dans ce chapitre, on s'intéresse à l'allocation de ressources pour les utilisateurs BE dans les réseaux OFDMA multi-cellulaires. Pour ces utilisateurs, un objectif de QdS possible consiste à éviter que des paquets ne soient perdus dans le buffer, ce qui arrive lorsque les files d'attente dépassent leur taille maximale. Le problème de la maximisation de la somme des débits pondérés (WSTM), où le poids de chaque utilisateur est proportionnel à la longueur de sa file d'attente, est un compromis entre la maximisation de la somme des débits, et l'équilibrage des longueurs des files d'attente, nécessaire pour éviter toute perte de paquets. Il s'écrit de la façon suivante :

$$\begin{aligned} &\max_{\{\mathbf{P},\boldsymbol{\Theta}\}} \sum_{k=1}^{K_{\mathcal{N}}} w_k R_k \qquad (0.8)\\ &\text{s. t. } \sum_{l=1}^{L_{\text{SC}}} P_{n_{\text{BS}}}^{l} \leq P_{\max}, \forall n_{\text{BS}} \in \{1, ..., N_{\text{BS}}\}\\ &\text{s. t. } P_{n_{\text{BS}}}^{l} \geq 0, \forall (n_{\text{BS}}, l) \in \{1, ..., N_{\text{BS}}\} \times \{1, ..., L_{\text{SC}}\}\\ &\text{s.t. } \Theta_k \cap \Theta_{k'} = \emptyset, \forall (k, k'),\ \text{servis par la même BS}, k \neq k' \end{aligned}$$

où w_k est le poids de l'utilisateur k.
Afin de résoudre le problème WSTM, on s'intéresse d'abord à la région de capacité du canal à interférences avec deux utilisateurs. Une caractérisation simple des cas où la transmission simultanée des deux utilisateurs est plus efficace que leur transmission en séquence pour maximiser la somme des débits pondérés est obtenue. Elle dépend de la convexité de la région de capacité, et des poids relatifs de chaque utilisateur. Différents exemples de régions de capacité, convexes ou concaves, sont représentés sur la figure suivante.



Il est démontré que la somme des débits pondérés est plus élevée lorsque les deux stations de base transmettent que lorsque l'une des deux ne transmet pas, sous les conditions suivantes :

$$A_{k,n} \geq 0,\ A_{n,k} \geq 0,\ \frac{w_k}{w_n} \geq m_n \text{ and } \frac{w_n}{w_k} \geq m_k$$

où $A_{k,n}$ et m_n dépendent des gains, de la puissance maximale par cellule et du bruit. $A_{k,n} \geq 0$ est la condition pour la fonction f_n, qui caractérise la dépendance du débit de

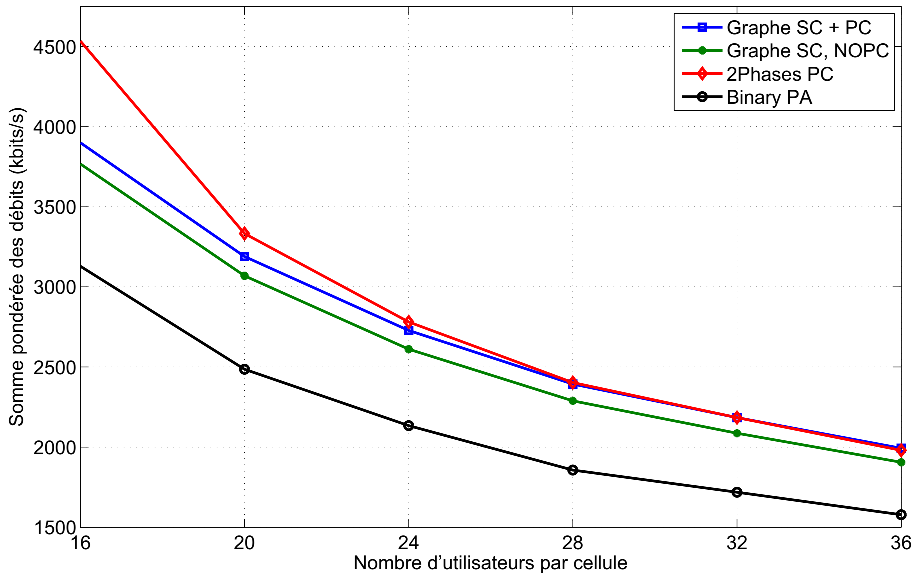
l'utilisateur n, R_n, par rapport au débit de l'utilisateur k, R_k, soit une fonction concave en 0. Dans le cas où $A_{k,n} \geq 0$ et $A_{n,k} \geq 0$, la région de capacité des utilisateurs n et k est convexe dans une certaine zone, ce qui implique que toute tangente à la région de capacité dans cette zone, de pente $-\frac{w_k}{w_n}$ (s'appuyant sur f_n), coupera l'axe des ordonnées en une valeur qui sera plus élevée que la valeur de f_n lorsque R_k est égal à 0. En conséquence, la somme des débits pondérés sera maximisée par la transmission conjointe des deux utilisateurs, correspondant à une $P_k^* \neq \{0, P_{\max}\}$ de la puissance de l'utilisateur k. La condition $\frac{w_k}{w_n}$ provient de la faisabilité des pentes. $\frac{w_k}{w_n} \geq m_n$ où $m_n = -f_n'(0)$ indique que w_k ne doit pas être trop faible par rapport à w_n, sinon $P_k = 0$ est optimal pour maximiser la somme des débits pondérés.
Ce critère est étendu au cas d'un réseau avec plus de deux cellules, et est utilisé pour construire le graphe d'interférence du réseau, qui permet ensuite de déterminer quels utilisateurs sont autorisés à transmettre simultanément dans la même sous-porteuse. Le graphe d'interférence est construit de sorte que deux utilisateurs dont la transmission conjointe maximise la somme de leurs débits pondérés peuvent transmettre simultanément. Il tient aussi compte de la contrainte d'orthogonalité entre utilisateurs d'une même cellule pour l'allocation de sous-porteuses. La transmission simultanée dans la même sous-porteuse de deux utilisateurs est interdite lorsqu'ils sont reliés dans le graphe d'interférence, dont les entrées valent 0 ou 1. Le graphe permet d'obtenir des sous-ensembles d'utilisateurs pour lesquels la transmission conjointe est permise. Il s'agit des utilisateurs qui ont la même couleur, à l'issue de la coloration du graphe, qui peut être obtenue avec des méthodes classiques comme l'algorithme DSATUR, ou encore avec des méthodes distribuées. L'allocation de sous-porteuses basée sur le graphe d'interférence est alors la suivante, pour les réseaux multi-cellulaires OFDMA :

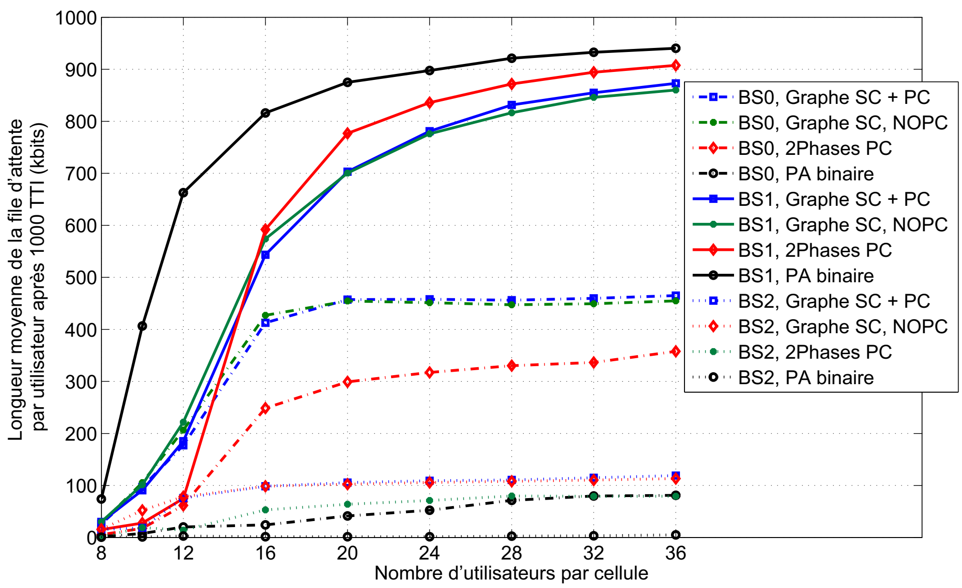
1. Sur chaque sous-porteuse, choisir l'utilisateur du réseau qui maximise le débit pondéré, dans lequel le débit est calculé en considérant le RSB seul. Cet utilisateur k se voit allouer la sous-porteuse dans sa cellule.

2. Puis sur cette même sous-porteuse, déterminer l'ensemble des utilisateurs des cellules adjacentes qui ont la même couleur que l'utilisateur k après coloration du graphe d'interférence. Sur chaque cellule, choisir l'utilisateur ayant la même couleur qui maximise le débit pondéré.

Le contrôle de puissance adapté au problème WSTM est ensuite étudié. Une méthode en deux étapes, qui est totalement distribuée et s'applique quel que soit le RSBI de chaque utilisateur, est proposée. La première étape consiste à effectuer le contrôle de puissance sur chaque sous-porteuse indépendamment. L'algorithme itératif utilisé, dans lequel chaque utilisateur tient compte de l'interférence qu'il génère sur les autres utilisateurs, peut converger vers un optimum local, car le problème d'optimisation correspondant n'est pas convexe. Néanmoins, les résultats numériques montrent que les utilisateurs dont la puissance tend vers zéro atteignent cette valeur en très peu d'itérations. A l'issue de cette première étape, un test est effectué par utilisateur et sous-porteuse, pour évaluer si le RSBI est élevé. Seuls les utilisateurs et les sous-porteuses vérifiant un critère de RSBI suffisant sont pris en compte dans la deuxième étape du contrôle de puissance. Celle-ci opère sur l'ensemble des sous-porteuses conjointement. Etant donnée l'hypothèse de RSBI élevé, le

problème d'optimisation est un problème de programmation géométrique, qui est équivalent à un problème d'optimisation convexe. La méthode itérative utilisée converge nécessairement vers l'optimum global de ce problème. Le contrôle de puissance en deux étapes que nous proposons est donc sous-optimal, mais il est totalement distribué.
Des simulations numériques sur le canal à interférence montrent que les résultats obtenus avec notre méthode et avec la méthode optimale, qui requiert une gestion totalement centralisée, sont très similaires, ce qui valide notre approche.





Enfin, l'allocation de sous-porteuses et le contrôle de puissance, utilisés conjointement ou séparément, sont testés avec des simulations numériques dynamiques. Dans ces simulations, le poids de chaque utilisateur est proportionnel à la longueur de sa file d'attente. L'allocation de ressources a lieu dans chaque TTI. Il est montré que lorsque le niveau de charge est fort ou moyen, les méthodes proposées permettent d'augmenter la somme des débits pondérés et d'obtenir des longueurs de files d'attente plus équitablement réparties entre les utilisateurs, par comparaison à l'allocation de puissance binaire. De plus, l'allocation de sous-porteuses basée sur le graphe d'interférence suivie du contrôle de puissance permet de diminuer de façon importante la consommation de puissance et de sous-porteuses.

Conclusion

En conclusion, il a été montré dans cette thèse que l'interférence inter-cellulaire peut être efficacement diminuée dans les réseaux distribués. Les trois techniques principales permettant d'atteindre ce résultat sont les suivantes :

- Coordonner l'accès aux ressources, afin de transformer l'interférence inter-cellulaire en signal utile. Cette technique est utilisée dans la coordination de réseaux.
- Eviter les situations dans lesquelles l'interférence est principalement responsable des dégradations des performances. Ceci peut être obtenu soit en utilisant l'allocation de sous-porteuses temporelle pour n'allouer simultanément que les utilisateurs qui s'interfèrent peu (cette méthode a été utilisée dans le chapitre 7 avec la définition du graphe d'interférence), soit en imposant une borne supérieure sur le débit par utilisateur et par sous-porteuse, qui permette de garantir qu'avec ces débits, l'interférence sera toujours maîtrisée (cette méthode a été utilisée dans les chapitres 5 et 6 pour les utilisateurs à Débit Contraint).
- Tenir compte de l'influence de l'interférence inter-cellulaire que chaque station de base génère sur les utilisateurs des cellules adjacentes, en équilibrant les niveaux de puissance entre cellules adjacentes afin d'éviter les situations menant à des niveaux d'interférence trop hétérogènes. Le contrôle de puissance itératif utilisant des informations d'interférence proposé dans le chapitre 7 remplit cet objectif.

Les résultats obtenus dans cette thèse pourraient être complétés par des travaux permettant de comparer l'efficacité des techniques de diminution de l'interférence inter-cellulaire présentées ici, avec d'autres techniques de gestion de l'interférence. Parmi celles-ci, l'alignement d'interférence pourrait être étudié, ainsi que les nouvelles méthodes proposant de transformer l'interférence en signal utile, par l'utilisation d'informations communes à tous les terminaux, transmises en plus de leur information privée.

Contents

List of Figures

List of Tables

Nomenclature

Abbreviations and Acronyms

3GPP	Third Generation Partnership Project
ADSL	Asymmetric Digital Subscriber Line
AF	Amplify and Forward
AKSK	Anas-Kim-Shin-Kim algorithm
AWGN	Additive White Gaussian Noise
BC	Broadcast Channel
BE	Best Effort
bits/c.u.	bits per channel use
BMI	Bilinear Matrix Inequality
BS	Base Station
C	network Coordination
CDF	Cumulative Distribution Function
CP	Cyclic Prefix
CSI	Channel State Information
CSIT	Channel State Information at Transmitter
DADP	Dual Asynchronous Distributed Pricing
DF	Decode and Forward
DFE	Decision Feedback
DFP	Differential Frequency Partitioning
DPP	Differential Power Partitioning
DSATUR	Degree SATURation
EPA	Equal Power Allocation
FDMA	Frequency Division Multiple Access
FRF	Frequency Reuse Factor
GO	Globally Optimal
HMF	Harmonic Mean Fair
HSPA	High Speed Packet Access
IEEE	Institute of Electrical and Electronics Engineers
i.i.d.	independent and identically distributed
(I) FFT	(Inverse) Fast Fourier Transform
INR	Interference-to-Noise Ratio
IP	Internet Protocol

ISI	Inter-Symbol-Interference
kbits	kilobits
kbits/s	kilobits per second
KKT	Karush-Kuhn-Tucker optimality conditions
LTE	Long Term Evolution
MA	Margin Adaptive
MAC	Multi-Access Channel
MIMO	Multiple-Input Multiple-Output
MISO	Multiple-Input Single-Output
ML	Maximum Likelihood
MMF	Max-Min Fair
MMSE	Minimum Mean Square Error
MSE	Mean Square Error
NAF	Non-orthogonal Amplify and Forward
NC	No network Coordination
NP-hard	Nondeterministic Polynomial-time hard
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
PA	Power Allocation
PC	Power Control
PF	Proportional Fair
PM	Power Minimization algorithm
QoS	Quality of Service
RA	Rate Adaptive
RC	Rate Constrained
RF	Radio Frequency
SC	Subcarrier Allocation
SIC	Successive Interference Cancellation
SINR	Signal-to-Interference-plus-Noise Ratio
SIR	Signal-to-Interference Ratio
SISO	Single-Input Single-Output
SIMO	Single-Input Multiple-Output
SNR	Signal-to-Noise Ratio
s.t.	subject to
SVD	Singular Value Decomposition
TCP	Transmission Control Protocol
TDMA	Time Division Multiple Access
TTI	Transmit Time Interval
UMTS	Universal Mobile Telecommunications System
WiMAX	Worldwide Interoperability for Microwave Access
WF	Water Filling
WSTM	Weighted Sum Throughput Maximization
ZF	Zero Forcing

Notations

Vectors and matrices are denoted by bold letters (e.g. $\mathbf{X}$). Other notational conventions are summarized as follows:

$\mathbb{C}^n$, $\mathbb{R}^n$	The sets of vectors of length n, with complex and real elements.
$\mathbb{R}^n_+$	The set of vectors of length n with real positive elements.
$\mathbb{R}^n_{++}$	The set of vectors of length n with real strictly positive elements.
$\mathbb{C}^{p\times n}$, $\mathbb{R}^{p\times n}$	The set of matrices with p rows and n columns, with complex and real elements.
$\mathbb{N}$	The set of positive integer elements.
$\log(\cdot)$	The natural logarithm.
$\log_2(\cdot)$	The base 2 logarithm.
$L(.)$	Lagrangian of an optimization problem.
∇f	The gradient of function f.
$\nabla^2 f$	The Hessian matrix of function f.
$\|x\|$	The absolute value of a scalar.
$\lfloor x \rfloor$	The floor operator, i.e. the smallest integer less than x.
$[x]^+$	max $\{0, x\}$.
$\|\Omega\|$	The cardinality of the set Ω, i.e. the number of elements in the finite set $\mathcal{X}$.
$\cap$	The intersection operator.
$\cup$	The union operator.
$\mathbb{E}\{\cdot\}$	The expectation operator.
$\mathcal{CN}(\mathbf{x}, \mathbf{X})$	The circularly symmetric complex Gaussian distribution with mean $\mathbf{x}$ and covariance matrix $\mathbf{X}$.
$\mathcal{N}(\mu, \sigma^2)$	The Gaussian distribution with mean μ and variance σ^2.
$(\cdot)^T$	The transpose operator.
$(\cdot)^H$	The complex conjugate (Hermitian) transpose operator.
$\mathbf{X}^{-1}$	The inverse of matrix $\mathbf{X}$.
$(\mathbf{X})_k$	The k^{th} row of matrix $\mathbf{X}$.
$(\mathbf{X})_{(k,n)}$	The element corresponding to the k^{th} and n^{th} column of matrix $\mathbf{X}$.
$\mathbf{I}_n$	The identity matrix of size n.
$\mathbf{0}_n$	The square zero matrix of size n.
$\mathbf{0}_{n_r\times n_t}$	The zero matrix of size $n_r \times n_t$.
$\mathrm{diag}\{x_1, ..., x_N\}$	The diagonal matrix containing elements $\{x_1, ..., x_N\}$.
$\mathrm{Tr}(\mathbf{X})$	The trace of matrix $\mathbf{X}$, i.e. the sum of the diagonal.
$\\|\mathbf{X}\\|$	Any norm of matrix $\mathbf{X}$.
$\\|\mathbf{X}\\|_\infty$	The infinity norm of matrix $\mathbf{X}$.
$\\|\mathbf{X}\\|_F^2 = \mathrm{Tr}(\mathbf{X}\mathbf{X}^H)$	The Frobenius norm of matrix $\mathbf{X}$.
$\rho(\mathbf{X})$	The spectral radius of matrix $\mathbf{X}$.

Thesis Specific Notations

The following list is not exhaustive, and consists of the most relevant notations used in the dissertation.

Notations applicable in all chapters:

$\mathcal{N}$	Considered network.
N_{BS}	Number of BSs in $\mathcal{N}$.
$K_{\mathcal{N}}$	Total number of users in $\mathcal{N}$.
N_{FFT}	OFDMA FFT size.
L_{SC}	Total number of subcarriers per BS.
B	Total available bandwidth.
B_{SC}	Bandwidth per subcarrier.
N_0	Variance of the AGWN noise.
$P_{\max}$	Maximum downlink transmit power per BS.
P_{out}	Outage capacity.
T_X	Transmitter.
R_X	Receiver.
n_t	Number of transmit antennas.
n_r	Number of receive antennas.
$\boldsymbol{I}(\mathbf{x};\mathbf{y}\vert\mathbf{H})$	Mutual information between $\mathbf{x}$ and $\mathbf{y}$ for a given channel realization $\mathbf{H}$.
$\mathbf{y}_k^l$	Receive vector for user k in subcarrier l.
$\mathbf{x}_k^l$	Transmit vector for user k in subcarrier l.
$\mathbf{n}_k^l$	AWGN noise vector for user k in subcarrier l.
R_k	Sum data rate of user k.
$R_{k,\mathrm{target}}$	Target data rate for RC user k.
R_k^l	Data rate of use k in subcarrier l.
C_k^l	Capacity of user k in subcarrier l.
Θ_k	Set of subcarriers allocated to user k.
$l_{\mathrm{SC},k}$	Number of subcarriers allocated to user k.
γ_k^l	SINR of user k in subcarrier l.
I_k^l	Noise plus interference received by user k in subcarrier l.
$P_{n_{\mathrm{BS}}}^l$	Power transmitted by BS n_{BS} in subcarrier l.
w_k	Weight of user k.

Notations specific to Chapters 3 and 4:

$BS_{d,k}$	Direct BS for user k.
$BS_{c,k}$	Coordinated BS for user k.
$P_{d,k}^l$	Power transmitted to user k by its direct BS in subcarrier l.
$P_{c,k}^l$	Power transmitted to user k by its coordinated BS in subcarrier l.
$h_{d,k}^l$	Fast fading coefficient between user k and its direct BS in subcarrier l.
$h_{c,k}^l$	Fast fading coefficient between user k and its coordinated BS in subcarrier l.
$g_{c,k}$	Path loss including shadowing between user k and its direct BS.

$g_{d,k}$	Path loss including shadowing between user k and its coordinated BS.
$\mathbf{H}_k^l$	Equivalent channel matrix for user k in subcarrier l.
$G_{d,k}^l$	Channel gain between user k and its direct BS in subcarrier l, including path loss, shadowing and fast fading.
$G_{c,k}^l$	Channel gain between user k and its coordinated BS in subcarrier l, including path loss, shadowing and fast fading.
$\mathcal{S}_d$	Set of users served on their direct link by a given BS.
$\mathcal{S}_c$	Set of users coordinated by a given BS.
$\mathcal{S}_{\mathrm{RC},d}$	Set of RC users served on their direct link by a given BS.
$\mathcal{S}_{\mathrm{RC},d}$	Set of RC users coordinated by a given BS.
$\mathcal{S}_{\mathrm{BE}}$	Set of BE users served by a given BS.
K	Number of users per BS.
K_d	Number of direct users per BS.
K_c	Number of coordinated users per BS .
$P_{\max,d}$	Maximum downlink transmit power dedicated to direct users per BS.
$P_{\max,c}$	Maximum downlink transmit power dedicated to coordinated users per BS.

Notations specific to Chapters 5, 6 and 7:

P_k^l	Power transmitted to user k by its serving BS in subcarrier l.
$g_{n,k}$	Path loss including shadowing between BS n and user k.
$G_{n,k}^l$	Channel gain between BS n and user k in subcarrier l, including path loss, shadowing and fast fading.
E_k^l	Power control convergence criterion for user k in subcarrier l.
τ_k^l	Estimate of the inter-cell interference level for user k in subcarrier l, used to trigger the power control convergence criterion.
Ω_l	Set of users that are active in subcarrier l.
ζ_k^l	Interference information for user k in subcarrier l.
$n_{\min}$	Minimum number of transmit and receive antennas.
$\mathbf{H}_{n,k}^l$	Channel matrix between BS n and user k in subcarrier l.
ρ_k^l	SNR of user k in subcarrier l.
$\mu_{n,k}^l$	INR corresponding to the interference received by user k from BS n in subcarrier l.
$\mathbf{\Phi}_k^l$	Covariance transmit matrix of user k in subcarrier l.
$\mathbf{V}_k^l$	Transmit matrix for user k in subcarrier l.
$\mathbf{W}_k^l$	Receive matrix for user k in subcarrier l.
$\mathbf{Q}_k^l$	Covariance matrix of the interference plus noise for user k in subcarrier l.
$d_{k,j}^l$	Normalized power per subcarrier l and per stream j for user k.
$\tilde{\gamma}_k^l$	Equivalent SINR of user k in subcarrier l, over all streams.

Chapter 1

Introduction

1.1 Background and motivation

In a few years, cellular networks have passed from circuit-switched networks with dedicated radio resources, to all-Internet Protocol (IP) networks where the Quality of Service (QoS) constraints of all users must be jointly considered. This evolution has placed resource allocation as the central point between the medium access layer and the physical layer. Resource allocation is responsible for managing multi-user radio access, depending on users' QoS demands, and on the channel state. Many work has been done in the previous years to determine opportunistic allocation methods that dynamically update the access resources according to the channel variations. Orthogonal Frequency Division Multiple Access (OFDMA) is the broadband system that has been chosen for the emerging cellular networks, WiMAX [1] and 3GPP LTE [2]. In OFDMA, the available bandwidth is separated into orthogonal subcarriers with flat fading channels. Opportunistic allocation is particularly efficient with OFDMA, as the probability that all users are in a deep fade in all subcarriers decreases when the number of subcarriers increases. Besides, users may be orthogonally multiplexed both in frequency and time, without generating intra-cell interference.

Opportunistic allocation requires resource allocation to be performed on the most accurate Channel State Information (CSI). CSI must then be fed back from each user to the resource allocation controller with the lowest possible delay. Therefore, the distance between the resource allocation controller and the users should be reduced. Resource allocation in cellular networks may be performed by a central controller responsible for all users of several adjacent cells, or may be distributed within each cell. UMTS R99 [3] [4] is based on centralized resource allocation. The benefits of centralized resource allocation concern multi-cell management. Inter-cell interference is decreased at cell's border, thanks to macrodiversity. Its main drawback is the latency induced by transmission of the signaling information from each base station (BS) to the central controller. Besides, UMTS R99 is mostly designed for real-time services. Circuit-switched transmission is possible for non real-time services, but is not efficient due to the waste of resources, and the shared channels cannot serve enough users. In order to bypass these limitations, HSPA has been introduced in UMTS Releases 5 and 6 for the downlink [5] [6]. A shared channel is then designed for all packet-switched transmissions. In HSPA, each BS is responsible for scheduling its users, and performing adaptive modulation and coding depending on the link's channel state. The most recent wireless cellular networks are WiMAX (IEEE 802.16e and 802.16m) and 3GPP LTE. Both systems are based on OFDMA: the physical layer is OFDM, and the multiple access scheme combines Time Division Multiple Access (TDMA) and Frequency Division Multiple Access (FDMA). Resource allocation may be updated every Transmit Time Interval (TTI). CSI should be sent by all users on all subcarriers, leading to a large amount of signaling. The 3GPP LTE standardization proceedings have chosen a totally 'flat' architecture [2] [7]: each BS is the network controller of its served users, and there should be minimum signaling exchange regarding resource allocation between adjacent BSs.

Centralized and distributed resource allocations are represented on Fig. 1.1 and Fig. 1.2.

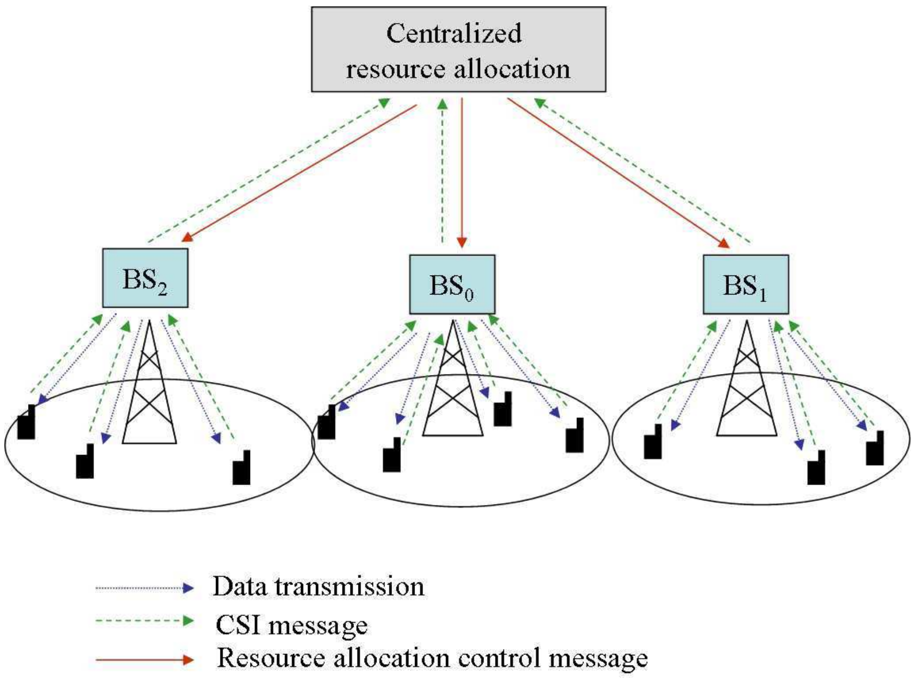


Figure 1.1: Centralized resource allocation

In the centralized architecture, the central controller decides for resource allocation on all the users of the network. It consequently performs a global optimization on the whole network. In the distributed architecture, each BS determines resource allocation for its own users, and is unaware of the resource allocation decisions on the adjacent BSs. Resource allocation therefore leads to a set of local optimizations per BS.

The emerging cellular networks also assume an all-IP packets transmission. This constraint results from the will to allow easy inter-connections of heterogeneous networks. The simplest architecture could then be obtained by connecting an IP server with any final access network, that may be either fixed or cellular. All-IP transmission implies that at radio

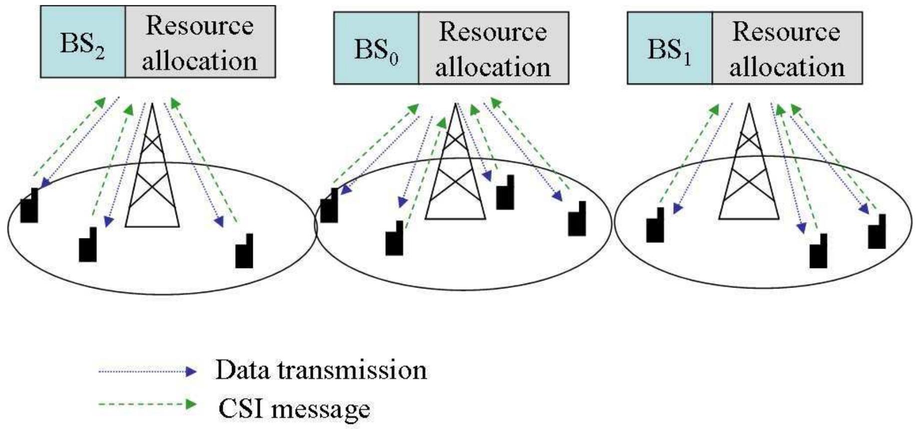


Figure 1.2: Distributed resource allocation

access, all users share the same set of radio resources, and resource allocation is responsible for dynamically allocating resources to the users, so that they reach their required QoS. Up to now, packet-based transmission protocols have mainly been defined for non-real-time services. The issue of fairness among users has been introduced with scheduling objectives such as Proportional Fair [8]. However, these classical scheduling algorithms provide some fairness among users on the channel averaged over several TTIs, but users with real-time services may require a minimum data rate per TTI. New radio access methods for real-time users requesting Voice over IP or streaming services over packet-based transmission should therefore be determined.

Removing the possibility to reserve dedicated resources adds flexibility in resource allocation, but leads to a complex resource allocation task for the resource controller. Indeed, resource allocation must consider all possible competing QoS requirements. OFDMA systems are however well-suited to cope with this problem, as they allow a fine granularity in resource allocation. In OFDMA, per cell resource allocation consists of subcarrier allocation and power allocation. Subcarrier allocation assigns each subcarrier to a specific user, and power allocation distributes the cell sum power over all assigned subcarriers. Resource allocation may correspond to an optimization problem representing users' QoS objectives. It is based on the CSI available at transmitter. It may be performed periodically, with a minimum period equal to one TTI, or occasionally, depending on the data arriving in the buffer, and on the QoS variations. In this thesis, we assume that resource allocation is performed every TTI.

The main drawback of distributed resource allocation is inter-cell interference. In OFDMA, the same bandwidth is shared by all cells. In the absence of coordination of subcarrier or power allocations among adjacent cells, the users located at the border of cells may be highly interfered. Static and dynamic subcarrier and power planning methods have been investigated in the literature to mitigate performance degradations at cell's border [9]. However, static planning is not optimized, and dynamic planning becomes useless at high load. Resource allocation should adequately adapt to the load level: it should allow full use of the bandwidth at low load, when inter-cell interference is not an issue, and remain efficient at high load. A major restriction regarding the viability of distributed networks based on flat architecture as cellular ones is whether they will be able to provide their QoS to all users, whatever their location in the cell is. In all-IP transmission, the users that will suffer the most from inter-cell interference are users with real-time QoS constraints located at the border of cells. However, as radio resources are shared between all users, resource allocation should not solely focus on a specific QoS category, but should establish priority levels between them. Similarly, as all adjacent cells influence each others through inter-cell interference, resource allocation within one cell should not be totally unaware of its effect on the other cells. New methods are required in order to take into account all these requirements in distributed networks.

1.2 Thesis objectives

In this thesis, we consider the distributed architecture of Fig. 1.2. Resource allocation is performed independently per BS. The aim of the thesis is to define distributed resource allocation methods that provide the required QoS to all users, at any location. The proposed methods should entail low inter-BS signaling and reasonable complexity, and allow us to efficiently mitigate inter-cell interference. The context study is OFDMA networks. Resource allocation is then composed of subcarrier allocation and power control. The proposed resource allocation methods are therefore applicable to WiMAX, 3GPP LTE, and future 4G networks based on OFDMA. Parts of this study can also be applied to ad hoc type networks.

The main objectives of the studies performed in this thesis are:

- Evaluate the feasibility of network coordination in distributed networks. Network coordination studies usually assume that all BSs simultaneously have access to the users' data, and that a central controller jointly allocates resources in all involved BSs. We here investigate network coordination with two constraints that are imposed by the distributed architecture: a causality constraint on BS to BS transmission, and a distributed resource allocation constraint.

- Determine a distributed resource allocation method applicable to users with a target data rate requirement, that manages inter-cell interference without any global information. The main issue is the feasibility of data rates' allocation. This has been characterized in previous literature for the single-channel case in Single-Input Single-Output (SISO). Our aim is to obtain a new characterization for the multi-channel case, in SISO as well as in Multiple-Input Multiple-Output (MIMO).

- Investigate distributed resource allocation for Best Effort users, with the consideration of their queue lengths. For these users, an additional degree of freedom is obtained via time-based resource allocation. Thanks to the characterization of whether joint transmission is more efficient than separate transmission to fulfill the resource allocation objective, time-based subcarrier allocation can then be optimized on top of power control.

In all cases, resource allocation corresponds to a global optimization objective that is decomposed over the BSs to address the distributed network constraint.

1.3 Assumptions

Throughout this dissertation, we make some assumptions that are detailed and discussed in this section.

- *Resource allocation per TTI*:
  The optimization objectives and the performance results are assessed in terms of data rates per TTI. We assume that the channel state remains unchanged during one TTI. Iterative power control within one TTI may consequently be performed over the different time slots.

- *Perfect synchronization*:
  All BSs are synchronized, and all mobile terminals receive the data and interference from all BSs with the same time delay. Thus, this delay is not considered within the analysis.

- *Signal constellation*:
  The signal constellation is assumed Gaussian. The modulation level is high enough so that, with these input symbols, the mutual information almost reaches the channel capacity. The influence of the modulation could be introduced in our analysis and numerical results by adding the Signal-to-Noise Ratio (SNR) gap Γ into the capacity formula [10]: $C = \log_2\left(1 + \frac{\mathrm{SNR}}{\Gamma}\right)$, where Γ is a function of the Bit Error Rate and of the modulation.

- *Downlink resource allocation*:
  Resource allocation is performed in the downlink, in order to take into account the per BS sum power constraint. However, the distributed methods from Chapters 5, 6 and 7 can easily be applied to the uplink, by replacing this constraint with per mobile terminal power constraints. On the contrary, the network coordination method studied in Chapters 3 and 4 assume perfect data transmission (in terms of SNR) between the BSs. To extend the proposed methods to the uplink, we should consider imperfect data transmission between two mobile terminals. In the uplink, we no longer deal with network coordination, but with more classical relaying techniques. Many papers concern this subject in the literature (references are given in Section 3.2.1).

- *Full channel state information at the receiver and transmitter*:
  Full CSI is available at the receiver (each user), and also at the transmitter (its serving BS), in all chapters but Chapter 6. In the case of network coordination, the mobile terminal provides full CSI to its direct and coordinated BSs on both links. The obtained results are consequently upper bounds to what could be expected in practical implementations. However, relaxing the full CSI assumption should not lead to a modification of the relative behaviors of the different methods, as shown in Chapter 6 when considering the outage capacity.

1.4 Thesis outline

Resource allocation in distributed networks is investigated from two perspectives: first, when network coordination between BSs is possible, and then, in fully distributed networks without network coordination. The first perspective, studied in Chapters 3 and 4, uses virtual MIMO to increase the link capacity. Resource allocation adapted to this specific problem is studied with an optimization objective applicable to all users in Chapter 3, and with QoS differentiation between Rate Constrained (RC) and Best Effort (BE) users in Chapter 4. The second perspective, which is more general, is investigated in Chapters 5, 6 and 7. Chapters 5 and 6 concern distributed resource allocation for RC users, with SISO and MIMO transmissions respectively. Chapter 7 studies the weighted sum throughput maximization problem in SISO for BE users.

Chapter 2 - Preliminaries

Chapter 2 provides the technical and mathematical background for our problem. The technical background concerns OFDMA, QoS, MIMO, and inter-cell interference. The mathematical background consists of elements on convex optimization, that are used throughout the dissertation to solve resource allocation problems.

Chapter 3 - Fairness increase through distributed network coordination

In Chapter 3, we propose a method suited to distributed networks, that brings additional fairness to the users located at the border of cells. It is based on causal network coordination, where two BSs form a virtual MIMO array to increase the diversity of their users. Causal network coordination can be used in distributed networks, where the inter-BS link is assumed perfect in terms of SNR, but does not allow instantaneous data transmission. First, the chapter proposes a review of the cooperative communications techniques that are relevant to our study. Then, resource allocation for OFDMA multi-cell networks with the proposed causal network coordination is detailed. As the main focus in this chapter is on power control, the same subcarrier allocation method is carried out with all power control objectives. Power control with four objectives leading to varying fairness among users' data rates is tested. These objectives are Globally Optimal, Max-Min Fair, Proportional Fair and Harmonic Mean Fair. Power control is an iterative process, performed over all BSs separately. In each BS, power allocation is further decomposed into power allocation for direct users, and power allocation for coordinated users. Each power allocation sub-problem is locally convex, as inter-cell interference is fixed to the value of the previous iteration. The impact of network coordination depending on the fairness of the power control objective is evaluated. The proposed method is assessed in terms of network throughput and fairness of the data rates' distribution among users. It leads to fairness increases with all power allocation objectives, and to sum throughput increases with all fair objectives.

Chapter 4 - QoS aware resource allocation with distributed network coordination

In this chapter, we extend the causal network coordination procedure to RC and BE users, and propose a method to jointly manage both QoS objectives. One of the conclusions of Chapter 3 is that network coordination is not efficient to maximize the sum throughput, which corresponds to Globally Optimal objective. As this is the QoS objective for BE users when queue lengths are not taken into consideration, network coordination is restricted to the RC users that are located at the border of cells. The chapter first details the state of the art on resource allocation in OFDMA. For RC users, resource allocation can be solved either with a Rate Adaptive objective, which is similar to Max-Min Fair resource allocation, or with a Margin Adaptive (MA) objective. The latter aims at minimizing the sum power required to reach the target data rates of all RC users. We focus on MA optimization for RC users, and Globally Optimal optimization for BE users. In downlink, RC and BE users are linked with a sum power constraint per cell. Subcarrier allocation as well as power control are differentiated among users to optimize resource allocation. They are performed separately for complexity purposes. The proposed method favors RC users over BE users in both subcarrier and power allocations. Network coordination for RC users located at the border of cells decreases the sum power required by these users, thanks to the diversity gain and inter-cell interference decrease. Even when network coordination is not used, users' prioritization increases the ratio of RC users that reach their target data rate, and the sum throughput of BE users. The prioritization scheme becomes even more efficient with network coordination, thanks to the power decrease on RC users.

Chapter 5 - Distributed resource allocation for Rate Constrained users

The MA resource allocation problem for RC users in fully distributed SISO OFDMA networks is investigated in this chapter. Based on the single-channel convergence criterion of distributed power control, we propose a method for determining the target SINR per user and subcarrier, that enables distributed power control to converge in multi-cell OFDMA. This corresponds to a per subcarrier upper bound on the target SINR, denoted as E, that depends on the channel gains with the interfering cells. This criterion is valid when inter-cell interference is not insignificant. Thus, it is adaptively used in a resource allocation method designed for MA optimization. Subcarrier allocation is first performed on each cell independently with an iterative method, that aims at minimizing the sum power required for each user to reach its target data rate, under equal power allocation and average inter-cell interference level assumptions. The subcarriers selected for each user should maximize the proposed E criterion, thus enabling a wider range of variations for the target SINR. Power allocation is then performed by first determining the set of target SINRs per user, and then running power control independently in each subcarrier. The proposed method is very efficient in terms of subcarriers and power consumption, and avoids the power divergence situations that frequently occur at medium to high load with iterative water-filling.

Chapter 6 - Distributed resource allocation in MIMO networks

This chapter extends the results from Chapter 5 to MIMO transmissions. The MA problem is solved in two cases: when full Channel State Information is available at transmitter (full CSIT), and when only statistical CSIT is available at transmitter. In the first case, the distributed E criterion for power control convergence is applied on the data rate per subcarrier. This upper bound is obtained by assuming average inter-cell interference and equal power allocation over all streams. Power allocation is distributed per BS. At each iteration, the precoder matrix that solves the MA problem is determined via water-filling over the subcarriers and streams of all users, and the power control convergence constraint is included within water-filling. In the second case, an approximate analytical expression is obtained for the outage capacity, as a function of the SNR and of the outage probability. This approximate expression is close to the results obtained via Monte-Carlo snapshots at practical SNR and outage capacity values, for various antennas configurations. It expresses the outage capacity for a fixed outage probability as a concave function of the SNR, that can be used to solve the MA problem as a convex optimization problem. An iterative power allocation method, including a distributed E convergence condition for power control, is then deduced. In the full CSIT as well as in the statistical CSIT cases, the proposed methods with criterion E avoid power control divergence situations at medium to high load, and thus lead to more RC users reaching their QoS objective.

Chapter 7 - Multi-cell weighted sum throughput maximization

Finally in Chapter 7, we consider SISO multi-cell OFDMA networks with BE users. The objective is the maximization of the weighted sum throughput (WSTM). Chapter 7 first reviews the relevant literature regarding resource allocation for BE users in OFDMA. Then it describes a graph-based subcarrier allocation method that allows joint transmission of two interfering links depending on the convexity of their capacity region, and on their relative links' weights. Joint transmission is authorized if it leads to the maximization of the weighted sum throughput. This criterion extends to the multi-cell case, and is used to build an interference graph for the network, that indicates which users should simultaneously transmit in the same subcarrier. We then investigate power control for WSTM. A distributed power control in two steps is proposed. It first determines which users and subcarriers should be transmitting, and then operates in high SINR regime for these users. Contrary to previous work that assumed high SINR on all links, this method can be used in all SINR regimes. Subcarrier allocation and power control, used jointly or separately, are assessed via dynamic numerical simulations, where the weight of each user is proportional to its queue length. At medium to high load, they lead to higher weighted sum throughput and fairer queue length management than binary power allocation. Besides, graph-based subcarrier allocation followed by power control importantly decreases the power and subcarriers' consumption.

1.5 List of publications

Journal papers

1. "Power control in distributed cooperative OFDMA cellular networks", M. Pischella and J-C. Belfiore, IEEE Transactions on Wireless Communications, vol 7, issue 5, part 2, pp.1900-1906, May 2008.

2. "Distributed resource allocation for rate-constrained users in multi-cell OFDMA networks", M. Pischella and J-C. Belfiore, IEEE Communication Letters, vol 12, issue 4, pp. 250-252, April 2008.

3. "Resource allocation for QoS-aware OFDMA systems using distributed network coordination", M. Pischella and J-C. Belfiore, accepted for future publication in IEEE Transactions on Vehicular Technology.

4. "Distributed margin adaptive resource allocation in MIMO OFDMA networks", M. Pischella and J-C. Belfiore, submitted to IEEE Transactions on Communications.

5. "Weighted sum throughput maximization in multi-cell OFDMA networks", M. Pischella and J-C. Belfiore, submitted to IEEE Transactions on Vehicular Technology.

Conference papers

1. "Distributed resource allocation in MIMO OFDMA networks with statistical CSIT", M. Pischella and J-C. Belfiore, SPAWC 2009, Perugia, Italy, June 2009.

2. "Distributed weighted sum throughput maximization in multi-cell wireless networks", M. Pischella and J-C. Belfiore, PIMRC 2008, Cannes, France, Sept. 2008.

3. "QoS-based resource allocation with cooperative diversity in OFDMA", M. Pischella and J-C. Belfiore, VTC Spring 2008, Singapore, May 2008.

4. "Achieving a Frequency Reuse Factor of 1 in OFDMA cellular networks with cooperative communications", M. Pischella and J-C. Belfiore, VTC Spring 2008, Singapore, May 2008.

5. "A Radio Resource Management strategy for downlink cooperation in distributed networks", M. Pischella and J-C. Belfiore, PIMRC 2007, Athens, Greece, Sept. 2007.

6. "Optimal power allocation for downlink cooperative cellular networks", M. Pischella and J-C. Belfiore, VTC Spring 2007, Dublin, Ireland, April 2007.

7. "Graph-based weighted sum throughput maximization in OFDMA cellular networks", M. Pischella and J-C. Belfiore, submitted to IWLCD 2009.

Chapter 2

Preliminaries

2.1 Technical background

2.1.1 Orthogonal Frequency Division Multiple Access

We consider OFDMA downlink cellular networks. OFDM is the chosen physical layer for many standards, such as IEEE 802.11a, 802.11g, 802.20, ADSL, and for both major emerging cellular networks' standards, WiMAX and 3GPP LTE.

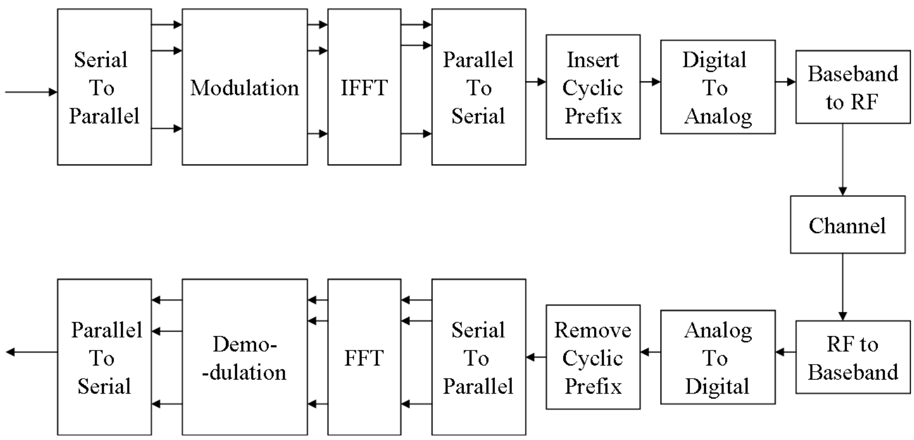


Figure 2.1: Implementation of an OFDM system

OFDM is a multi-carrier transmission technique that divides the available bandwidth into several orthogonal subcarriers with equal bandwidth. The information symbols are then transmitted in parallel over the wireless channel. Orthogonal transmission is obtained by using an Inverse Fast Fourier Transform (IFFT) at transmission, and a Fast Fourier Transform (FFT) at reception (Fig. 2.1). The data to be transmitted are first divided into N_{FFT} parallel groups that are independently modulated. Let $\mathbf{x}^l$ be the complex subsymbols at the output of the l^{th} modulator, that correspond to the l^{th} subcarrier. The IFFT module transforms all subsymbols into N_{FFT} time samples $\mathbf{X}^n$ with:

$$\mathbf{X}^n = \frac{1}{\sqrt{N_{\text{FFT}}}} \sum_{l=0}^{N_{\text{FFT}}-1} \mathbf{x}^l e^{j2\Pi \frac{nl}{N_{\text{FFT}}}} \tag{2.1}$$

A cyclic prefix of N_{CP} time samples is appended to the N_{FFT} time samples (Fig. 2.2). It copies the N_{CP} last samples of each OFDM symbol, and is inserted at the beginning of the symbol. N_{CP} must be greater than the maximum multipath delay spread. If this condition is fulfilled, then the multipath delay of one symbol will only affect the next symbol's cyclic prefix, leaving the useful data samples free from Inter-Symbol-Interference (ISI).
At the receiver, the cyclic prefix is removed, and $\mathbf{x}^l$ is retrieved by FFT. The received signal before the FFT is the result of a circular convolution between $\mathbf{X}^n$ and the channel response $\mathbf{H}$. Thus the result of the FFT is :

$$\hat{\mathbf{x}}^l = \mathbf{h}^l \mathbf{x}^l + \mathbf{n}^l \tag{2.2}$$

where $\mathbf{n}^l$ is the additive channel noise for subcarrier l. $\mathbf{h}^l$ is the complex channel transfer factor for subcarrier l.

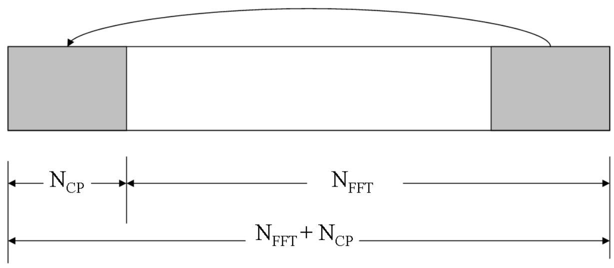


Figure 2.2: Cyclic prefix of an OFDM system

OFDMA is a multiple access technique based on OFDM physical layer. The multiple access scheme combines TDMA and FDMA. The minimum resource level that may be allocated to each user is one subcarrier within one time slot. This basic resource is orthogonally allocated to users in each cell, so that there is no intra-cell interference. In multi-cellular networks, however, adjacent cells may be transmitting on the same subcarriers. Each subcarrier can then be viewed as an interference channel. If the power level per subcarrier is an optimization variable, the subcarriers are linked by a sum power constraint, which applies to all the subcarriers for downlink transmission, or to the subset of subcarriers assigned to the same user, for uplink transmission. Due to this sum power constraint, resource allocation on the different interference channels is not parallel. Resource allocation in multi-cell OFDMA must consequently take into account both multi-user QoS constraints per cell (namely, how to assign subcarriers and distribute power to the different users of a cell, so as to fulfill their required QoS) and multi-cell interference per subcarrier.
In the following, we consider complex OFDMA signals where each narrowband subcarrier is assumed to experiment a flat fading quasi-static channel.

2.1.2 Quality of Service

In wireless networks, each user requests one or several uplink and downlink services. These services are characterized by specific QoS constraints. The network layer may map the requested services to predefined sets of QoS constraints that apply to the end-to-end wireless transmission. These constraints may correspond to a minimum data rate, a tolerated jitter, a maximum transmission delay, ...

We investigate resource allocation per TTI. Consequently, only per TTI QoS parameters are relevant to our study, and the QoS constraints then correspond to two categories of users:

1. Rate Constrained (RC) users: these users must achieve a target data rate to fulfill their QoS constraint. We assume that the target data rate should be reached within one TTI. From a network viewpoint, the objective is to minimize the sum power required to reach the target data rate of RC users.

2. Best Effort (BE) users: these users are not constrained by any objective regarding their individual per TTI quality indicators. From a network viewpoint, the objective is to maximize the sum data rate of all BE users. If the time variations are considered, then the queue length of BE users should also be taken into account. Indeed, buffers have a limited maximum size, and buffer overflow leads to data loss.

In Chapter 3, we assume that all users have an infinite backlog of data to send. This assumption also stands for BE users in Chapter 4. In Chapters 4, 5 and 6, the RC users have exactly the data corresponding to their required data rate to send in each TTI. In Chapter 7, the actual data traffic is modeled for each BE user.

2.1.3 Multiple-Input Multiple-Output channel

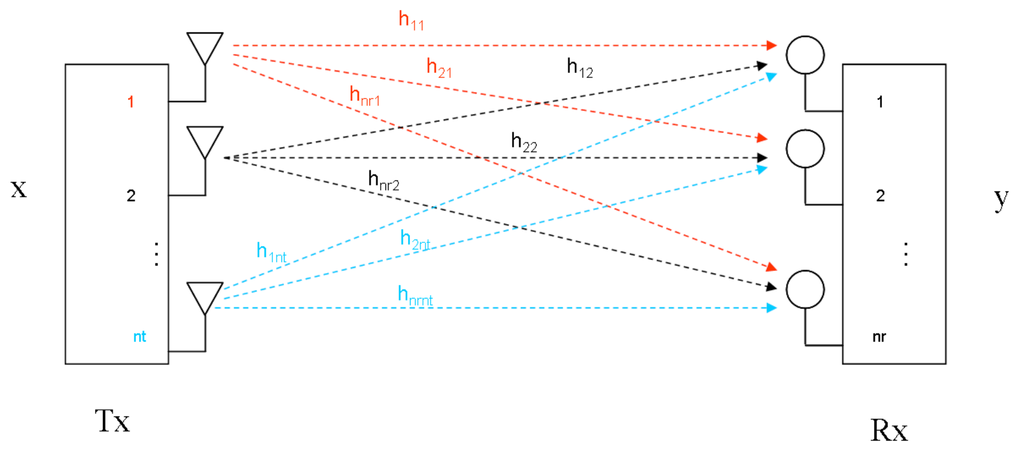


Figure 2.3: MIMO channel model

We consider a point-to-point MIMO channel with n_t transmit antennas and n_r receive antennas, subject to Additive White Gaussian Noise (AWGN) but not to interference. The channel model is [11]:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{2.3}$$

where $\mathbf{x} \in \mathbb{C}^{n_t \times 1}$ is the transmitted vector, $\mathbf{y} \in \mathbb{C}^{n_r \times 1}$ is the received vector, $\mathbf{n} \in \mathbb{C}^{n_r \times 1}$ is an AWGN noise vector with variance N_0, and $\mathbf{H} \in \mathbb{C}^{n_r \times n_t}$ is the channel matrix.

$$\mathbf{H} = \begin{pmatrix} h_{11} & h_{12} & ... & h_{1n_t} \\ h_{21} & h_{22} & ... & h_{2n_t} \\ & & ... & \\ h_{n_r1} & h_{n_r2} & ... & h_{n_rn_t} \end{pmatrix}$$

h_{ij} is the channel gain between the j-th transmit antenna and the i-th receive antenna. It consists of path loss, shadowing and fast fading. If the channel is zero-mean i.i.d., then the fading element of each channel gain is composed of independent zero-mean complex Gaussians, each with independent real and imaginary parts with variance $\frac{1}{2}$.
The mutual information for a given channel realization $\mathbf{H}$ is [12] :

$$\boldsymbol{I}(\mathbf{x};\mathbf{y}|\mathbf{H}) = \log_2\left(\det\left(\mathbf{I}_{n_r} + \frac{1}{N_0}\mathbf{HQH}^H\right)\right) \tag{2.4}$$

where $\mathbf{Q} = \mathbb{E}\{\mathbf{xx}^H\}$ is the covariance matrix of the transmit vector. The mutual information is maximized when $\mathbf{x}$ is complex Gaussian. The capacity is then defined as the maximum of the mutual information between $\mathbf{x}$ and $\mathbf{y}$ given a constraint on the total power transmission, $\mathrm{Tr}(\mathbf{Q}) \leq P$. In the following of the dissertation, with an abuse of notation, we denote $\boldsymbol{I}(\mathbf{x};\mathbf{y}|\mathbf{H})$ as the capacity of the channel, C, even in the absence of optimization on the input symbols. C is expressed in bits per channel use.
If CSI is known at the transmitter, then the capacity is maximized by using water-filling over the singular values of the channel as a power allocation policy [13]. If no CSI is available at transmitter, a good assumption regarding capacity maximization is to distribute the power equally among the transmit antennas.
MIMO systems can be used to multiplex several streams over the same point-to-point transmission. These streams may either belong to the same user (therefore increasing its data rate), or to several users (therefore leading to multi-user transmission). The number of separable streams is equal to the rank of the channel matrix $\mathbf{H}$. MIMO can be used to increase the total data rate through data multiplexing, and to increase the robustness of communications by sending the same data over multiple paths subject to different fading conditions. MIMO thus provides additional degrees of freedom, that may be expressed in terms of diversity and multiplexing gains, compared to SISO transmission.
Let us now assume that a linear precoder is used at transmission. $\mathbf{s} \in \mathbb{C}^{M \times 1}$ is a complex Gaussian vector containing the M input symbols, and $\mathbf{V} \in \mathbb{C}^{n_t \times M}$ is the linear precoder. The transmitted vector is then $\mathbf{x} = \mathbf{Vs}$. The receiver aims at retrieving the input symbols, during the detection step. The most efficient detection technique is Maximum Likelihood (ML). However, its complexity is often prohibitive. Zero-Forcing (ZF) and Minimum Mean Square Error (MMSE) equalizers are simple linear processing methods. They are also used as basis for non-linear Decision Feedback (DFE) treatments (DFE-ZF and DFE-MMSE). The linear receiver aims at obtaining $\mathbf{s}$ from the received vector $\mathbf{y}$ through a linear operation

$$\hat{\mathbf{s}} = \mathbf{W}^H\mathbf{y} \tag{2.5}$$

where $\mathbf{W}^H \in \mathbb{C}^{M \times n_r}$.
Let $\tilde{\mathbf{H}} = \mathbf{HV}$. The ZF receiver inverts the channel by applying the pseudo-inverse matrix

of $\tilde{\mathbf{H}}$:

$$\mathbf{W} = \left(\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\right)^{-1}\tilde{\mathbf{H}}^H \tag{2.6}$$

The main problem of ZF receiver is noise amplification. The MMSE receiver minimizes the mean square error $MSE = \mathbb{E}\{\|\hat{\mathbf{s}} - \mathbf{s}\|^2\}$, and is defined as:

$$\mathbf{W} = \left(\tilde{\mathbf{H}}^H\tilde{\mathbf{H}} + \frac{N_0}{P}\mathbf{I}_M\right)^{-1}\tilde{\mathbf{H}}^H \tag{2.7}$$

where $P = \mathrm{Tr}(\mathbf{s}\mathbf{s}^H)$. The MMSE receiver is equivalent to the ZF receiver when the noise becomes negligible.
For the most general slow fading channel case, if we assume that each channel realization is known at transmitter, then optimization over the instantaneous capacity corresponding to a given channel realization (2.4) can be performed. If the time average converges to the same limit for almost all channels realizations of the fading process, the channel is ergodic. In that case, optimization methods should apply to the ergodic capacity. If the channel realization is not known at transmitter, then the optimization concerns the outage probability and the corresponding outage capacity.

2.1.4 Interference management in multi-cell networks

Interference channel

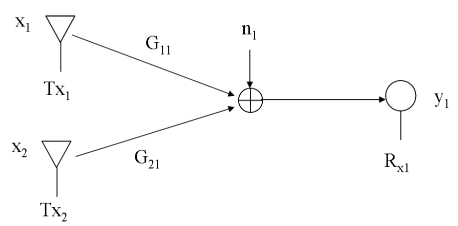


Figure 2.4: Multi-access channel

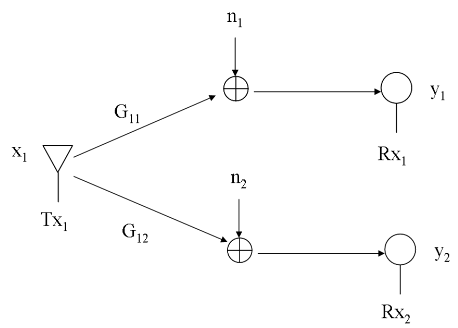


Figure 2.5: Broadcast channel

Point-to-point transmission is the simplest case where one transmitter communicates with one receiver. In SISO transmission with one transmit antenna and one receive antenna,

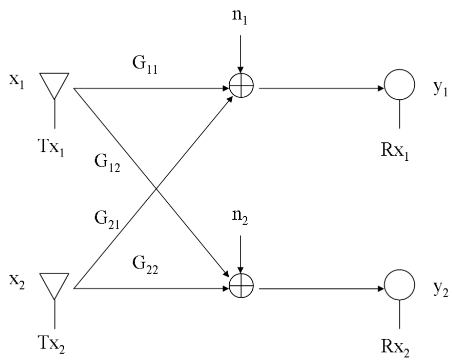


Figure 2.6: Interference channel

the capacity of a point-to-point link with AWGN noise is

$$C = \log_2(1 + \mathrm{SNR}) \tag{2.8}$$

where SNR is the Signal-to-Noise Ratio. We hereunder review the different SISO multi-user channels.

The multi-access channel (MAC) corresponds to several transmitters and only one receiver (Fig. 2.4). The capacity region of the MAC is the set of individual data rates that can be jointly achieved for all transmitters. The corner points of the capacity region are reached by using Successive Interference Cancellation (SIC) [14]. The decoding order determines the final data rates. Whatever the cancellation order, the same sum optimal data rate is achieved with SIC. The broadcast channel corresponds to one transmitter and several receivers (Fig. 2.5). The sum data rate is maximized by superposition coding at the transmitter, and SIC at each of the receivers. Each receiver must decode the weaker users before decoding its own data, so as to achieve the maximum sum data rate. Information-theoretic problems are often more easily solved on the MAC than on the broadcast channel. The uplink-downlink duality enables us to study the broadcast channel via the MAC, and then reverse the roles of transmitters and receivers [15].

The interference channel (Fig. 2.6) is the most complex case, as it involves several unrelated transmitters and receivers that are interfering each other. No joint treatment such as SIC is possible in transmission or reception, as neither the set of transmitters nor the set of receivers are communicating. The capacity region of the interference channel has not yet been fully characterized. It has been determined in the strong interference case [16], for a class of deterministic interference channels [17], and for a class of degraded interference channels [18]. For the general case, Han and Kobayashi [19] have derived an achievable rate region, whose characterization is quite complex. It involves splitting the transmitted information into two parts: a private information to be decoded only by a single receiver, and common information to be decoded by all receivers. This method is applied by Etkin et al. in [20] for a two-users interference channel with AWGN noise. By using an outer bound on the capacity region, it is shown that the achieved rate region is within one bit of the capacity region.

Recently, the notion of degrees of freedom on the symmetric interference channel has been introduced in [20] for a two users Gaussian interference channel, and extended to the K users Gaussian symmetric interference channel, where all users are subject to the same SNR and the same Interference to Noise Ratio (INR), in [21]. At high SNR, the capacity of a point-to-point AWGN link is $C_{\text{awgn}} \approx \log(\text{SNR})$. Let $\alpha = \frac{\log \text{INR}}{\log \text{SNR}}$ be the ratio of the INR and SNR in dB. Let C_{sym} be the symmetric capacity, i.e. the best rate that all users can simultaneously achieve. Then the generalized degree of freedom is defined as:

$$d_{\text{sym}}(\alpha) = \lim_{(\text{SNR},\text{INR}) \to \infty} \frac{C_{\text{sym}}(\text{INR}, \text{SNR})}{C_{\text{awgn}}(\text{SNR})} \tag{2.9}$$

Jafar and Vishwanath have shown in [21] that these values can actually be reached. The main conclusion is that when $\alpha \geq 2$, $d_{\text{sym}} = 1$. This means that at very high interference, the channel can be seen as interference-free.

Complex treatments in transmission and reception, involving the use of private and common information, are required to obtain these results. We do not investigate these signal processing issues in the thesis. Our aim is to mitigate interference thanks to resource allocation. Therefore in the following, we consider interference as noise when working on the interference channel. Consequently, we always express the SISO capacity of a given user as:

$$C = \log_2(1 + \text{SINR}) \tag{2.10}$$

where SINR is the Signal-to-Noise-plus-Interference Ratio of this user.

Interference in OFDMA cellular networks

There is no intra-cell interference in OFDMA, due to the orthogonality between subcarriers, and to FDMA multiple access. However, several cells in OFDMA multi-cellular networks may be using the same subcarriers for unrelated transmissions. In downlink, each user may receive inter-cell interference coming from adjacent cells transmitting in the same subcarriers. The users located at the border of cells are the most affected by inter-cell interference, as their propagation loss to the interfering BSs is of the same order as their propagation loss to their serving BS.

In order to circumvent this limitation, several methods have been proposed, some of them in 3GPP LTE study items [22] [23]. A review of the relevant literature is provided in [9]. The main techniques are:

1. *Subcarrier planning*: Use subcarrier allocation so as to avoid joint transmission of interfering users in the same subcarrier. This may be performed through static planning with fixed reuse patterns, or with Differential Frequency Partitioning (DFP). In the first case, groups of non-interfering adjacent cells are formed, and each group is assigned an independent subset of the total available set of subcarriers. In the second case, the reuse pattern only concerns the users that are located at the border of each cell, whereas all subcarriers can be allocated in all cells for the users located inside each cell. An example of DFP with Frequency Reuse Factor (FRF) equal to 3 is depicted on Fig. 2.7. The main limitation of static planning is that it reduces the bandwidth available for transmission. Dynamic subcarrier allocation may also be considered: each cell may choose the subcarriers with lowest inter-cell interference level for its transmission. However, it is only useful at low load, and it may not converge to a stable state when distributed per BS.

2. *Power planning*: Use power allocation in order to mitigate inter-cell interference. Power planning determines adjacent subsets of subcarriers. Each group of non- interfering cells then transmits at full power within its assigned subcarriers' subset, and has a bounded transmit power level in the other subcarriers' subsets. The subsets may be static, or they may be dynamically updated, depending on the load level, allowing less concurrent full power transmission among cells when the load increases. Similarly to the subcarrier planning case, Differential Power Partitioning (DPP) may be defined. In that case, adjacent subcarriers' subsets are only used at cell's border.

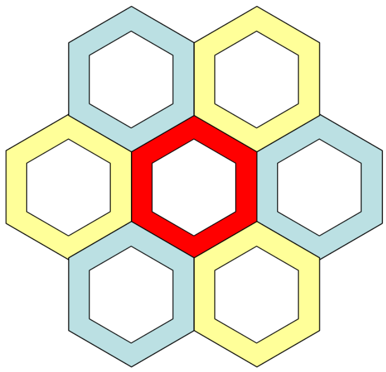


Figure 2.7: Example of DFP with FRF equal to 3

IEEE 802.16e standard [1] defines two modes for subchannel allocation that may be relevant to inter-cell interference mitigation: Partially Used SubChannelization (PUSC), and Fully Used SubChannelization (FUSC). Subcarriers are grouped into subchannels using pseudo-random permutations. In downlink, the minimum set of allocated resources per user is one subchannel. In PUSC, for sectored antennas, each sector is assigned adjacent sets of subchannels, whereas in FUSC, the same subchannels are available on the whole cell. At low to medium load, the probability of using the same subcarrier on two adjacent cells is lowered thanks to the permutations. Inter-cell interference is however not handled at high load, when almost all subcarriers are allocated.

2.2 Elements of convex optimization

Throughout the dissertation, we consider resource allocation problems with a network-wide objective. However, we are bound by several constraints to solve these problems:

- Network architecture constraint: our study is limited to distributed networks, where downlink resource allocation should be performed per BS. The network-wide optimization problem must consequently be split into N_{BS} sub-problems, one per BS.

- Inter-cell interference and QoS constraints: in distributed multi-cell networks, inter-cell interference compels us to use iterative resource allocation over the BSs. Because of users' QoS constraints such as target data rate constraints for RC users, or queue length constraints for BE users, all the subcarriers allocated to a given user must be jointly considered. Inter-cell interference management can consequently not be independently performed in each subcarrier.

- Complexity constraint: joint subcarrier and power allocation may be limited by complexity (details are given in the state of the art on resource allocation in OFDMA, in Section 4.2). Besides, in most cases, iterative methods are used, thus the complexity issue is even more problematic as the resource allocation procedures is repeatedly performed. Consequently, sub-optimal heuristics are preferred, at the expense of some performance loss.

Most of the problems dealt with in the thesis are or can be decomposed into convex optimization problems. The main advantage of convex optimization problems is that they have a unique global optimum. Besides, thanks to duality, under certain conditions, the primal problem can be solved by solving the dual problem. The dual problem is often separable into several sub-problems. This decomposition feature is particularly relevant to our problem. The first level of decomposition considered is per BS decomposition. This is of course required in distributed networks. Then a per user or per subcarrier decomposition may be used. In all the cases, we must keep in mind that if the initial resource allocation problem is not convex, then even if its decomposed sub-problems are convex, the final result will not necessarily be the global optimum of the original problem.
In this section, we detail the properties of convex optimization problems, as well as the main methods that will be used in the thesis to solve them. Our main references are Boyd and Vanderbergue [24] for convex optimization, and Chiang [25] for geometric programming. We first start with the definitions and main types of convex problems. Then we investigate the duality theory, and explain how to analytical solve simple convex problems with the Karush-Kuhn-Tucker (KKT) conditions. An example of decomposition through Lagrangian decomposition is also provided. In Section 2.2.3, Newton's method to numerically solve convex problems is detailed. Finally in Section 2.2.4, we provide some insight into the class of geometric programming.

2.2.1 Convex optimization problems

General mono-objective optimization problems

A mono-objective optimization problem has the form:

$$\begin{aligned} &\min_{\mathbf{x}} \ f_0(\mathbf{x}) \\ &\text{s.t. } f_i(\mathbf{x}) \leq 0, \forall\ i \in \{1, ..., m\} \\ &\text{s.t. } h_i(\mathbf{x}) = 0, \forall\ i \in \{1, ..., p\} \end{aligned} \tag{2.11}$$

Where 'min' stands for 'minimize'. The vector $\mathbf{x} = [x_1, ..., x_n]'$ is the optimization variable, $f_0 : \mathbb{R}^n \to \mathbb{R}$ is the objective function, $f_i : \mathbb{R}^n \to \mathbb{R}$, $\forall i \in \{1, ..., m\}$ are the inequality constraint functions, and $h_i : \mathbb{R}^n \to \mathbb{R}$, $\forall i \in \{1, ..., p\}$ are the equality constraint functions. A vector $\mathbf{x}^*$ is optimal if it has the smallest objective value, among all the vectors that satisfy the constraints.

Convex sets

A set $S \subset \mathbb{R}^n$ is convex if the line segment between any two points in S lies in S: for any $(\mathbf{x}_1, \mathbf{x}_2) \in C^2$ and any θ with $0 \leq \theta \leq 1$:

$$\theta \mathbf{x}_1 + (1-\theta)\mathbf{x}_2 \in S$$

Convex functions

A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if for all $(\mathbf{x}, \mathbf{y}) \in (\mathbb{R}^n)^2$ and all $(\alpha, \beta) \in \mathbb{R}^2$ with $\alpha + \beta = 1$, $\alpha \geq 0$, $\beta \geq 0$,

$$f(\alpha \mathbf{x} + \beta \mathbf{y}) \leq \alpha f(\mathbf{x}) + \beta f(\mathbf{y})$$

A function f is convex over a given set $\mathrm{dom} f$ if and only if $\mathrm{dom} f$ is convex and the Hessian matrix of f is semidefinite positive for all $\mathbf{x} \in \mathrm{dom} f$. If f is a function on $\mathbb{R}$, this reduces to $f^{''}(x) \geq 0$, and $\mathrm{dom} f$ is an interval. The opposite of a convex function is a concave function. The convexity of some important functions that will be used throughout the dissertation is discussed hereunder:

- Linear functions are convex and concave on $\mathbb{R}$.
- Exponential: $\exp(ax)$ is convex on $\mathbb{R}$, for any $a \in \mathbb{R}$.
- Powers: x^a is convex on $\mathbb{R}_+$ when $a \geq 1$ or $a \leq 0$, and concave for $0 < a < 1$.
- Logarithm: $\log(x)$ is concave on $\mathbb{R}_+$.
- Max function: $f(\mathbf{x}) = \max\{x_1, ..., x_n\}$ is convex on $\mathbb{R}^n$.
- Log-sum-exp: The function $f(\mathbf{x}) = \log(\exp x_1 + ... + \exp x_n)$ is convex on $\mathbb{R}^n$.
- Geometric mean: $f(\mathbf{x}) = \left(\prod_{i=1}^n x_i\right)^{1/n}$ is concave on $\mathbb{R}^n_{++}$.
- Log-determinant: $f(\mathbf{X}) = \log(\det(\mathbf{X}))$ is concave on the set of symmetric positive definite matrices.

Convex minimization problems

A convex optimization problem is a mono-objective optimization problem of the form (2.11), where the objective and inequality constraint functions are convex, and the equality constraint functions are affine.
A convex optimization problem is expressed in standard form as:

$$\begin{aligned}
&\min\nolimits_{\mathbf{x}}\ f_0(\mathbf{x}) \\
&\text{s.t. } f_i(\mathbf{x}) \leq 0, \forall\ i \in \{1, ..., m\} \\
&\text{s.t. } \mathbf{a}_i^T \mathbf{x} = \mathbf{b}_i, \forall\ i \in \{1, ..., p\}
\end{aligned} \tag{2.12}$$

If $m = 0$ and $p = 0$, the minimization problem is unconstrained. If $m = 0$ but $p \geq 1$, it is an equality constrained minimization problem. If $m \geq 1$ and $p \geq 1$, it is an inequality constrained minimization problem. If $p \geq 1$ then the equality constraints can be equivalently denoted as $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{A} \in \mathbb{R}^{p \times n}$ with rank $\mathbf{A} = p < n$.

Concave maximization problems

$$\begin{aligned} &\max_{\mathbf{x}}\ f_0(\mathbf{x}) \\ &\text{s.t. } f_i(\mathbf{x}) \leq 0, \forall\ i \in \{1, ..., m\} \\ &\text{s.t. } \mathbf{a}_i^T \mathbf{x} = \mathbf{b}_i, \forall\ i \in \{1, ..., p\} \end{aligned} \tag{2.13}$$

If the objective function f_0 is concave and inequality functions are convex, then problem (2.13) is equivalent to a convex optimization problem, where the objective is to minimize the convex objective function $-f_0$.

Main property of convex optimization problems

A convex optimization problem has only one global optimum. Any locally optimal point is therefore also globally optimal.
For an unconstrained problem ($m = 0$ and $p = 0$), the global optimum is obtained by nulling the gradient of the objective function:

$$\nabla f_0(\mathbf{x}) = \mathbf{0}$$

For equality constrained optimization problems, efficient solving methods are provided in the two following sections. It should be noted that we do not consider methods for inequality constrained optimization problems in this chapter. Indeed, all the problems treated throughout the dissertation are equality constrained optimization problems, where the equality constraint either corresponds to a sum power constraint per cell, or to a sum data rate constraint per user. Interior points methods are detailed in [24].

2.2.2 Duality theory and applications for convex optimization problems

Duality

In this section, we consider general optimization problems of the form (2.11), without any assumption on the convexity. Let D be the non-empty domain of (2.11). The initial problem is called the primal problem.
The Lagrangian $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$ associated with problem (2.11) is:

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{i=1}^{p} \nu_i h_i(\mathbf{x}) \tag{2.14}$$

where vectors $\boldsymbol{\lambda}$ and $\boldsymbol{\nu}$ are the dual variables or Lagrange multiplier vectors associated with problem (2.11). λ_i is the Lagrange multiplier associated with the i^{th} inequality constraint $f_i(\mathbf{x}) \leq 0$, and ν_i is the Lagrange multiplier associated with the i^{th} equality constraint $h_i(\mathbf{x}) = 0$.
The dual function $g : \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$ is defined as the minimum value of the Lagrangian over $\mathbf{x} \in D$:

$$g(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \min_D\ L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \min_D\ \left(f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{i=1}^{p} \nu_i h_i(\mathbf{x}) \right) \tag{2.15}$$

The dual function is the minimum of a family of affine functions of $(\boldsymbol{\lambda}, \boldsymbol{\nu})$. Therefore, it is always concave, even when the primal problem (2.11) is not convex.
If we suppose that all $\lambda_i \geq 0$, then for any $\tilde{\mathbf{x}}$,

$$g(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \min_D \, L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) \leq L(\tilde{\mathbf{x}}, \boldsymbol{\lambda}, \boldsymbol{\nu}) \leq f_0(\tilde{\mathbf{x}})$$

Consequently, $g(\boldsymbol{\lambda}, \boldsymbol{\nu}) \leq \mathbf{p}^*$, where $\mathbf{p}^*$ is the optimal value of the primal problem.
The dual function provides a lower bound on the optimal value of the primal problem for each pair $(\boldsymbol{\lambda}, \boldsymbol{\nu})$, with $\lambda_i \geq 0, \forall i \in \{1, ..., m\}$. The optimal value of the primal problem can consequently be approximated by solving the Lagrange dual problem :

$$\begin{aligned} &\max_{\{\boldsymbol{\lambda}, \boldsymbol{\nu}\}} \, g(\boldsymbol{\lambda}, \boldsymbol{\nu}) \\ &\text{s.t. } \lambda_i \geq 0, \forall \, i \in \{1, ..., m\} \end{aligned} \tag{2.16}$$

The solutions of problem (2.16), $(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$, are called dual optimal or optimal Lagrange multipliers. As the Lagrange dual problem (2.16) is a convex optimization problem, $(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ are unique global optimal solutions.
Let $\mathbf{d}^* = \mathbf{g}(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ be the optimal value of the Lagrange dual problem. The inequality $\mathbf{d}^* \leq \mathbf{p}^*$ holds even if the primal problem is not convex. $\mathbf{p}^* - \mathbf{d}^*$ is referred to as the optimal duality gap of the primal problem. If the duality gap is equal to zero, then the optimal solution of the primal problem is obtained by solving the dual problem. The duality gap is equal to zero if both following conditions are verified:

1. The primal problem is convex.
2. Slater's condition holds: $\exists \, \mathbf{x}$ such that $f_i(\mathbf{x}) < 0, \forall \, i \in \{1, ..., m\}$, and $\mathbf{A}\mathbf{x} = \mathbf{b}$.

It should be noted that for equality-constrained convex problems, Slater's condition is fulfilled as soon as the feasible set is not empty (i.e., $\exists \, \mathbf{x}$ such that $\mathbf{A}\mathbf{x} = \mathbf{b}$.). Consequently, strong duality holds for any feasible convex problem.

Karush-Kuhn-Tucker conditions

The Karush-Kuhn-Tucker (KKT) conditions are:

$$\begin{aligned} f_i(\mathbf{x}^*) &\leq 0, \forall \, i \in \{1, ..., m\} \\ h_i(\mathbf{x}^*) &= 0, \forall \, i \in \{1, ..., p\} \\ \lambda_i^* &\geq 0, \forall \, i \in \{1, ..., m\} \\ \lambda_i^* f_i(\mathbf{x}^*) &= 0, \forall \, i \in \{1, ..., m\} \\ \nabla f_0(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i^* f_i(\mathbf{x}^*) + \sum_{i=1}^{p} \nu_i^* h_i(\mathbf{x}^*) &= 0 \end{aligned} \tag{2.17}$$

For any optimization problem with differentiable objective and constraints functions for which strong duality holds, any pair of primal and dual optimal points $(\mathbf{x}^*, (\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*))$ must verify the KKT conditions.
Besides, if the primal problem is convex, then the KKT conditions are also sufficient for the points to be primal and dual optimal. Consequently, for a convex optimization problem with differentiable objective and constraints functions satisfying Slater's condition, the KKT

conditions provide necessary and sufficient conditions for optimality.
The KKT conditions are useful in convex optimization to obtain analytical solutions to simple problems. We will use them throughout the dissertation, for simple problems stemming from water-filling, or more complex problems as the one studied for network coordination of RC users in Chapter 4.

Applications of duality theory to decomposition

Convex problems can be decomposed into several sub-problems in the dual space. Indeed, the dual space often allows relaxing some constraints, and consequently working on parallel optimization problems. A detailed review of the decomposition methods and of their applications to cross-layer optimization is performed in [26].
Dual decomposition can be used when the problem has a coupling variable such that, when relaxed, the optimization problem decouples into several sub-problems. This method is also called 'Lagrangian relaxation'. The decomposition takes place in the dual space and the coupling variable is the set of Lagrange multipliers. Each sub-problem is solved in parallel, for a fixed value of the dual variables. Then the dual variables are updated with the aim to minimize the dual function.
For instance, let us consider an inequality constrained convex optimization problem of the following form:

$$\begin{aligned} &\min_{\{\mathbf{x}_1,..,\mathbf{x}_m\}} \sum_{i=1}^{m} f_i(\mathbf{x}_i) \\ &\text{s.t. } \sum_{i=1}^{m} h_i(\mathbf{x}_i) \leq \mathbf{c} \end{aligned} \tag{2.18}$$

The Lagrangian is:

$$L(\mathbf{x}_i, \boldsymbol{\lambda}) = \sum_{i=1}^{m} f_i(\mathbf{x}_i) + \boldsymbol{\lambda}^T \left(\sum_{i=1}^{m} h_i(\mathbf{x}_i) - \mathbf{c} \right)$$

The dual problem maximizes the minimum of $L(\mathbf{x}_i, \boldsymbol{\lambda})$ over $\boldsymbol{\lambda}$. It can be solved with two levels of optimization. At the lower level, the Lagrangian is decoupled into m sub-problems over i.

$$\min_{\mathbf{x}_i} \ f_i(\mathbf{x}_i) + \boldsymbol{\lambda}^T h_i(\mathbf{x}_i), \ \forall i \in \{1, ..., m\} \tag{2.19}$$

Each sub-problem is solved for a fixed value of the Lagrange multiplier $\boldsymbol{\lambda}$. The obtained solution, for each $i \in \{1, ..., m\}$, is denoted as $g_i(\boldsymbol{\lambda})$. It is then used in the master dual problem, that updates $\boldsymbol{\lambda}$ by solving the dual problem:

$$\begin{aligned} &\max_{\boldsymbol{\lambda}} \ g(\boldsymbol{\lambda}) = \sum_{i=1}^{m} g_i(\boldsymbol{\lambda}) + \boldsymbol{\lambda}^T \mathbf{c} \\ &\text{s.t. } \boldsymbol{\lambda} \geq \mathbf{0} \end{aligned} \tag{2.20}$$

The master problem can be solved with the sub-gradient method. The two steps are iterated until convergence. As the dual problem is convex, the obtained solution is necessarily globally optimal. If the duality gap is zero, then the dual optimum is also equal to the primal optimum.
Remark: Primal decompositions are also possible in some cases. However, as we will not use them in the thesis, we do not detail them in this section.

2.2.3 Newton's method

The KKT conditions may not always lead to a simple analytical solution. In that case, it is necessary to use numerical methods in order to solve the convex optimization problem (2.12). An efficient and classical method for unconstrained and equality-constrained convex optimization problem is the Newton's method. This section provides an overview of this method. Detailed convergence proofs can be found in [24].

Descent methods

Unconstrained convex optimization problems are solved by descent methods. The solution $\mathbf{x}^*$ that minimizes $f(\mathbf{x})$ is obtained at $\nabla f(\mathbf{x}^*) = \mathbf{0}$. Descent methods produce a minimizing sequence $\mathbf{x}^k$ such that

$$f(\mathbf{x}^{k+1}) < f(\mathbf{x}^k),\ \forall k \geq 0$$

The sequence is obtained with :

$$\mathbf{x}^{k+1} = \mathbf{x}^k + t^k \Delta\mathbf{x}^k,\ \forall k \geq 0$$

where $t^k \geq 0$ is the step size, and $\Delta\mathbf{x}^k$ is the descent direction. Due to the convexity of f, $\Delta\mathbf{x}^k$ must verify $\nabla f(\mathbf{x}^k)^T \Delta\mathbf{x}^k < 0$.
The general descent method iteratively updates $\mathbf{x}^k$ until reaching the optimum value. It starts from an initial vector $\mathbf{x}^0 \in \mathrm{dom} f$. Then it uses the following iterative process:

1. Determine a descent direction $\Delta\mathbf{x}$.
2. Choose a step size $t > 0$ (line search).
3. Update $\mathbf{x} = \mathbf{x} + t\Delta\mathbf{x}$.

The process stops when a given accuracy criterion on $\nabla f(\mathbf{x}) = 0$ is fulfilled. Line search determines the step size t useful to update the direction. An efficient line search method is backtracking line search. It depends on two parameters $0 < \alpha < 0.5$ and $0 < \beta < 1$, and consists in decreasing t as long as the stopping criterion is not fulfilled. Starting from $t = 1$, the condition $f(\mathbf{x}+t\Delta\mathbf{x}) \leq f(\mathbf{x})+\alpha t\ \nabla f(\mathbf{x}^T\Delta\mathbf{x})$ is checked. If it is not fulfilled, then $t = \beta t$.

Newton's method for unconstrained convex optimization problems

The convergence of the iterative process is improved when the descent direction is a steepest descent direction. For a given norm $\|.\|$, a normalized steepest descent direction gives the largest decrease in the linear approximation of function f : $\Delta\mathbf{x} = \arg\ \min\ \left\{\nabla f(\mathbf{x})^T\mathbf{v} / \|v\| = 1\right\}$
Newton's method for unconstrained convex optimization problems chooses the descent direction as the steepest descent direction at $\mathbf{x}$, for the quadratic norm defined by the Hessian $\nabla^2 f(\mathbf{x})$: $\|v\|_{\nabla^2 f(\mathbf{x})} = \left(\mathbf{v}^T\nabla^2 f(\mathbf{x})\mathbf{v}\right)^{1/2}$. It is equal to:

$$\Delta\mathbf{x}_{nt} = -\nabla^2 f(\mathbf{x})^{-1}\nabla f(\mathbf{x}) \tag{2.21}$$

And the stopping criterion for the iterative algorithm is: $\lambda(\mathbf{x})^2/2 < \epsilon$, where $\epsilon > 0$ is the target accuracy, and $\lambda(\mathbf{x})$ is the Newton decrement:

$$\lambda(\mathbf{x})^2 = \nabla f(\mathbf{x})^T\nabla^2 f(\mathbf{x})^{-1}\nabla f(\mathbf{x}) \tag{2.22}$$

Newton's method for equality constrained convex optimization problems

Newton's method can be extended to equality-constrained convex optimization problems. Two modifications are included: first, the initial vector must be feasible (it must lie in $\mathrm{dom} f$ and satisfy the equality constraints: $\mathbf{Ax}^0 = \mathbf{b}$), and the Newton step must take into account the equality constraint. The Newton step is characterized by:

$$\begin{pmatrix} \nabla^2 f(\mathbf{x}) & \mathbf{A}^T \\ \mathbf{A} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{\Delta}\mathbf{x_{nt}} \\ \mathbf{w} \end{pmatrix} = \begin{pmatrix} -\nabla f(\mathbf{x}) \\ \mathbf{0} \end{pmatrix} \tag{2.23}$$

$\begin{pmatrix} \nabla^2 f(\mathbf{x}) & \mathbf{A}^T \\ \mathbf{A} & \mathbf{0} \end{pmatrix}$ is called the KKT matrix. It must be invertible at each $\mathbf{x}$. $\mathbf{w}$ is the optimal dual variable for the second-order Taylor approximation of the original problem near $\mathbf{x}$. Finally, Newton's method for equality-constrained convex optimization problems is:
Start from $\mathbf{x}^0 \in \mathrm{dom} f$ with $\mathbf{Ax}^0 = \mathbf{b}$. Define an accuracy level $\epsilon > 0$. Then use the following iterative process:

1. Compute the Newton step $\Delta\mathbf{x}_{nt}$ by solving (2.23). Determine the Newton decrement $\lambda(\mathbf{x}) = \left(\nabla f(\mathbf{x})^T \nabla^2 f(\mathbf{x})^{-1} \nabla f(\mathbf{x})\right)^{1/2}$
2. Quit if $\lambda(\mathbf{x})^2/2 < \epsilon$.
3. Choose a step size $t > 0$ by backtracking line search.
4. Update $\mathbf{x} = \mathbf{x} + t\Delta\mathbf{x}_{nt}$.

2.2.4 Geometric programming

Geometric programming is a class of non-linear optimization problems that can be turned into convex optimization problems through a logarithmic change of variables. It is often used in wireless applications, as many optimization problems belong to the class of geometric programming.
A monomial is a function $f : \mathbb{R}^m_{++} \to \mathbb{R}$ of the following form:

$$f(\mathbf{x}) = dx_1^{a(1)} x_2^{a(2)} ... x_n^{a(n)}$$

where $d \geq 0$ is a multiplicative constant and $a(j)$ are exponential constants in $\mathbb{R}$, $j \in \{1, ..., n\}$. A posynomial is a sum of monomial:

$$f(\mathbf{x}) = \sum_{k=1}^{K} d_k x_1^{a_k^{(1)}} x_2^{a_k^{(2)}} ... x_n^{a_k^{(n)}}$$

A geometric program in standard form is defined as:

$$\begin{aligned} &\min_{\mathbf{x}} \ f_0(\mathbf{x}) \\ &\text{s.t. } f_i(\mathbf{x}) \leq 0, \forall\ i \in \{1, ..., m\} \\ &\text{s.t. } h_l(\mathbf{x}) = 0, \forall\ l \in \{1, ..., p\} \end{aligned} \tag{2.24}$$

where $f_i,\ i \in \{0, ..., m\}$ are posynomials:

$$f_i(\mathbf{x}) = \sum_{k=1}^{K_i} d_{ik} x_1^{a_{ik}^{(1)}} x_2^{a_{ik}^{(2)}} ... x_n^{a_{ik}^{(n)}}$$

and $h_l,\ l \in \{1, ..., p\}$, are monomials:

$$h_l(\mathbf{x}) = d_l x_1^{a_l^1} x_2^{a_l^2} ... x_n^{a_l^n}$$

In standard form, a geometric program is not a convex optimization problem, because posynomials are not convex functions. Let us introduce a logarithmic change of variables: $y_i = \log(x_i)$ and multiplicative constants: $b_{ik} = \log(d_{ik})$ and $b_l = \log(d_l)$. We also set $\mathbf{a}_{ik} = \left[a_{ik}^1, a_{ik}^2, ..., a_{lk}^n\right]^T$. The optimization problem becomes:

$$\begin{aligned}
&\min_{\mathbf{y}} \sum_{k=1}^{K_0} \exp(\mathbf{a}_{0k}^T \mathbf{y} + b_{0k}) \\
&\text{s.t. } \sum_{k=1}^{K_i} \exp(\mathbf{a}_{ik}^T \mathbf{y} + b_{ik}) \leq 1, \forall\ i \in \{1, ..., m\} \\
&\text{s.t. } \mathbf{a}_l^T \mathbf{y} + b_l = 0,\ l \in \{1, ..., p\}
\end{aligned} \tag{2.25}$$

This optimization problem is equivalent to:

$$\begin{aligned}
&\min_{\mathbf{y}} \log \sum_{k=1}^{K_0} \exp(\mathbf{a}_{0k}^T \mathbf{y} + b_{0k}) \\
&\text{s.t. } \log \sum_{k=1}^{K_i} \exp(\mathbf{a}_{ik}^T \mathbf{y} + b_{ik}) \leq 1, \forall\ i \in \{1, ..., m\} \\
&\text{s.t. } \mathbf{a}_l^T \mathbf{y} + b_l = 0, \forall\ l \in \{1, ..., p\}
\end{aligned} \tag{2.26}$$

Problem (2.26) is convex because the log-sum-exp function $f(\mathbf{x}) = \log(\exp x_1 + ... + \exp x_n)$ is convex on $\mathbb{R}^n$.

Geometric programming is often used in the literature to solve resource allocation problems in communications. For instance, power control with various objectives is studied with geometric programming in [27]. It should be noted that many optimization problems in multi-cell wireless communications become geometric programs only when a high SINR assumption is fulfilled, i.e. when the approximation $\log(1 + \text{SINR}) \approx \log(\text{SINR})$ holds.

Chapter 3

Fairness increase through distributed network coordination

3.1 Introduction

In Chapters 3 and 4, we propose a method to increase fairness in distributed networks. It uses network coordination, where two BSs equipped with a single antenna form a virtual MIMO array, in order to bring additional diversity to their users.
Virtual MIMO consists in transmitting signals to the same receiver from various locations. From the receiver's perspective, the channel is then equivalent to a MIMO channel with several transmit antennas. The diversity gain brought by virtual MIMO depends on the path loss and shadowing loss differences between the involved transmitters and the receiver's locations, on top of the differences in fading conditions. Virtual MIMO has been introduced for multi-hop networks in [28] [29]. It belongs to the set of cooperative communication techniques. This set contains all the techniques where several transmitters or receivers share some of their resources to increase the link capacity. The most usual cooperative techniques are relaying techniques and network coordination. Relays are fixed or mobile radio nodes, that are used to forward data from a source node to a destination node. For instance, in cellular networks, mobile terminals may serve as relays to other mobile terminals in uplink, or to BSs in downlink, in order to increase cells' coverage. Fixed nodes, dedicated to relaying, may also be considered. These nodes are then called repeaters. Whereas the 'relay' designation may apply to various scenarios, network coordination refers to a specific downlink technique for cellular networks. In that case, the transmission from several BSs to their users is jointly coordinated, thus turning the channel into an equivalent MIMO broadcast channel.
In most network coordination studies, it is assumed that all coordinated BSs simultaneously have access to the data to be transmitted. Thus, a central controller managing the buffers

of all users and connected to all BSs through high speed links must be included in the network architecture. In Chapters 3 and 4, we consider distributed networks where each BS manages its terminals' buffers, and where the inter-BS link may not allow instantaneous transmission. Consequently, a causality constraint is imposed on the inter-BS transmission. Our method can be either seen under the scope of network coordination with a causality constraint, or under the scope of relaying with a source-to-relay link with perfect SNR. To simplify notations, in the following, we will refer to our method as a network coordination method.

We study downlink resource allocation when causal network coordination is triggered for the users located at the border of cells. The impact of network coordination depending on the fairness of the power allocation objective is evaluated. Four power allocation objectives are considered. Our proposed method is assessed in terms of cell throughput, and in terms of fairness of the data rates' distribution among users.

The chapter first presents the state of the art on cooperative communication techniques. Then in Section 3.3, we detail the proposed causal network coordination process, and the adapted power control designed for this process. Network coordination is performed over two TTIs and involves two BSs. The direct BS is responsible for determining its set of users requiring an additional capacity gain. Coordination is triggered without knowledge of the complete CSI, prior to coordinated transmission, using a path-loss based criterion. Then, a simple subcarrier allocation method, aiming at assigning the same number of subcarriers to all direct users, is carried out. The direct BS is responsible for requesting the required subcarriers for all coordinated users on their respective coordinated BS. Finally, power control is performed independently per BS. It is further decomposed into two sub-problems per BS: power allocation for direct users, and power allocation for coordinated users. Power allocation is performed iteratively over all BSs, in order to take inter-cell interference into consideration. It is detailed in Section 3.4, with four power allocation objectives: Globally Optimal, Max-Min Fair, Proportional Fair and Harmonic Mean Fair. In Section 3.5, the performance results of the proposed method are assessed and compared with a simple case without network coordination. Numerical results show that network coordination increases the fairness with all objectives, and increases the sum throughput with fair objectives. With the unfair objective, Globally Optimal, the fairness increase comes at the expense of a high peak data rate decrease, thus leading to a sum throughput decrease.

The main contributions of this chapter are:

- Causal network coordination is introduced. The proposed scheme is adapted for distributed networks where the inter-BS link may not allow instantaneous data transmission. A resource allocation strategy for this cooperative scheme is then detailed. As network coordination is costly in terms of resource consumption, it is restricted to the users located at the border of cells.

- Power control for the proposed causal network coordination is detailed. It is based on an iterative process over all BSs, and on a separation of power allocation between direct and coordinated users of the same BS.

- The proposed network coordination and power control strategy is tested with different power allocation objectives. The conclusions are that network coordination is very useful to increase the data rate of the users located at the border of cells, thanks to diversity gain and inter-cell interference decrease. This leads to a fairness improvement with all objectives, and to a throughput increase with fair objectives. When the power allocation objective is unfair, the sum throughput decreases due to the limitation in the resources allocated to the users in best radio conditions.

3.2 State of the art

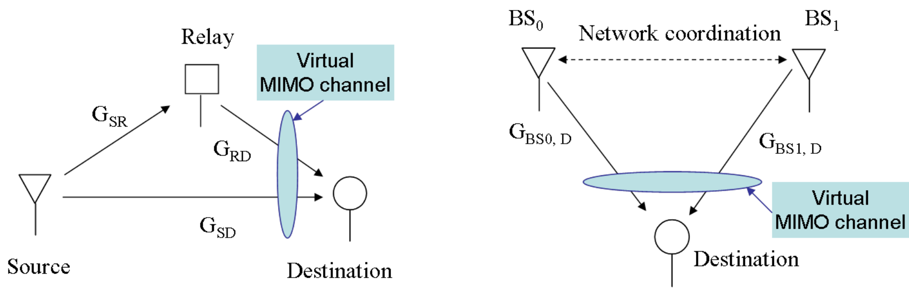


Figure 3.1: Cooperation techniques: relaying (left) and network coordination (right)

3.2.1 Relaying techniques

Relaying techniques contain many different methods for wireless applications, in cellular and ad hoc networks [30]. The most general definition is the following: a source node transmits its data to a relay node, that retransmits it to the destination node. There may be several relays retransmitting the data simultaneously, and the relaying process may consists of several hops. If the source and relays transmit the same information, possibly not within the same time slot due to causality constraints, and if the destination node is able to jointly process this information, then the relay channel is equivalent to a MIMO channel. This type of relaying protocol is called 'Non Orthogonal'. Relaying is orthogonal whenever the destination does not directly receive information from the source. Besides, the causality constraint is almost always assumed when studying relays. It means that the relay cannot instantaneously retransmit the data it receives. A special causality case is the half-duplex constraint, where the relay cannot simultaneously listen and transmit. Relaying protocols have been widely studied in uplink cellular networks [31] [32] [33]. A major research area is the definition of adapted protocols derived from Amplify-and-Forward (AF) and Decode-and-Forward (DF) [34] [35] [36] [37]. In AF, the relay simply forwards the received data with an amplification factor. Its main limitation is the induced noise amplification. In DF, the relay first decodes the received data, and then re-encodes and retransmits it. It avoids noise amplification, but requires numerical processing at the relay. Resource allocation adapted to cooperative protocols has rarely been treated in the literature. Power allocation for uplink cooperation has been studied in [38] [39] [40]. Most papers on downlink relaying deal with coverage improvement, using relays as repeaters without diversity [41].

3.2.2 Network coordination

Network coordination consists in coordinating the transmission of several BSs of a cellular network toward the same users. The involved BSs should be linked by a high-speed backbone. The BSs may coordinate their transmission to increase diversity and to decrease interference. The time-based transmission of adjacent BSs may be coordinated via cooperative scheduling [42], so as to prevent interfering users from simultaneously transmitting. This method is an extension of the dynamic planning techniques described in [9] to time-based scheduling. The BSs involved in network coordination act together as a MIMO transmitter. If several users are served by this set of BSs, the equivalent downlink channel is a broadcast channel, with one transmitter and several independent receivers. Thus, MIMO transmit processing can be used on this virtual MIMO channel in order to increase links' capacity. All precoding methods applicable to the broadcast channel extend to network coordination. The optimal joint encoding technique is dirty paper coding [43]. It provides an upper bound on the achievable multi-user capacity. However, dirty paper coding is far too complex to be implemented in real practical coordinated networks. In order to derive more practical methods, MIMO linear precoding has been investigated for network coordination [44] [45] [46]. A generalized ZF precoder, also referred to as block diagonalization, eliminates the interference from other users, as the transmission of each user lies within the null space of the others' transmission. An MMSE precoder may be used instead of ZF to avoid noise increase. It should be noted that dirty paper coding and linear precoding require full CSI on all links, available at all BSs or, equivalently, at the central controller responsible for coordinated transmission. The main limitation of network coordination is consequently the amount of feedback to be managed, and the additional backbone load induced by multi-cell treatments [47].

3.3 Resource allocation for network coordination

3.3.1 Causal network coordination

Let us consider a network $\mathcal{N}$ with N_{BS} base stations and $K_{\mathcal{N}}$ users. All BSs use OFDMA with the same FFT size, N_{FFT}. The total available bandwidth is B, L_{SC} is the number of subcarriers per BS, and $B_{\mathrm{SC}} = B/L_{\mathrm{SC}}$ is the bandwidth per subcarrier. Only one user is served per subcarrier in each cell.

The studied network coordination method coordinates subcarrier and power allocations between two BSs serving the same mobile terminal. Each user k is attached to a serving (denoted as 'direct') BS, $BS_{d,k}$. This BS is responsible for getting the user's data from the network layer, and transmitting it to the user. If $BS_{d,k}$ estimates that user k is in bad radio conditions and consequently has a high probability of achieving a low data rate, then $BS_{d,k}$ identifies a neighboring BS that may coordinate its subcarrier and power allocations so as to serve user k efficiently. Let $BS_{c,k}$ be the identified BS for coordination. $BS_{d,k}$ sends a message to $BS_{c,k}$ via the inter-BS link, composed of the data to be transmitted within each subcarrier, and of the subcarriers' indexes. Both direct and coordinated BSs should be transmitting on the same set of subcarriers, so as to form a virtual MIMO array. Let Θ_k be that set of subcarriers.

In order to derive general results that will be valid for different types of network infrastructures, we impose the causality constraint on the link between the direct BS and the BS chosen for network coordination. However, the inter-BS channel is assumed perfect in terms of SNR (SNR $\rightarrow \infty$). This assumption is reasonable if the BSs are connected by a wired link (which may be required for seamless handover management in 3GPP LTE or WiMAX), or by directional radio links. Besides, we assume that the cyclic prefix compensates for the delay spread from both BSs, so that there is no ISI induced by network coordination.
The proposed transmission technique can be either seen under the scope of network coordination, or under the scope of downlink cooperation. In the second case, the considered transmission may be viewed as the Non Orthogonal Amplify-and-Forward (NAF) cooperation protocol from [36] with one relay and two TTIs. In this protocol, the source communicates with the relay and with the destination during the first TTI. In the second TTI, both the source and the relay communicate with the destination. It has been shown in [36] that NAF leads to higher mutual information than the other AF one-relay, two-TTIs protocols (Orthogonal relaying protocol and simultaneous transmission of source and relay in the second time-slot only). NAF is therefore the best AF protocol in terms of achievable rate.
Let $\mathbf{x}_k^l = \left[x_{k,1}^l, x_{k,2}^l\right]'$ be the vector of symbols transmitted by the direct BS in subcarrier l. The vector of symbols received by user k in the two TTIs, $\mathbf{y}_k^l = \left[y_{k,1}^l, y_{k,2}^l\right]'$, is equal to

$$y_{k,1}^l = h_{d,k}^l \sqrt{\frac{g_{d,k} P_{d,k}^l}{I_k^l}} x_{k,1}^l + n_{k,1}^l \tag{3.1}$$

$$y_{k,2}^l = h_{c,k}^l \sqrt{\frac{g_{c,k} P_{c,k}^l}{I_k^l}} x_{k,1}^l + h_{d,k}^l \sqrt{\frac{g_{d,k} P_{d,k}^l}{I_k^l}} x_{k,2}^l + n_{k,2}^l \tag{3.2}$$

where

- $P_{d,k}^l$ (resp. $P_{c,k}^l$) is the transmit power from the direct (resp. coordinated) BS to user k in subcarrier l.
- $g_{d,k}$ (resp. $g_{c,k}$) is the power path loss (including shadowing) from the direct (resp. coordinated) BS to user k.
- $h_{d,k}^l$ (resp. $h_{c,k}^l$) is the amplitude of the fast fading channel coefficient between the direct (resp. coordinated) BS and user k, in subcarrier l. It remains constant over two TTIs.
- I_k^l is the inter-cell interference plus noise received by user k in subcarrier l. $\mathbf{n}_k^l \sim \mathcal{CN}(0, \mathbf{I}_2)$ is AWGN with variance N_0.

The coordinated network is represented on Figure 3.2, and an example of coordination process is illustrated on Figure 3.3 with BS_1 as the direct BS, and BS_0 as the coordinated BS of user k.

We consider an idealistic case with full CSI at the BSs. Practical implementations should be based on partial knowledge of the channel state, provided by statistical estimation of the channel parameters. However, it has been shown for the MIMO channel, for the most general case of Ricean fading, that there is no analytical solution for optimizing the ergodic capacity,

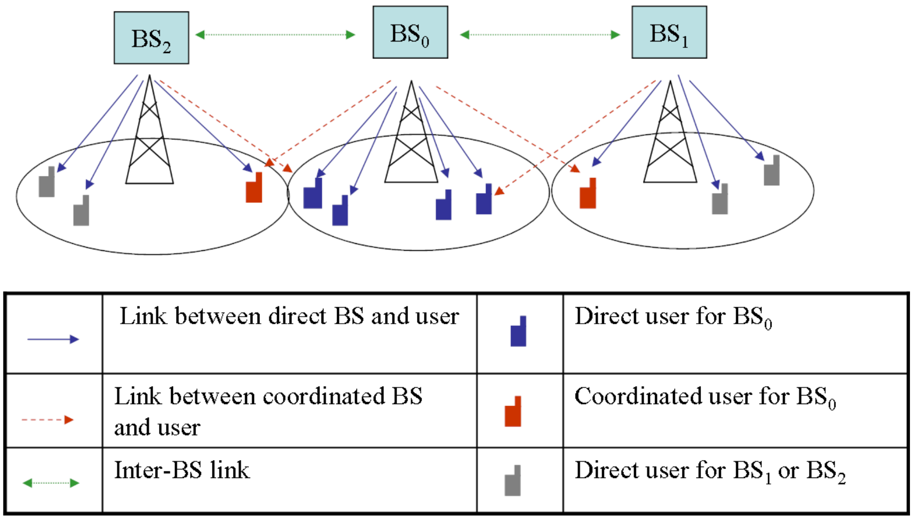


Figure 3.2: Causal network coordination

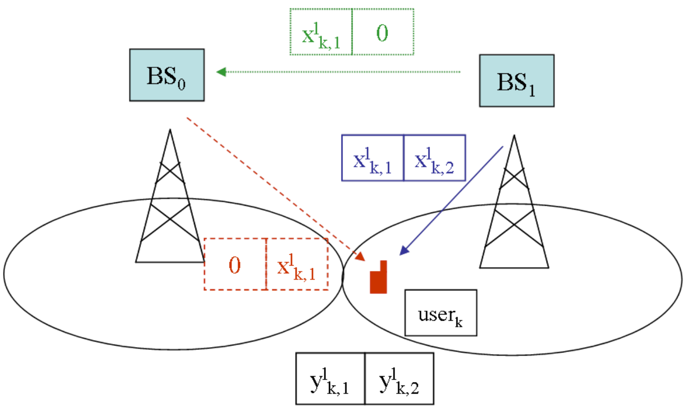


Figure 3.3: Example of coordination

even when the channel mean and covariance matrices are available [48]. Therefore, complex numerical techniques are required. The same limitations should apply to the considered cooperative channel which is similar to a MIMO channel. Thus, we limit our study to full CSI, so as to provide a first, idealistic assessment of the proposed method.
The transmission channel can be modeled as follows:

$$\mathbf{y}_k^l = \mathbf{H}_k^l \mathbf{x}_k^l + \mathbf{n}_k^l \tag{3.3}$$

$\mathbf{H}_k^l$ is the equivalent channel matrix for user k in subcarrier l,

$$\mathbf{H}_k^l = \begin{pmatrix} h_{d,k}^l \sqrt{\frac{g_{d,k} P_{d,k}^l}{I_k^l}} & 0 \\ h_{c,k}^l \sqrt{\frac{g_{c,k} P_{c,k}^l}{I_k^l}} & h_{d,k}^l \sqrt{\frac{g_{d,k} P_{d,k}^l}{I_k^l}} \end{pmatrix} \tag{3.4}$$

We assume that $\mathbb{E}\{\mathbf{x}_k^l {\mathbf{x}_k^l}^H\} = \mathbf{I}_2$. The capacity of user k in subcarrier l is [12]

$$\begin{aligned} \mathrm{C}_k^l &= \frac{1}{2} \log_2 \left(\det \left(\mathbf{I}_2 + \mathbf{H}_k^l (\mathbf{H}_k^l)^H \right) \right) \\ &= \frac{1}{2} \log_2 \left(\left(1 + \frac{G_{d,k}^l P_{d,k}^l}{I_k^l} \right)^2 + \frac{G_{c,k}^l P_{c,k}^l}{I_k^l} \right) \end{aligned} \tag{3.5}$$

where $G_{d,k}^l = \left| h_{d,k}^l \right|^2 g_{d,k}$ and $G_{c,k}^l = \left| h_{c,k}^l \right|^2 g_{c,k}$.
The data rate of user k is consequently equal to

$$\begin{aligned} R_k &= B_{\mathrm{SC}} \sum_{l \in \Theta_k} \mathrm{C}_k^l \\ &= \frac{B_{\mathrm{SC}}}{2} \sum_{l \in \Theta_k} \log_2 \left(\det \left(\mathbf{I}_2 + \mathbf{H}_k^l (\mathbf{H}_k^l)^H \right) \right) \end{aligned} \tag{3.6}$$

3.3.2 Distributed power control

We now consider the influence of the proposed network coordination protocol on power control. Let $\phi(R_k)$ be a concave function of the data rate R_k of user $k \in \{1, ..., K_{\mathcal{N}}\}$. Power control aims at allocating powers to all users of the network, in order to solve an optimization problem of the global form [49] [8]

$$\begin{aligned} &\max_{\{\mathbf{P}_1, ..., \mathbf{P}_{N_{\mathrm{BS}}}\}} \sum_{k=1}^{K_{\mathcal{N}}} \frac{\phi(R_k)^{1-\alpha}}{1-\alpha} \\ &\text{s. t. } \sum_{l=1}^{L_{\mathrm{SC}}} P_{n_{\mathrm{BS}}}^l \le P_{\max}, \forall n_{\mathrm{BS}} \in \{1, ..., N_{\mathrm{BS}}\} \\ &\text{s. t. } P_{n_{\mathrm{BS}}}^l \ge 0, \ \forall (n_{\mathrm{BS}}, l) \in \{1, ..., N_{\mathrm{BS}}\} \times \{1, ..., L_{\mathrm{SC}}\} \end{aligned} \tag{3.7}$$

where $P_{n_{\mathrm{BS}}}^l$ is the power transmitted by BS n_{BS} in subcarrier l, and $P_{\max}$ is the maximum power per BS for downlink transmission. The optimization variables consist of the set of power values transmitted by all BSs in all subcarriers. α is a coefficient that indicates the fairness of the problem.
Although (3.7) may be turned into a convex optimization problem under certain assumptions on the utility function (see for instance [50]), we cannot conclude in the general case.

Besides, because of network coordination, the total SINR is an additive function of the SINR on the direct and coordinated links, therefore its inverse is not a posynomial function, which could be solved through geometric programming [50]. If inter-cell interference was not considered, then the problem would be convex and a global optimum could be obtained [24]. However, its determination would require a global knowledge of all channel gains, which is not feasible in distributed networks.

As inter-cell interference must be considered, we propose to perform power allocation for each BS independently and to iterate the process. At each iteration, the inter-cell interference received by each user is re-computed, depending on the power transmit values of the previous iteration. The iterative process does not necessarily converge for all power values, as the distributed problem may not converge toward a fixed point -which may be the global optimum. A numerical convergence study is carried out in Section 3.5.1. It shows that it is possible to set the number of iterations for a given convergence rate on the power values.

Besides, in order to turn problem (3.7) into a convex optimization problem, whatever the concave function ϕ is, (3.7) must not only be distributed over BS, but also be separated into two problems per BS: optimization over the direct users and optimization over the coordinated users.

Therefore, we impose that a fixed part of the total power be dedicated to coordinated users. Let $\mathcal{S}_d$ be the set of users served on their direct link by BS n_{BS}, and $\mathcal{S}_c$ be the set of users that are coordinated by n_{BS}, with $|\mathcal{S}_d| = K_d$ and $|\mathcal{S}_c| = K_c$. The power dedicated to coordinated users is then defined as

$$P_{\max,c} = \frac{K_c}{(K_c + K_d)} P_{\max}$$

and the power dedicated to direct users is $P_{\max,d} = P_{\max} - P_{\max,c}$. As a consequence, for each BS, two power controls are performed in parallel: on the one hand, power is allocated to direct users with power constraint $\sum_{k \in \mathcal{S}_d} \sum_{l \in \Theta_k} P_{d,k}^l = P_{\max,d}$; and on the other hand, power is allocated to coordinated users with power constraint $\sum_{k \in \mathcal{S}_c} \sum_{l \in \Theta_k} P_{c,k}^l = P_{\max,c}$.

3.3.3 Joint coordination - resource allocation strategy

We propose a joint coordination - resource allocation strategy for the coordinated channel, which is made of three steps. It is performed independently in each BS, as it only requires feedback from each user to its direct and coordinated BSs.

1. *Identification of the users of the BS that require coordination, and of their coordinated BS.* We assume that coordination is useful for users at the border of each cell: therefore, it is requested for users which have a path loss difference of less than Δ dB between their direct BS and the best (in terms of path loss) neighboring BS. Δ is a parameter that should be set depending on the cellular environment and deployment. Δ should lead to a trade-off between improving the performance of coordinated users, and maintaining the performance of the other users.

2. *Subcarrier allocation*: The same subcarrier allocation method is used whatever the power control objective is, in order to compare them. A proportion of the subcarriers of each BS is dedicated to coordinated users, with the restriction that each user should

have at least one subcarrier allocated on its direct link. The proportion of subcarriers dedicated to coordinated users is

$$L_{\mathrm{SC},c} = \min \left\{ \left\lfloor \frac{K_c}{(K_c + K_d)} L_{\mathrm{SC}} \right\rfloor, L_{\mathrm{SC}} - K_d \right\}$$

As a path loss criterion Δ is used for coordination request, the BS knows its number of coordinated users K_c before resource allocation. Subcarriers are first allocated to direct users on the remaining $L_{\mathrm{SC},d} = L_{\mathrm{SC}} - L_{\mathrm{SC},c}$ subcarriers. The subcarrier allocation process aims at allocating the same number of subcarriers to all users. Each subcarrier l is allocated to the user that maximizes channel coefficient $G^l_{d,k}$, and that has less allocated subcarriers than the user with the maximum number of subcarriers. Then, the direct BS requests the list of subcarriers allocated to user k on its coordinated BS. Coordinated users are favored over direct users regarding subcarrier allocation. Consequently, if the subcarrier required for coordination is already assigned to a direct user on BS_c, and if some subcarriers are still available, the direct user is re-allocated to the free subcarrier of BS_c that maximizes its channel coefficient. At the end of the subcarrier allocation step, if a user cannot obtain the same subcarriers on its coordinated link as on its direct link, then network coordination is eventually not triggered for that user.

3. *Iterative power control*: Iterative power control performs power allocation over coordinated users, and over direct users, within each BS independently. The inter-cell interference values of the previous iteration are considered. With a fixed inter-cell interference, each local power allocation problem is convex and has a global optimum. The general iterative power control framework is depicted on Fig. 3.4. It is detailed in Section 3.4.

3.4 Power allocation objectives with varying fairness

In this section, we consider problem (3.7) distributed over all BSs, and focus on BS n_{BS} of $\mathcal{N}$. For each user $k \in \mathcal{S}_d \cup \mathcal{S}_c$ served by n_{BS}, the performance indicator is the data rate R_k (3.6):

$$\phi(R_k) = R_k = B_{\mathrm{SC}} \sum_{l \in \Theta_k} \mathrm{C}^l_k$$

3.4.1 Definitions

In the following, we evaluate the performance of the proposed causal network coordination method with different power allocation objectives: Globally Optimal, Max-Min Fair, Proportional Fair and Harmonic Mean Fair. $\mathbf{P}_d$ (resp. $\mathbf{P}_c$) is the vector of power levels on the subcarriers allocated to direct (resp. coordinated) users by n_{BS}. It should be noted that, as R_k is an increasing function of $\mathbf{P}_d$ and $\mathbf{P}_c$ (for fixed inter-cell interference values), in all studied cases, the objective function is maximized when the sum power reaches the maximum allowed power, $P_{\max,d}$ or $P_{\max,c}$. Thus we use the equality constraint in the optimization problems.

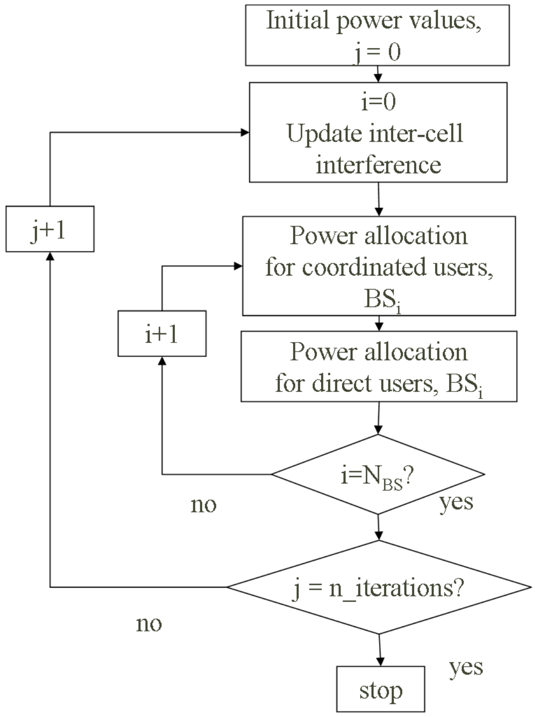


Figure 3.4: Iterative power control with network coordination

1. Globally Optimal allocation is achieved when the fairness coefficient in (3.7) is $\alpha = 0$. It is a greedy optimization method that favors users in good radio conditions, but may leave users in bad conditions unserved. It is defined as

$$\begin{aligned}
&\max_{\mathbf{P}_d} \sum_{k\in\mathcal{S}_d} R_k \\
&\text{s.t. } \sum_{k\in\mathcal{S}_d}\sum_{l\in\Theta_k} P^l_{d,k} = P_{\max,d} \\
&\text{s.t. } P^l_{d,k} \geq 0,\ \forall k\in\mathcal{S}_d,\ l\in\Theta_k
\end{aligned} \tag{3.8}$$

for direct users, and

$$\begin{aligned}
&\max_{\mathbf{P}_c} \sum_{k\in\mathcal{S}_c} R_k \\
&\text{s.t. } \sum_{k\in\mathcal{S}_c}\sum_{l\in\Theta_k} P^l_{c,k} = P_{\max,c} \\
&\text{s.t. } P^l_{c,k} \geq 0,\ \forall k\in\mathcal{S}_c,\ l\in\Theta_k
\end{aligned} \tag{3.9}$$

for coordinated users. In the following, we only write down the optimization problem for direct users. The optimization problems for coordinated users are similar.

2. Max-Min Fairness is obtained when $\alpha \to \infty$. It aims at serving all users, but is constrained by the users in worst radio conditions. Therefore, it may lead to poor

cell's sum throughput values.

$$\begin{aligned}
&\max_{\mathbf{P}_d} \ \min_{k\in\mathcal{S}_d} R_k \\
&\text{s.t.} \ \sum_{k\in\mathcal{S}_d}\sum_{l\in\Theta_k} P_{d,k}^l = P_{\max,d} \\
&\text{s.t.} \ P_{d,k}^l \geq 0, \ \forall k\in\mathcal{S}_d, \ l\in\Theta_k
\end{aligned} \tag{3.10}$$

3. Proportional Fairness corresponds to $\alpha \to 1$:

$$\begin{aligned}
&\max_{\mathbf{P}_d} \ \sum_{k\in\mathcal{S}_d} \log(R_k) \\
&\text{s.t.} \ \sum_{k\in\mathcal{S}_d}\sum_{l\in\Theta_k} P_{d,k}^l = P_{\max,d} \\
&\text{s.t.} \ P_{d,k}^l \geq 0, \ \forall k\in\mathcal{S}_d, \ l\in\Theta_k
\end{aligned} \tag{3.11}$$

4. Harmonic Mean Fairness is achieved when $\alpha = 2$:

$$\begin{aligned}
&\max_{\mathbf{P}_d} \ -\left(\sum_{k\in\mathcal{S}_d} \frac{1}{R_k}\right) \\
&\text{s.t.} \ \sum_{k\in\mathcal{S}_d}\sum_{l\in\Theta_k} P_{d,k}^l = P_{\max,d} \\
&\text{s.t.} \ P_{d,k}^l \geq 0, \ \forall k\in\mathcal{S}_d, \ l\in\Theta_k
\end{aligned} \tag{3.12}$$

Proportional Fair and Harmonic Mean Fair power allocations are trade-offs between Globally Optimal and Max-Min Fair allocations.

3.4.2 Detailed power allocation process

For direct and coordinated users independently, power allocation is separated into two steps on each BS. Power is first allocated globally per user, considering the global optimization problem and sum power constraint. Then for each user, the previously determined power is allocated to its subcarriers, with the objective to maximize the data rate. Each power allocation problem is convex, and can be solved via classical convex optimization techniques such as Newton's method.
More specifically, in the first step, we consider that each subcarrier's coefficient is equal to the average direct link channel coefficient $\tilde{G}_{d,k} = \frac{1}{l_{\mathrm{SC},k}} \sum_{l\in\Theta_k} G_{d,k}^l$ and to the average coordinated link channel coefficient $\tilde{G}_{c,k} = \frac{1}{l_{\mathrm{SC},k}} \sum_{l\in\Theta_k} G_{c,k}^l$, where $|\Theta_k| = l_{\mathrm{SC},k}$. The approximate data rate of user k is then

$$\tilde{R}_k = \frac{B_{\mathrm{SC}} \ l_{\mathrm{SC},k}}{2} \log_2 \left(\det\left(\mathbf{I}_2 + \tilde{\mathbf{H}}_k (\tilde{\mathbf{H}}_k)^H\right)\right)$$

where $\tilde{\mathbf{H}}_k$ is the equivalent channel matrix corresponding to $\tilde{G}_{d,k}$ and $\tilde{G}_{c,k}$. As the subcarrier allocation method defined in Section 3.3.3 leads to the same number of subcarriers for all users, the global optimization problem can be simplified as follows (for instance for

Globally Optimal allocation and direct users):

$$\begin{aligned}
&\max_{\mathbf{P}_d} \sum_{k\in\mathcal{S}_d} \log_2\left(\left(1+\frac{\tilde{G}_{d,k}P_{d,k}}{I_k}\right)^2+\frac{\tilde{G}_{c,k}P_{c,k}}{I_k}\right)\\
&\text{s.t. } \sum_{k\in\mathcal{S}_d} P_{d,k}=P_{\max,d}\\
&\text{s.t. } P_{d,k}\geq 0,\ \forall k\in\mathcal{S}_d
\end{aligned} \tag{3.13}$$

where $I_k = \sum_{l\in\Theta_k} I^l_{d,k}$ is the interference plus noise received by user k on all its subcarriers, and $P_{d,k} = \sum_{l\in\Theta_k} P^l_{d,k}$ (resp. $P_{c,k}$) is the sum power transmitted by the direct (resp. coordinated) BS to user k. The optimization variable is the sum power per user. This optimization problem can be solved with numerical methods in the four cases: a simple iterative process is used for Max-Min Fair allocation, while Newton's method is used with the three other objectives. The details of both methods are provided in Appendices 3.A and 3.B.

Then, $P_{d,k}$ and $P_{c,k}$ are distributed over the subcarriers of user k with the following objective (represented here for the direct link, the coordinated link's optimization is similar):

$$\begin{aligned}
&\max_{\mathbf{P}_{d,k}} \sum_{l\in\Theta_k} \log_2\left(\left(1+\frac{G^l_{d,k}P^l_{d,k}}{I^l_k}\right)^2+\frac{G^l_{c,k}P^l_{c,k}}{I^l_k}\right)\\
&\text{s.t. } \sum_{l\in\Theta_k} P^l_{d,k}=P_{d,k}\\
&\text{s.t. } P^l_{d,k}\geq 0,\ \forall l\in\Theta_k
\end{aligned} \tag{3.14}$$

(3.14) is a Globally Optimal optimization problem, that can be solved with Newton's method (see Appendix 3.B).

3.4.3 Single carrier case: analytical study

In this section, we consider the single carrier case with Globally Optimal power allocation objective, in order to get some insight into the proposed joint coordination - power control method. The single carrier case can be analytically solved for coordinated users. An analytical solution can also be obtained for direct users, if all direct users are also coordinated by neighboring BSs.

Power allocation for coordinated users

The optimization problem (3.9) is convex in $\mathbf{P}_c$. It can be solved with the KKT conditions. Details are given in Appendix 3.C. The solution is:

$$P_{c,k}=\left[\frac{1}{\mu_c}-\frac{I_k}{G_{c,k}}\left(1+\frac{G_{d,k}P_{d,k}}{I_k}\right)^2\right]^+ \tag{3.15}$$

where the constant μ_c must be chosen so that the power constraint $\sum_{k\in\mathcal{S}_c} P_{c,k} = P_{\max,c}$ is fulfilled.

This solution is consequently a modified water-filling, where the SINR from the direct link, $\frac{G_{d,k}P_{d,k}}{I_k}$, must be considered. From this analytical expression, we can deduce that the power on the coordinated link for user k, $P_{c,k}$, will decrease when the SINR on the direct link increases. This indicates that the coordinated link aims at compensating for the direct link, thus setting more power when the SINR on the direct link is too low, and on the contrary, setting less power when the SINR on the direct link is high enough. These conclusions stand for a fixed μ_c value, whenever $P_{c,k}$ is different from zero, which depends both on the direct link SINR, and on the SINR of the other coordinated users, that are considered through μ_c setting.

Power allocation for direct users

If a direct user is coordinated, then its power has an influence in $(P_{d,k})^2$ in the data rate; whereas for non-coordinated users, the power has an influence in $P_{d,k}$. Consequently, if both types of users are present, we cannot find a simple analytical expression for the optimization problem's solution. It is then necessary to use Newton's method (see Appendix 3.B).
In the case where all the direct users are coordinated by a neighboring BS, the solution can be obtained analytically with the KKT conditions :

$$P_{d,k} = \left[\frac{1}{\mu_d} + \frac{1}{\mu_d}\sqrt{a_k} - \frac{I_k}{G_{d,k}}\right]^+ \tag{3.16}$$

where

$$a_k = \left[1 - \frac{\frac{G_{c,k}P_{c,k}}{I_k}}{(\frac{G_{d,k}}{I_k})^2}\mu_d^2\right]^+$$

and μ_d is a constant parameter that must be chosen in order to fulfill the power constraint $\sum_{k\in\mathcal{S}_d} P_{d,k} = P_{\max,d}$.
The power on the coordinated link is only considered if condition $\left(1 - \frac{\frac{G_{c,k}P_{c,k}}{I_k}}{(\frac{G_{d,k}}{I_k})^2}\mu_d^2\right) > 0$ is fulfilled. In that case, the power on the direct link $P_{d,k}$ decreases when the SINR on the coordinated link increases. However, $\sqrt{P_{c,k}}$ may be considered within $P_{d,k}$, whereas $(P_{d,k})^2$ is considered in $P_{c,k}$, thus emphasizing that both links asymmetrically contribute to the data rate.

3.5 Numerical results

Network $\mathcal{N}$ is modeled by two rings of interfering BSs with omnidirectional antennas and same cell radius (see Fig. 3.5). The different power allocation objectives are compared by using Monte-Carlo simulations, with $K_d = \{32, 64, 96, .., 224\}$ direct users for each BS of $\mathcal{N}$. The total bandwidth is $B = 10$ MHz and the FFT size is $N_{\text{FFT}} = 256$. The number of available subcarriers per BS is $L_{\text{SC}} = 256$. The inter-site distance is $\text{d}_{\text{is}} = 0.7\sqrt{3} = 1.212$ km. The path loss model is Okumura-Hata [4]: $pl(\text{d}) = 137.74 + 35.22\log(\text{d})$ in dB. The shadowing follows a log-normal law with standard deviation 7 dB, and the fast fading is Rayleigh. The noise is equal to -105 dBm. It contains the thermal noise and non-managed interference coming from non-modeled cells. The maximum transmit power for each BS is $P_{\max} = 43$ dBm. The coordination triggering parameter Δ is set to 3 dB. This value restricts network coordination to users at cell's border. It is also consistent with the handover triggering value usually chosen for this type of deployment in cellular networks.

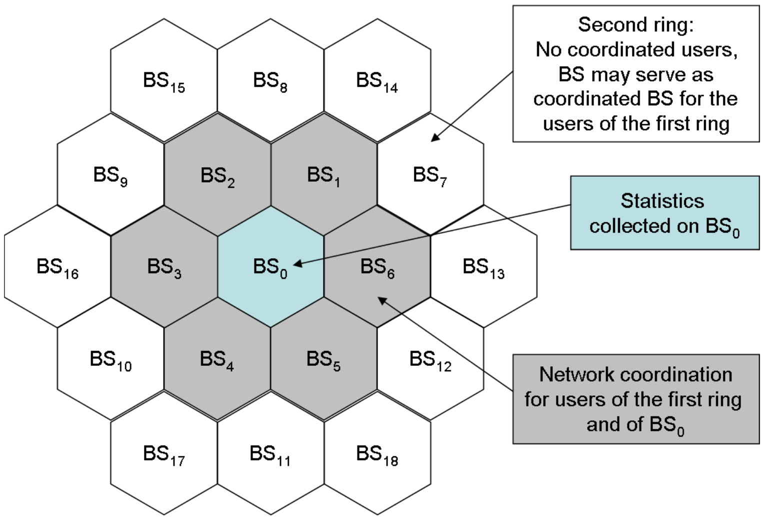


Figure 3.5: Network model

Only the users of the 7 central BSs (BS_0 to BS_6) may be coordinated. The second ring of BSs generates inter-cell interference and helps in coordinating users of the 7 central BSs. Statistics are averaged over BS_0, in order to avoid side effects. In the figures and tables, we refer to Globally Optimal allocation as 'GO', Proportional Fairness as 'PF', Harmonic Mean Fairness as 'HMF', and Max-Min Fairness as 'MM'. 'C' indicates that network coordination is used, whereas 'NC' indicates that network coordination is not used.

3.5.1 Numerical convergence study

Each power allocation problem corresponding to one BS and one user type (direct or coordinated) is locally convex and leads to an optimal value. However, this does not imply that the global problem over the two rings of BSs actually converges. In order to evaluate the relevance of our method, we have performed a numerical convergence study on direct and coordinated power for users of BS_0. The convergence statistics for $K_d = 128$, with a convergence criterion of 5%, are gathered in Table 3.1.

Convergence is faster at high load. Indeed, at low to medium load, each user has several allocated subcarriers, and power control is performed in two steps. The second step allocates power to subcarriers with Globally Optimal objective, which leads to some subcarriers' power set to zero if the channel coefficients have a high standard deviation. This explains for users' power variations at low to medium load. This phenomenon is amplified with fair power allocation objectives, as in these cases, users at cell's border have high power values. The power variations over the subcarriers of these users are consequently higher. However, at high load (when $K_d \geq 128$), the convergence rate for direct and coordinated powers exceeds 99% with all objectives.

Finally, the average number of iterations required to achieve convergence for all users also increases with the fairness of the power allocation objective. 6 iterations are required with Globally Optimal objective, whereas up to 19 iterations are required with Max-Min Fair objective. Therefore, the algorithm's convergence time remains quite reasonable, and iterative power control may be run within the time slots composing the two considered TTIs.

Table 3.1: Convergence of power control for $K_d = 128$

BS Power	For direct users	For coordinated users
Convergence Rate (%)		
GO	99.09	99.21
PF	99.54	99.86
HMF	99.85	99.86
MM	99.31	99.29
Required number of iterations		
GO	5	6
PF	5	6
HMF	7	8
MM	18	19

3.5.2 Throughput comparison

Network coordination increases the cell's sum throughput at any load with all power allocation objectives except Globally Optimal, and the gain gets higher when the load increases. On the contrary, network coordination decreases the sum throughput in all cases with Globally Optimal, but the loss gets lower when the load increases (see Fig. 3.6). With Globally Optimal objective, the throughput loss is between 25 and 9%. With Proportional Fair, the throughput gain is around 5%. With Harmonic Mean Fair, the minimum throughput gain is 11%, and a maximum gain of 36% is achieved for $K_d = 224$. The highest gains are obtained with Max-Min Fair, as the sum throughput gain is between 28% and 85%. The fairer the power allocation objective, the higher network coordination gain is. Indeed, network coordination increases the data rate of users at cell's border. With Max-Min Fair objective, the data rate of all users is increased, as it is aligned with the lowest data rate. Harmonic Mean Fair and Proportional Fair objectives follow the same tendency. With Globally Optimal objective however, increasing the data rate at cell's border implies decreasing the data rate of users that are close to the cell, and as this power allocation objective importantly favors users in good radio conditions at the expense of users at the border of cells, the sum throughput is negatively impacted.

3.5.3 Fairness comparison

In order to study the influence of network coordination depending on user's locations, BS_0 is divided into 10 rings with equal area around the BS. The same number of users are served within each ring.

Fig. 3.7 and 3.8 represent the average data rate within each ring, for a medium load of $K_d = 128$. We can first notice that, with Max-Min Fair, Harmonic Mean Fair and Proportional Fair objectives, network coordination brings data rates gain in all the rings.

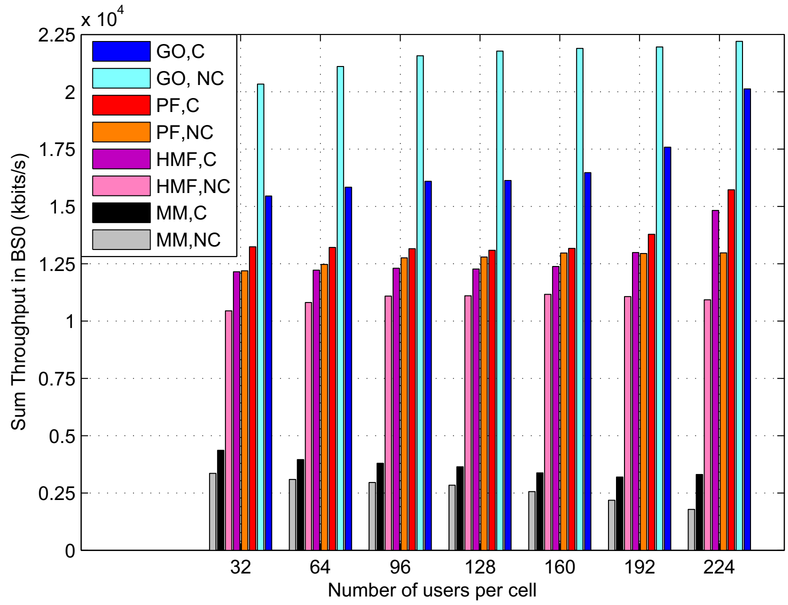


Figure 3.6: Sum throughput, depending on the load and on the power allocation objective

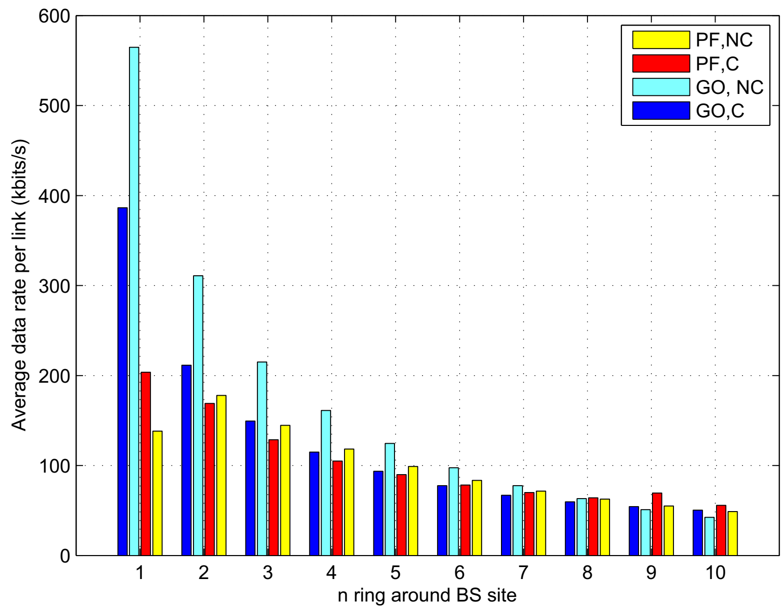


Figure 3.7: Average users' data rate per ring for $K_d = 128$, GO and PF power allocation objectives

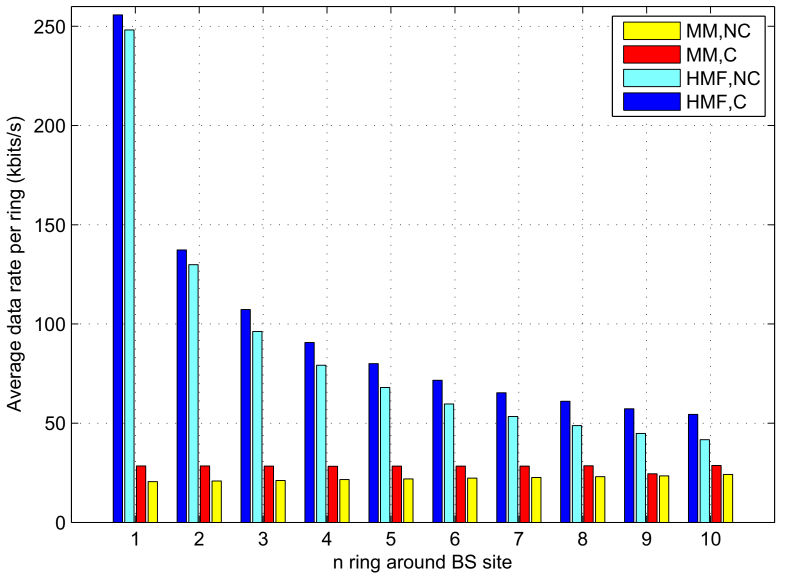


Figure 3.8: Average users' data rate per ring for $K_d = 128$, HMF and MM power allocation objectives

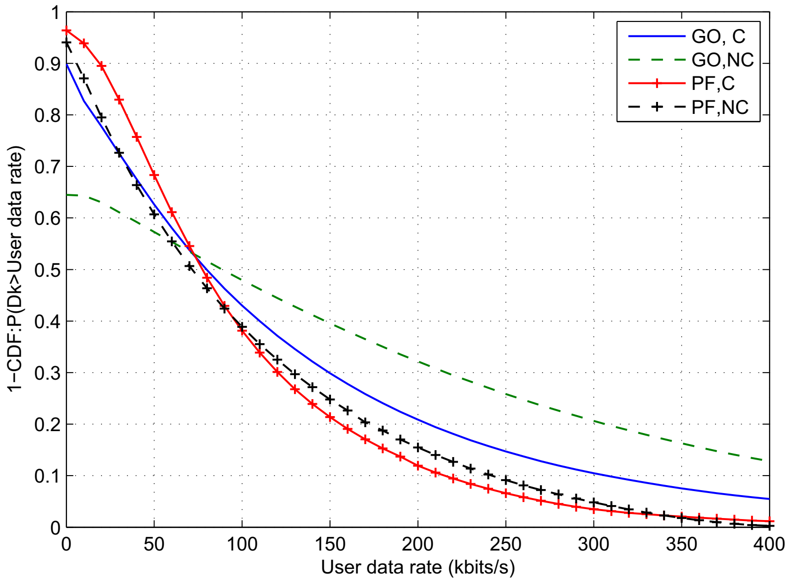


Figure 3.9: Distribution of users' data rates for $K_d = 128$, GO and PF power allocation objectives

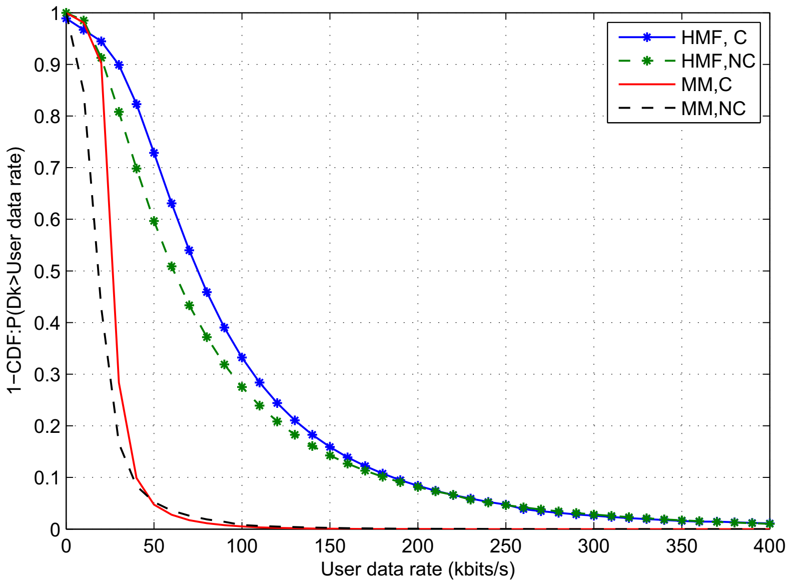


Figure 3.10: Distribution of users' data rates for $K_d = 128$, HMF and MM power allocation objectives

With Globally Optimal objective, only the last two rings benefit from coordination, and the data rates loss is high in the first ring. Consequently, the fairness gain for users at the border of the cell is obtained at the expense of a peak data rate decrease equal to 31% with this objective.

Fig. 3.9 and Fig. 3.10 represent the influence of relaying on users' data rates' distribution, for $K_d = 128$. Fairness is improved with all power allocation objectives. The probability of exceeding a given data rate is higher with network coordination for all data rates with Max-Min Fair and Harmonic Mean Fair, and for data rates up to 94 kbits/s with Proportional Fair. Although these data rates may seem low, it should be noted that the data rate increase concerns 59% of the users with Proportional Fair. With Globally Optimal objective, network coordination decreases the ratio of unserved users from 35% to 10%.

3.6 Conclusion

This chapter has introduced a resource allocation strategy for causal network coordination. The influence of network coordination on four power allocation objectives, Globally Optimal, Max-Min Fair, Proportional Fair and Harmonic Mean Fair, has been tested. Simulation results show that network coordination brings additional fairness at any load with all power allocation strategies, and increases the system throughput with fair power allocation objectives, whereas it decreases the sum throughput with the unfair Globally Optimal objective. Indeed, with this objective that favors users in good radio conditions, additional fairness is obtained for users at cell's border at the expense of a decrease of the peak data rate. On the contrary, with fair power allocation objectives, all users benefit from the data rate gain provided to the users at cell's border.

In this chapter, we have assumed that all users were subject to the same resource allocation objective. However, this conjecture no longer holds when considering more realistic wireless networks, where users may be requesting various services. As seen in Section 2.1.2, the users' QoS characteristics regarding per TTI resource allocation can be mapped onto two categories: RC users, whose objective is to reach a minimum target data rate, and BE users, that do not have any individual QoS objective. The first category of users follows a fair resource allocation objective, as the QoS fulfillment of each user is compelled by the achievement of a given quality indicator (namely, its target data rate), whereas the second one follows an unfair resource allocation objective, which is the maximization of the sum data rates, also referred to as Globally Optimal. The results obtained in this chapter have shown that network coordination should be used for the users whose QoS constraints correspond to a fair resource allocation objective, but not for the users following a Globally Optimal resource allocation objective. Consequently, in Chapter 4, we will study how to efficiently use network coordination in order to differentiate users per QoS constraint.

APPENDIX

3.A Max-min fair power allocation

To simplify the notations, we set: $a_k = \frac{\tilde{G}_{c,k}}{I_k} P_{c,k}$ and $b_k = \frac{\tilde{G}_{d,k}}{I_k}$. Let us consider Max-Min Fair power allocation on direct users:

$$\begin{aligned}
&\max_{\mathbf{P}_d} \min_{k \in \mathcal{S}_d} \log_2 \left((1 + b_k P_{d,k})^2 + a_k \right) \\
&\text{s.t.} \sum_{k \in \mathcal{S}_d} P_{d,k} = P_{\max,d} \\
&\text{s.t. } P^l_{d,k} \geq 0, \ \forall k \in \mathcal{S}_d, \ l \in \Theta_k
\end{aligned} \tag{3.17}$$

The Max-Min Fair routine aims at providing the same data rate to all users, which is equivalent to setting the same capacity, as all users are allocated the same number of subcarriers: $\mathrm{C} = \log_2(\Delta_{\min})$. The power allocation method is consequently:

1. Compute $\Delta_{\min}$, solution of

$$P_{\max,d} + \sum_{\{k \in \mathcal{S}_d | P_{d,k} \neq 0\}} \frac{1}{b_k} = \sum_{\{k \in \mathcal{S}_d | P_{d,k} \neq 0\}} \frac{\sqrt{\Delta_{\min} - a_k}}{b_k}$$

   For each k,

   (a) if $\Delta_{\min} - a_k > 0$ and $\sqrt{\Delta_{\min} - a_k} - 1 \geq 0$ set $P_{d,k} = \frac{1}{b_k}\left[\sqrt{\Delta_{\min} - a_k} - 1\right]$.

   (b) Else, set $P_{d,k} = 0$.

2. Iterate until $\sum_{k \in \mathcal{S}_d} P_{d,k} = P_{\max,d}$.

3.B Newton's method

The general convex problem for power allocation is:

$$\begin{aligned}
&\min_{\mathbf{P}} \sum f\left(\log(g(\mathbf{P}))\right) \\
&\text{s. t. } \mathbf{1}^T \mathbf{P} = P_{\max} \\
&\text{s. t. } \mathbf{P} \geq \mathbf{0}
\end{aligned} \tag{3.18}$$

Where $g(\mathbf{P})_k = \log_2\left(\left(1 + \frac{\tilde{G}_{d,k} P_{d,k}}{I_k}\right)^2 + \frac{\tilde{G}_{c,k} P_{c,k}}{I_k}\right)$, and f is a convex function in $\mathbf{P}_d$ (resp. $\mathbf{P}_c$ if we consider the coordinated link). $P_{\max} = P_{\max,d}$ for direct users, and $P_{\max} = P_{\max,c}$ for coordinated users. With Globally Optimal objective, $(f(\mathbf{x}))_k = -(\mathbf{x})_k$; with Proportional Fair objective, $(f(\mathbf{x}))_k = -(\log(\mathbf{x}))_k$ and with Harmonic Mean Fair objective, $(f(\mathbf{x}))_k = \left(\frac{1}{\mathbf{x}}\right)_k$, $\forall k \in \mathcal{S}_d$ for direct users, and $\forall k \in \mathcal{S}_c$ for coordinated users.
This problem is a convex equality-constrained problem that can be solved with Newton's method [24]. Newton's method is simplified in our case as the Hessian matrix is diagonal. The Newton step $\mathbf{\Delta P_{nt}}$ is defined by the KKT system:

$$\begin{pmatrix} \mathbf{B} & \mathbf{A}^T \\ \mathbf{A} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{\Delta P_{nt}} \\ \mathbf{w} \end{pmatrix} = \begin{pmatrix} -\mathbf{c} \\ \mathbf{0} \end{pmatrix} \tag{3.19}$$

where $\mathbf{A} = \mathbf{1}^T$, $\mathbf{c} = \nabla f(\mathbf{P})$, and $\mathbf{B}$ is a diagonal matrix with elements: $(\mathbf{B})_{k,k} = \nabla^2(f(\mathbf{P}))_k$. The KKT system can be solved by elimination, *i.e.* by solving $\mathbf{A}\mathbf{B}^{-1}\mathbf{A}^T\mathbf{w} = -\mathbf{A}\mathbf{B}^{-1}\mathbf{c}$, and setting $\mathbf{\Delta P_{nt}} = -\mathbf{B}^{-1}(\mathbf{A}^T\mathbf{w} + \mathbf{c})$.
As $\mathbf{A} = \mathbf{1}$, we directly obtain:

$$\mathbf{w} = -\frac{\sum_{k\in\mathcal{S}}((\nabla^2 f(\mathbf{P}))_k)^{-1} \times \nabla(f(\mathbf{P}))_k}{\sum_{k\in\mathcal{S}}(\nabla^2(f(\mathbf{P}))_k)^{-1}} \tag{3.20}$$

where $\mathcal{S} = \mathcal{S}_d$ for direct users, and $\mathcal{S} = \mathcal{S}_c$ for coordinated users.
and $\mathbf{\Delta P_{nt}}$ is deduced from:

$$\mathbf{\Delta P_{nt}} = -\text{diag}\ ((\nabla^2(f(\mathbf{P}))_k)^{-1})(\mathbf{A}^T\mathbf{w} + \nabla f(\mathbf{P})) \tag{3.21}$$

Finally, the algorithm to solve the power allocation problem is:

1. Compute the gradient and Hessian of $f(\mathbf{P})$.
2. Compute $\mathbf{w}$ from equation (3.20).
3. Compute $\mathbf{\Delta P_{nt}}$ from (3.21).
4. Compute $\lambda(\mathbf{P})^2 = -\bigtriangledown f(\mathbf{P})^T\mathbf{\Delta P_{nt}}$. Quit if $\lambda(\mathbf{P})^2/2 < \epsilon$, where ϵ is a parameter evaluating the convergence.
5. Compute t with backtracking line search.
6. Update $\mathbf{P} := \mathbf{P} + t\mathbf{\Delta P_{nt}}$ and iterate.

3.C Single carrier case, power allocation for coordinated users

The Lagrangian corresponding to problem (3.9), normalized by $\frac{2\log(2)}{B_{\text{SC}}}$ is:

$$\begin{aligned} L(\mathbf{P_c}, \boldsymbol{\lambda}, \mu_c) = &\sum_{k\in\mathcal{S}_c} \log\left(\left(1 + \frac{G_{d,k}P_{d,k}}{I_k}\right)^2 + \frac{G_{c,k}P_{c,k}}{I_k}\right) \\ &+ \sum_{k\in\mathcal{S}_c} \lambda_k P_{c,k} - \mu_c\left(\sum_{k\in\mathcal{S}_c} P_{c,k} - P_{\max,c}\right) \end{aligned} \tag{3.22}$$

The derivative of the Lagrangian with regard to $P_{c,k}$ is given by:

$$\frac{\partial L(\mathbf{P_c}, \boldsymbol{\lambda}, \mu_c)}{\partial P_{c,k}} = \frac{\frac{G_{c,k}}{I_k}}{\left(1 + \frac{G_{d,k}P_{d,k}}{I_k}\right)^2 + \frac{G_{c,k}P_{c,k}}{I_k}} + \lambda_k - \mu_c$$

The KKT conditions impose that: $\frac{\partial L(\mathbf{P_c}, \boldsymbol{\lambda}, \mu_c)}{\partial P_{c,k}} = 0$ and $\lambda_k \geq 0$.
$\lambda_k \geq 0$ condition corresponds to :

$$\mu_c \geq \frac{\frac{G_{c,k}}{I_k}}{\left(1 + \frac{G_{d,k}P_{d,k}}{I_k}\right)^2 + \frac{G_{c,k}P_{c,k}}{I_k}}$$

Another KKT condition is that $\lambda_k P_{c,k} = 0$, i.e.,

$$\left(\mu_c - \frac{\frac{G_{c,k}}{I_k}}{\left(1 + \frac{G_{d,k}P_{d,k}}{I_k}\right)^2 + \frac{G_{c,k}P_{c,k}}{I_k}} \right) P_{c,k} = 0$$

If $\mu_c > \frac{\frac{G_{c,k}}{I_k}}{\left(1+\frac{G_{d,k}P_{d,k}}{I_k}\right)^2 + \frac{G_{c,k}P_{c,k}}{I_k}}$ then this condition can only be fulfilled if $P_{c,k} = 0$. Else, $\mu_c = \frac{\frac{G_{c,k}}{I_k}}{\left(1+\frac{G_{d,k}P_{d,k}}{I_k}\right)^2 + \frac{G_{c,k}P_{c,k}}{I_k}}$ Finally, we obtain :

$$P_{c,k} = \left[\frac{1}{\mu_c} - \frac{I_k}{G_{c,k}} \left(1 + \frac{G_{d,k}P_{d,k}}{I_k}\right)^2 \right]^+ \tag{3.23}$$

where $[x]^+ = \max\{0, x\}$. The constant μ_c must be chosen so that the power constraint $\sum_{k \in \mathcal{S}_c} P_{c,k} = P_{\max,c}$ is fulfilled.

Chapter 4

QoS aware resource allocation with distributed network coordination

4.1 Introduction

Power control with network coordination has been studied in the previous chapter with a global optimization objective for all users. Four power allocation objectives have been tested with the same subcarrier allocation method: Globally Optimal, Max-Min Fair, Proportional Fair and Harmonic Mean Fair. In Chapter 3, network coordination was triggered for all the users located at the border of cells. Numerical results have shown that network coordination increases fairness among users. Thus, high throughput gains are obtained with fair power allocation objectives, whereas the peak data rate decreases when the objective is unfair. As a consequence, network coordination is useful for users requesting a minimum data rate, but should not be triggered for BE users, whose per TTI QoS performance criterion is the cell's sum throughput.

Based on the results from Chapter 3, we propose a resource allocation method that uses network coordination to provide the required QoS of RC users, that are subject to target data rate requirements, and to increase the sum data rate of BE users. Subcarrier allocation as well as power control are used to optimize resource allocation. The proposed method arranges users for resource access according to priority. The priority order depends on QoS constraints, and favors RC users. RC users should consequently get their required data rate, but at the expense of the least possible radio resources, so that the remaining resources can be allocated to BE users. The proposed method is suited to distributed networks, as it requires small information exchange between the coordinated BSs, and takes into account the causality constraint.

The chapter first details the literature relevant to resource allocation in OFDMA. In Chapter 3, subcarrier allocation was not optimized, and the main focus was on power control. In this chapter, both subcarrier and power allocations need to be optimized so as to use radio resources as efficiently as possible. Therefore, we review the different methods proposed in the literature for joint resource allocation in OFDMA. Aside from the case of Globally Optimal objective in single-cell OFDMA, none of the resource allocation problems can be optimally solved by a decomposition into two sequential stages (subcarrier allocation followed by power allocation). We detail the current status of joint resource allocation for RC users in single-cell OFDMA. We then study some relevant separate methods proposed to solve the corresponding optimization problems. Finally, a review of the proposed techniques for resource allocation in multi-cell OFDMA is performed. It shows that up to now, no efficient solution has been determined for joint resource allocation in multi-cell OFDMA networks. Our focus in this thesis will be on separate resource allocation methods, due to the distributed network's constraint.

The chapter then describes a QoS-aware resource allocation method. In the considered scenario, two sets of users with different QoS requirements are present, RC and BE users. Our resource allocation method uses the network coordination procedure investigated in Chapter 3 to serve the RC users located at the border of cells. For these users, resource allocation corresponds to a Margin Adaptive (MA) objective: the aim is to minimize the sum power required to fulfill their data rate requirement. Then BE users are served following a Globally Optimal objective. The chapter details the prioritization among the set of users regarding subcarrier and power allocations. The allocation methods for the two sets of users are then described. Finally, the QoS-aware resource allocation method is compared with Equal Power Allocation with the same subcarrier allocation, and with a modified version of the method proposed in [51].

The main contributions of this chapter are:

- Network coordination is efficiently used, by limiting its triggering to the cases when a fairness gain is required. Only RC users located at the border of cells are coordinated. This leads to a decrease in the sum power required to reach the target data rates of these users, which is beneficial to the whole network.

- The global resource allocation problem per cell is decomposed thanks to a prioritization among RC and BE users. Resource allocation is first performed for RC users, and then for BE users. This decomposition of the original problem greatly simplifies it.

- Subcarrier allocation for RC users with network coordination is proposed as an extension of the method from [52]. The analytical solution of the Margin Adaptive resource allocation problem involving network coordination is derived, by using the KKT conditions.

- The performance results of both QoS sets are highly increased compared to the existing methods from the literature, both with and without network coordination. Consequently, the proposed prioritization and resource allocation methods could be used even if network coordination is not available. Network coordination brings additional performance gains, both in terms of ratio of RC users achieving their target data rate, and in terms of the sum throughput of BE users.

4.2 State of the art on resource allocation in OFDMA

In this section, we detail the state of the art on downlink resource allocation in OFDMA. Most articles consider a single cell without inter-cell interference. This comes from the fact that early works on resource allocation in OFDMA stem from Digital Subscriber Line systems (xDSL) studies. Consequently, we first review single-cell cases, and then the relevant literature on multi-cell networks. In the following optimization problems, the optimization variables are the subcarrier allocation set and the power values.

4.2.1 Optimization problems

Best Effort users: Globally Optimal objective

The Globally Optimal objective, maximization of the sum data rate under a sum power constraint $P_{\max}$, is studied in [53]. The authors consider a single cell with L_{SC} subcarriers, containing K users. They first allow multiple transmission of several users on the same subcarrier. The assignment index c_k^l is a real continuous value in $[0, 1]$. It is equal to 0 if subcarrier l is not allocated to user k. R_k^l is the data rate of user k in subcarrier l, and P_k^l is the power transmitted to user k in subcarrier l.

$$\begin{aligned}
&\max_{\{\mathbf{P},\mathbf{c}\}} \sum_{k=1}^{K}\sum_{l=1}^{L_{\mathrm{SC}}} R_k^l \\
&\text{s.t. } \sum_{k=1}^{K}\sum_{l=1}^{L_{\mathrm{SC}}} c_k^l P_k^l \leq P_{\max} \\
&\text{s.t. } P_k^l \geq 0, \forall (k,l) \in \{1, ..., K\} \times \{1, ..., L_{\mathrm{SC}}\}
\end{aligned} \tag{4.1}$$

The authors analytically prove that the sum data rate is maximized when each subcarrier is assigned to one user only, the user with the best channel gain for that subcarrier, and when the transmit power is distributed over the subcarriers through water-filling. Consequently, for BE users, OFDMA is optimal, i.e. c_k^l is an integer in $\{0, 1\}$. Besides, the solution of the joint optimization problem over subcarrier and power allocations is equal to the solution of the separated optimization problems, where subcarrier allocation is first performed, and then followed by power allocation.

Rate Constrained users: Rate Adaptive and Margin Adaptive problems

Two optimization problems can be used to study resource allocation for RC users: Rate Adaptive (RA) and Margin Adaptive (MA). In OFDMA, the assignment index c_k^l is equal to 1 if subcarrier l is allocated to user k by its serving BS, and to 0 otherwise. The set of subcarriers allocated to user k, $\Theta_k = \left\{l \mid c_k^l = 1\right\}$, is consequently orthogonal to the sets of subcarriers allocated to the other users, and it is equivalent to consider the set $\mathbf{\Theta}$ instead of $\mathbf{c}$ as subcarrier allocation's optimization variable. The sum data rate for user k is $R_k = \sum_{l\in\Theta_k} R_k^l$, and the sum power allocated to user k is $P_k = \sum_{l\in\Theta_k} P_k^l$.
The RA optimization problem in downlink OFDMA has been introduced for the single-cell case in [52]. Its objective is to maximize the minimum data rate over all K users, under a

maximum sum power constraint:

$$
\begin{aligned}
&\max_{\{\mathbf{P},\boldsymbol{\Theta}\}} \min_{k=\{1,\ldots,K\}} \sum_{l\in\Theta_k} R_k^l \\
&\text{s.t. } \sum_{k=1}^{K}\sum_{l\in\Theta_k} P_k^l \leq P_{\max} \\
&\text{s.t. } P_k^l \geq 0, \forall (k,l) \in \{1,...,K\} \times \{1,...,L_{\mathrm{SC}}\} \\
&\text{s.t. } \Theta_k \cap \Theta_{k'} = \emptyset, \forall (k,k') \in \{1,...,K\}^2, k \neq k'
\end{aligned}
\tag{4.2}
$$

The MA allocation problem [54] aims at minimizing the sum power under a set of users' data rates constraints:

$$
\begin{aligned}
&\min_{\{\mathbf{P},\boldsymbol{\Theta}\}} \sum_{k=1}^{K}\sum_{l\in\Theta_k} P_k^l \\
&\text{s.t. } \sum_{l\in\Theta_k} R_k^l \geq R_{k,\mathrm{target}}, \forall k \in \{1,...,K\} \\
&\text{s.t. } P_k^l \geq 0, \forall (k,l) \in \{1,...,K\} \times \{1,...,L_{\mathrm{SC}}\} \\
&\text{s.t. } \Theta_k \cap \Theta_{k'} = \emptyset, \forall (k,k') \in \{1,...,K\}^2, k \neq k'
\end{aligned}
\tag{4.3}
$$

In [55], it is shown that if all users have the same target data rate $R_{k,\mathrm{target}} = R_{\mathrm{target}}, \forall\, k$, the RA problem can be viewed as an iterative MA problem where the target data rate increases until the total power constraint no longer holds.

4.2.2 Joint resource allocation for RC users

In OFDMA, MA and RA resource allocation problems are not convex because subcarrier allocation is a discrete problem, as only one user can be assigned per subcarrier. They can be turned into convex optimization problems by relaxing the discrete constraint in subcarrier allocation [52] [54], i.e., allowing $c_k^l \in [0,1]$. However, in both cases, the obtained solution may lead to some subcarriers being shared between several users. Consequently, an additional process is required to retrieve an OFDMA solution, which may therefore be sub-optimal. Besides, as stated in [56], the methods proposed in [54] and [52] involve higher computational load than solving the MA and RA via Integer Programming. However, even with Integer Programming, the complexity increases exponentially in the number of constraints and variables. Consequently, most papers propose sub-optimal heuristics to reduce complexity [56] [57].

Recently, it has been shown by Seong et al. in [58] that the Lagrange dual decomposition method can solve the joint resource allocation problem in OFDMA, provided that the channel consists of a minimum number of subcarriers. Seong et al. use an important result from [59] on non-convex optimization. In [59], the authors study the problem of maximization of the sum data rate under a sum power constraint. They show that under a certain condition called the time-sharing condition, the duality gap of the optimization problem is always zero, regardless of the convexity of the objective function. The time-sharing condition is satisfied for practical multi-user spectrum optimization problems in multi-carrier systems in the limit, as the number of subcarriers goes to infinity. Consequently, it is possible to use efficient numerical algorithms to solve non-convex problems in the dual domain.

These results are applicable to xDSL and to OFDM systems. Seong et al. [58] study the Weighted Sum Rate Maximization problem (WSRMax), that corresponds to Globally Optimal objective with a weight per user, and the Weighted Sum Power Minimization problem (WSPMin), that is equivalent to MA optimization problems with a given weight per user, for single-cell OFDMA. They show that with a minimum number of $L_{\mathrm{SC}} = 8$ subcarriers and with $K = 2$ users, the duality gap is virtually equal to zero.

4.2.3 Separate subcarrier and power allocations for RC users

For the RA problem, Rhee and Cioffi [52] perform subcarrier allocation, assuming that equal amount of power is allocated to each subcarrier, with an iterative process. After an initialization phase during which each user is allocated its best subcarrier (with regard to the channel coefficient), in each iterative step, the user with the lowest data rate is identified, and is assigned its best subcarrier, among the list of subcarriers that are still available. The iterative process stops when all subcarriers are allocated.

In [60], the authors add power control in sequence after the subcarrier allocation method from [52]. Power allocation follows a RA objective. The considered RA problem may impose proportionality constraints among the data rates of users: $R_1 : R_2 : ... : R_k = \alpha_1 : \alpha_2 : ... : \alpha_k$, where $\{\alpha_k\}_{k=1}^{K}$ is a set of positive coefficients. Power allocation is solved with Newton's method when $\alpha_k = 1, \forall k$. Numerical results show that the proposed power allocation method leads to higher capacity than the equal power allocation of [52]. The capacity gain increases with the number of users, thanks to multi-user diversity.

In [51], the method of [60] is adapted in order to serve two types of users with different QoS requirements: RC users and BE users. The optimization problem aims at maximizing the minimum data rate for RC users, while maximizing the sum data rate for BE users. Subcarrier allocation is first performed for RC users with the method from [52], and the remaining subcarriers are then allocated to BE users in a greedy manner. The RA problem is solved for power allocation of RC users, under the constraint that the sum power should be lower than P_{RC}, which is proportional to the number of subcarriers allocated to RC users. For BE users, power allocation aims at maximizing the sum data rate under the sum power constraint determined by the remaining power, $P_{\max} - P_{\mathrm{RC}}$. The main limitation of this method is that the sum power dedicated to RC users, P_{RC}, is not determined according to any optimization rule.

4.2.4 Resource allocation in multi-cell networks

To the best of our knowledge, joint subcarrier and power allocation has not been treated for the multi-cell case yet. The interesting results from [58] are only applicable to the single-cell case.

Resource allocation in multi-cell can be performed independently in each BS, considering inter-cell interference as noise. This method is called iterative water-filling [61]. If inter-cell interference cannot be neglected, then iterative water-filling may lead to power divergence situations. In [62] [63], the MA optimization problem is studied under the scope of game theory. Iterative water-filling can be viewed as a non-cooperative game where each player (i.e., each BS) aims at minimizing the sum power. The authors in [62] suggest to avoid power divergence situations by including a centralized mediator, that prevents some users from transmitting in the subcarriers where their SINR is too low. Other heuristics have been proposed in [64] [65]. In order to account for inter-cell interference, these heuristics are either centralized procedures [64] or distributed procedures requiring an iterative process [65]. This dissertation focuses on distributed multi-cell resource allocation. Based on the conclusions that, even in the single-cell case, separate resource allocation leads to satisfying performance results with reasonable complexity compared to joint resource allocation methods, and that joint resource allocation in the multi-cell case remains an open problem, we propose to separate subcarrier allocation from power control.

4.3 QoS-aware resource allocation

In this section, a QoS-aware resource allocation method for multi-cell OFDMA with both RC and BE users is determined. Unlike [51], the optimization problem for RC users is MA. In order to provide the required QoS to all RC users whatever their location, we use network coordination for the RC users located at the border of cells. The causal network coordination method has been described in Chapter 3. A major conclusion of Chapter 3 is that the sum throughput is not improved by network coordination when the power allocation objective is Globally Optimal. As a consequence, BE users are not coordinated.

4.3.1 Description of the proposed method

The resource allocation problem in each BS aims at providing their target data rate to RC users, while maximizing the sum data rate of BE users. The sum data rate for BE users is an increasing function of the power dedicated to BE users. Consequently, the optimization goal is equivalent to minimizing the sum power required to satisfy all RC users, so that the power dedicated to BE users is maximized. The resource allocation problem on each BS n_{BS} is:

$$
\begin{aligned}
&\min_{\{\mathbf{P},\boldsymbol{\Theta}\}} \sum_{k\in\mathcal{S}_{\mathrm{RC},d}} \sum_{l\in\Theta_k} \left(P_{d,k}^l + P_{c,k}^l\right) \\
&\text{s.t. } \sum_{l\in\Theta_k} R_k^l \geq R_{k,\mathrm{target}},\ \forall k \in \mathcal{S}_{\mathrm{RC},d} \\
&\max \sum_{k\in\mathcal{S}_{\mathrm{BE}}} R_k \\
&\text{s.t. } \sum_{l=1}^{L_{\mathrm{SC}}} P_{n_{\mathrm{BS}}}^l \leq P_{\max} \\
&\text{s.t. } P_{n_{\mathrm{BS}}}^l \geq 0, \forall l \in \{1, ..., L_{\mathrm{SC}}\} \\
&\text{s.t. } \Theta_k \cap \Theta_{k'} = \emptyset, \forall (k,k') \in \{\mathcal{S}_{\mathrm{RC},d} \cup \mathcal{S}_{\mathrm{BE}}\}^2, k \neq k'
\end{aligned}
\tag{4.4}
$$

$\mathcal{S}_{\mathrm{RC},d}$ is the set of direct RC users served by BS n_{BS}, $\mathcal{S}_{\mathrm{RC},c}$ is the set of RC users coordinated by n_{BS}, and $\mathcal{S}_{\mathrm{BE}}$ is the set of BE users served by n_{BS}. The same notations are used in this chapter as in Chapter 3 to characterize network coordination.

Network coordination is restricted to the RC users located at the border of cells. The direct BS requests coordination to a neighboring BS for RC users that have a path loss difference of less than Δ dB between their direct BS and the best neighboring BS.

As justified in Section 4.2.4, subcarrier allocation is separated from power control to reduce complexity. Subcarrier allocation for RC users is performed with an iterative algorithm that aims at minimizing the number of subcarriers required to satisfy the target data rate requirements of all RC users. Assuming equal power and interference levels on all subcarriers, the proposed algorithm consequently minimizes the sum power. The RC users are grouped within groups with the same target data rate. These groups are ordered by increasing target data rate, so that the users with the lowest target data rate are more likely to be satisfied. The subcarrier allocation process (detailed in Section 4.3.2) on the direct link is performed for each group of RC users in decreasing priority order. Then subcarriers on the coordinated link are allocated. Finally, subcarrier allocation for the BE users is performed on the remaining subcarriers with a greedy algorithm: each available subcarrier is allocated to the user that maximizes the channel coefficient in the subcarrier.

Once subcarrier allocation is determined, an iterative process is used to perform power control in a distributed way. At each iteration, power allocation is independently carried out within each BS, considering the inter-cell interference level computed in the previous iteration as a fixed value. The power control problem then becomes equivalent to a set of convex optimization problems that are distributed over the users and over the BSs. Regarding RC users, minimizing the sum power over all users in a given BS is equivalent to minimizing the sum power per user, because we do not impose a sum power constraint at that stage. The sum power constraint is indeed considered after power allocation through an independent admission control procedure. For the RC users requesting coordination, the power allocation process computes the power values of the direct BS and of the coordinated BS that minimize the sum power over both BSs. The details are given in Section 4.3.3. It should be noted that, as the data rate is an increasing function of the power, the minimum sum power is obtained when $\sum_{l\in\Theta_k} R_k^l = R_{k,\text{target}}$. Besides, BE users' sum data rate maximization can be performed independently in each BS. The power allocation problem in BS n_{BS} is consequently:

$$\begin{aligned}
&\min_{\{\mathbf{P}_{d,k},\mathbf{P}_{c,k}\}} \sum_{l\in\Theta_k} \left(P_{d,k}^l + P_{c,k}^l\right) \\
&\text{s.t. } \sum_{l\in\Theta_k} R_k^l = R_{k,\text{target}},\ \forall k \in \mathcal{S}_{\text{RC},d} \\
&\text{Then } \max \sum_{k\in\mathcal{S}_{\text{BE}}} R_k \\
&\text{s.t. } \sum_{k\in\mathcal{S}_{\text{BE}}} \sum_{l\in\Theta_k} P_{d,k}^l \le P' \\
&\text{where } P' = P_{\max} - \sum_{k\in\mathcal{S}_{\text{RC},d}} \sum_{l\in\Theta_k} P_{d,k}^l - \sum_{k\in\mathcal{S}_{\text{RC},c}} \sum_{l\in\Theta_k} P_{c,k}^l \\
&\text{s.t. } P_{d,k}^l \ge 0, \forall k \in \{\mathcal{S}_{\text{RC},d} \cup \mathcal{S}_{\text{BE}}\},\ l \in \Theta_k \\
&\text{s.t. } P_{c,k}^l \ge 0, \forall k \in \mathcal{S}_{\text{RC},c},\ l \in \Theta_k
\end{aligned} \tag{4.5}$$

On each BS, if the sum power used for all RC users to reach their target data rates is less than the maximum power $P_{\max}$, then the remaining power is used for BE users. Power allocation for BE users is performed with a Globally Optimal objective. The details of the corresponding water-filling algorithm are given in Appendix 4.A.

On the contrary, if the sum power used for RC users is more than the maximum power

$P_{\max}$, BE users are not served, and some RC users must be rejected. The admission control process starts with the RC users with the lowest priority levels. Among these users, the most demanding in terms of power level are rejected. This process stops when the sum power becomes less than $P_{\max}$.

4.3.2 Subcarrier allocation for RC users

The subcarrier allocation method for RC users is an iterative algorithm that determines the minimum set of subcarriers required to satisfy all RC users' data rate requirements.
At subcarrier allocation stage, no information is available on power and inter-cell interference. As a consequence, the BS performs subcarrier allocation under the assumption that each subcarrier is allocated the same power, $\frac{P_{\max}}{L_{\mathrm{SC}}}$, and that each user undergoes the same inter-cell interference plus noise within all subcarriers ($\tilde{I}$, which depends on the load and deployment's characteristics). Under these assumptions, minimizing the set of subcarriers is equivalent to minimizing the required sum power.
In this algorithm, the direct BS also considers the influence of the coordinated link on the users' data rate. To simplify subcarrier allocation, and avoid having to request CSI on the coordinated BS-to-user link at that stage, the direct BS only takes into account the path loss (including shadowing) between the coordinated BS and the user.
A proportion of the subcarriers of each BS is dedicated to coordinated users, with the restriction that each user should have at least one subcarrier allocated on its direct link. The proportion of subcarriers dedicated to coordinated users is

$$L_{\mathrm{SC},c} = \min \left\{ \left\lfloor \frac{K_c}{(K_c + K_d)} L_{\mathrm{SC}} \right\rfloor, L_{\mathrm{SC}} - K_d \right\}$$

where K_d is the total number of direct users, including RC and BE users, and K_c is the number of coordinated RC users. We denote the approximate data rate, obtained under the assumptions of equal power allocation and average interference, as $\tilde{R}_k$. The algorithm detailed hereunder is run for a set of users with the same target data rate, $R_{k,\mathrm{target}} = R_{\mathrm{target}}, \forall k$.

1. Initialization:

   (a) Set $\tilde{R}_k = 0$ for all users k

   (b) For each user k, find the subcarrier l on the direct link with highest channel coefficient $G^l_{d,k}$, and assign it to user k. If user k is a coordinated user, compute $\tilde{R}_k$ by taking into account the coordinated link.

2. Iterative phase: as long as there are free subcarriers and as long as at least one user does not reach the target data rate R_{target},

   (a) Find the user k that has the lowest data rate $\tilde{R}_k$.

   (b) For that user k, find the free subcarrier l on the direct link with the highest channel coefficient $G^l_{d,k}$, and assign it to user k.

   (c) Update $\tilde{R}_k$ with the additional approximate data rate obtained on subcarrier l, taking into account the coordinated link if user k is a coordinated user.

After direct subcarriers have been assigned to all RC users, for all coordinated users, the direct BS requests the subcarriers corresponding to the direct subcarriers of that user, on the coordinated BS. If a required subcarrier is already assigned to a direct user, and if some subcarriers are still available, the direct user is re-allocated to the free subcarrier that maximizes its channel coefficient.
Each coordinated user should have the same number of subcarriers on the coordinated link as on the direct link. No transmission takes place on the coordinated link for users who cannot meet this requirement.

4.3.3 Power control for RC users

Power control aims at minimizing the sum power over all RC users. At that step, we do not consider the maximum sum power constraint. Consequently, the power control objective can be distributed per user: the aim is to minimize the sum power, over the direct subcarriers and the coordinated subcarriers, required to reach the target data rate. For direct users, the solution is simply to water-fill over the user's subcarriers, under the data rate objective. For each coordinated user k, we consider the joint direct and coordinated power optimization problem:

$$\begin{aligned}
&\min_{\{\mathbf{P}_{d,k},\mathbf{P}_{c,k}\}} \sum_{l\in\Theta_k} \left(P_{d,k}^l + P_{c,k}^l\right) \\
&\text{s.t.} \sum_{l\in\Theta_k} R_k^l = R_{k,\text{target}} \\
&\text{s.t. } P_{c,k}^l \geq 0 \text{ and } P_{d,k}^l \geq 0, l \in \Theta_k
\end{aligned} \tag{4.6}$$

where the sum data rate is equal to

$$R_k = \frac{B_{\text{SC}}}{2} \sum_{l\in\Theta_k} \log_2\left(\left(1 + \frac{G_{d,k}^l P_{d,k}^l}{I_k^l}\right)^2 + \frac{G_{c,k}^l P_{c,k}^l}{I_k^l}\right)$$

This problem is solved jointly over $\{\mathbf{P}_{d,k}, \mathbf{P}_{c,k}\}$ through convex optimization. The analytical solution is detailed in Appendix 4.B. Power allocation requires both direct BS-to-user and coordinated BS-to-user CSI. Consequently, the distributed implementation may be performed in two ways: either in the terminal or in the BSs. In the first implementation, the mobile terminal performs power control and then feeds back the required transmit power values to the direct and coordinated BSs independently. In the second implementation, the mobile terminal feeds back CSI on the coordinated BS-to-user link to the direct BS. The direct BS can then emulate the joint solution, and only provide the direct power. It also transmits the coordinated BS-to-user link CSI to the coordinated BS (this is less resource-consuming than a direct CSI transmission between the user and the coordinated BS, as we assume that the inter-BS link is perfect, contrary to the terminal-user link). Finally the coordinated BS provides the additional power required in order to reach the target data rate.
The first implementation is less costly than the second one in terms of resource consumption associated with signaling. However, it may only be feasible if the terminal has the required hardware and software to iterativaly compute the power values.

4.4 Numerical results

The studied network and the simulation parameters are the same as in Chapter 3 (see Section 3.5).
Our method is compared with two methods that use the same subcarrier allocation, but different power allocations. The first method is Equal Power Allocation (EPA), where power value $\frac{P_{\max}}{L_{\mathrm{SC}}}$ is allocated to each subcarrier. The second method has been proposed in [51] and summarized in Section 4.2.3., for scenarios with both RC and BE users. In the following, we refer to our method as 'PM', for Power Minimization. 'C' means that network coordination is used, and 'NC' means that there is no network coordination.

4.4.1 Performance of RC users

Scenarios with RC users only are first studied, in order to evaluate the rejection probability for given target data rates and load levels. A RC user is rejected if it does not reach its required data rate. We consider three cases: 32 users per cell requesting 144 kbits/s, 96 users per cell requesting 64 kbits/s, and 192 users per cell requesting 32 kbits/s. The rejection probability with and without coordination, with EPA and PM methods, is represented on Fig. 4.1. We can first notice that PM is efficient compared to EPA, with and without coordination. Besides, network coordination decreases the rejection probability in all the scenarios. Fig. 4.2 shows the average data rate per ring, and Fig. 4.3 shows the percentage of rejected users per ring, both for the 32 users per cell scenario. With the proposed method PM, the rejected users mainly belong to the rings at the border of the cell, because admission control rejects the users with the highest power requests if the maximum sum power is exceeded. Network coordination decreases the power required to satisfy all users' data rate demands, therefore restricting the need for admission control, and preserving fairness. We can notice that with EPA, in all scenarios, the average data rate in the central rings is very high. This triggers an inefficient use of power for some RC users, as the target data rate is importantly exceeded. Besides, the high average data rates obtained in almost all rings with EPA should be considered with caution as the rejection probability is very high, even with network coordination (see Fig. 4.3): it exceeds 40% after the third ring with 32 and 96 users per cell, and after the sixth ring with 192 users per cell. Therefore, EPA is efficient in terms of sum data rate, but is totally unfair.

4.4.2 Performance with RC and BE users

We now consider a scenario with both RC and BE users, where the RC users do not require all the power, so that the BE users may also served. Our proposed method is compared with the method from [51], where the proportion of power dedicated to RC and BE users is fixed before power control: P_{RC} and $P_{\mathrm{BE}} = P_{\max} - P_{\mathrm{RC}}$. The power for RC users is proportional to the number of subcarriers allocated to RC users. However, contrary to [51], we do not use RA objective for RC users power allocation, but MA objective. Admission control rejects the users with highest power values, until the sum power becomes lower than P_{RC}. Numerical results are given for 16 RC users per cell requesting 64 kbits/s, and 16 BE users per cell. The method adapted from [51] is denoted as 'AKSK'.
Network coordination's cost is evaluated through power consumption. Table 4.1 shows that

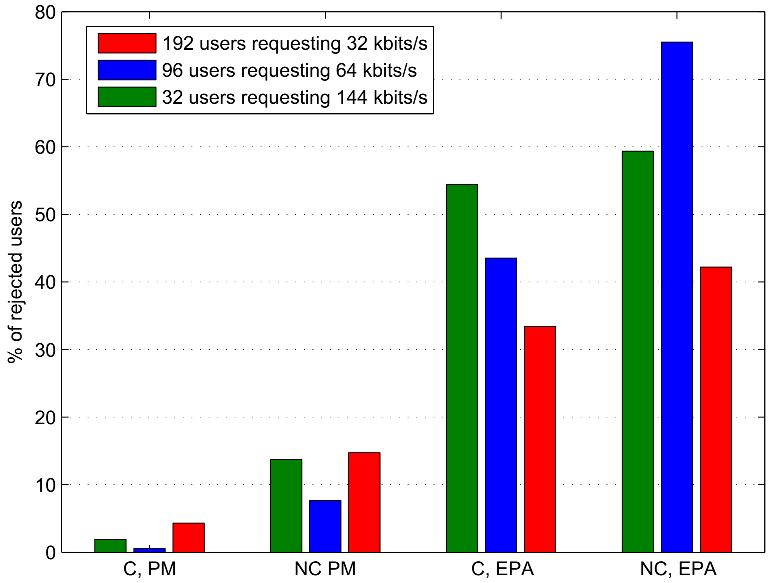


Figure 4.1: Percentage of rejected RC users with different load and target data rate scenarios

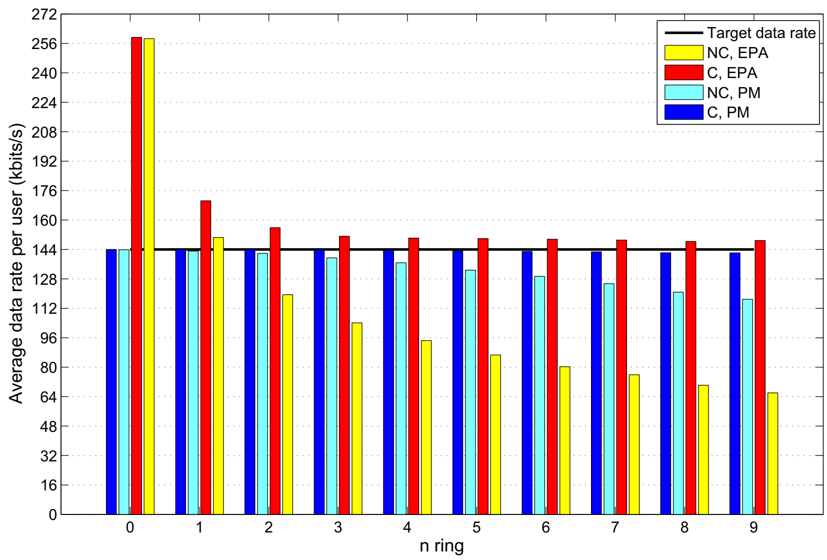


Figure 4.2: Average data rate per ring, 32 users per cell requesting 144 kbits/s

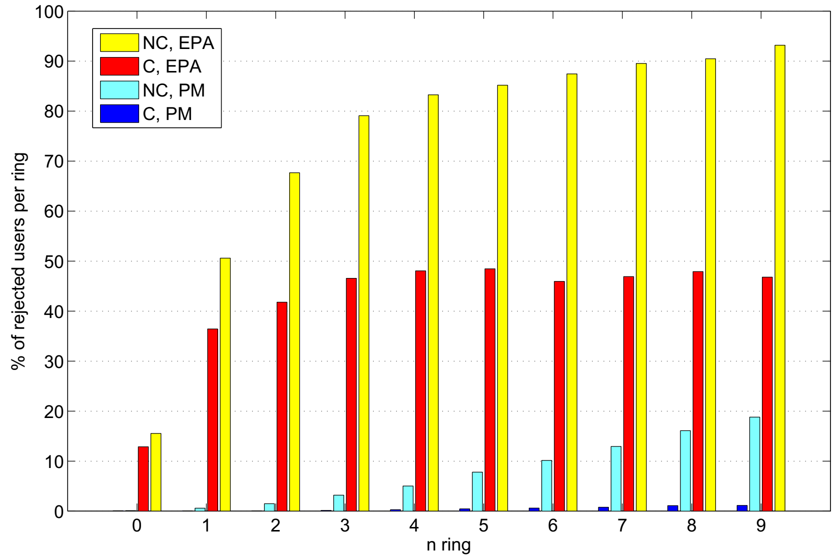


Figure 4.3: Percentage of rejected users per ring, 32 users per cell requesting 144 kbits/s

Table 4.1: Comparison of power consumption

	Coordination	No coordination
PM (required power)	1.95 W	5.19 W
AKSK (evaluated by algorithm)	11.07 W	16.15 W
EPA (required power)	10.99 W	10.31 W

the sum power required to serve both direct and coordinated users in coordination case is lower than the power required to serve direct users only without coordination. This is due to the inter-cell interference decrease and to the fact that under some shadowing and fading circumstances, transmitting on the coordinated BS avoids high power transmissions from the direct BS. These results show that for RC users located at the border of cells, network coordination spares some power. AKSK highly overestimates the power required for RC users. Consequently, few power is allocated to BE users, which explains for their low data rate (see Fig. 4.4 and 4.5). EPA leads to even worse results, although more power is dedicated to BE users than with AKSK. This is due to inter-cell interference, which is higher with EPA than with the two other methods, as there is no power control on RC users. Power control on BE users does not make much difference, as maximizing the sum data rate over BE users almost leads to equal power allocation [53].

Besides, we can notice that using network coordination importantly increases the data rate for BE users. The gain brought by network coordination comes from both subcarrier and power allocations. Indeed, the average number of subcarriers allocated per BE user in the network coordination case is 7, whereas it is only 3 in the no-coordination case. This comes

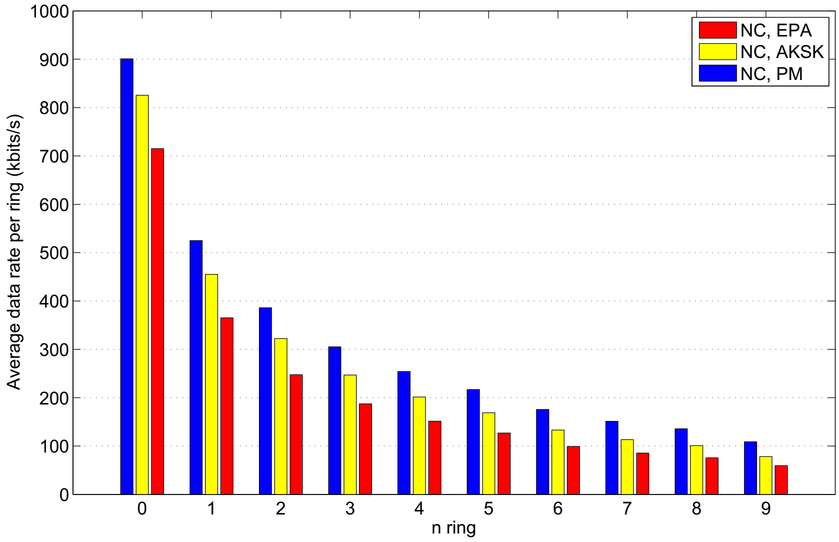


Figure 4.4: Average data rate per ring without network coordination for BE users

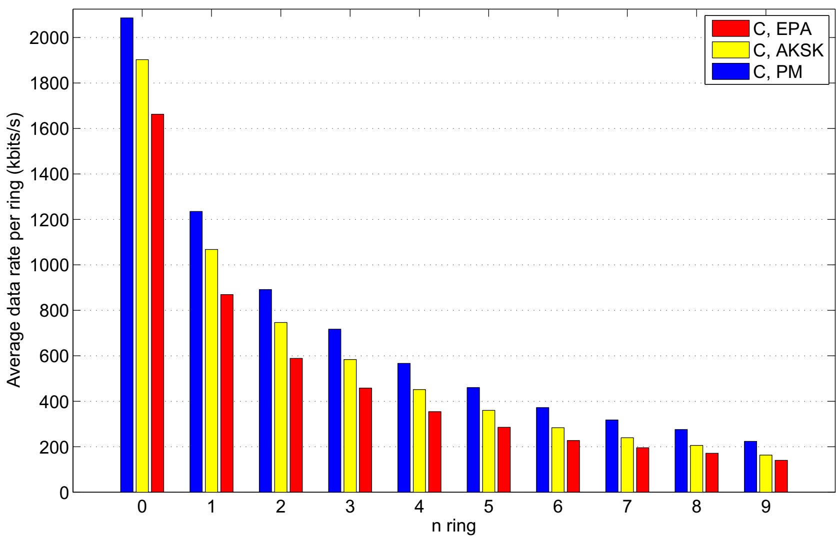


Figure 4.5: Average data rate per ring with network coordination for BE users

from the fact that less subcarriers are required for RC users when network coordination is used, thanks to the capacity gain brought in each subcarrier. As a consequence, more subcarriers are then available for BE users. Besides, the power available for BE users is far lower without coordination than with coordination, so even the BE users that actually get subcarriers are limited by power allocation. With PM, the average power per BE user is 2.5 to 3.5 times higher with coordination than without coordination.

The subcarrier allocation process overestimates the required number of subcarriers, because it assumes equal power allocation and equal inter-cell interference in each subcarrier. With our proposed method, power control allows RC users to reach their target data rate with less subcarriers than estimated at subcarrier allocation step. This is mainly important for the users located at the border of cells, that require the most subcarriers. This limitation is somewhat mitigated by network coordination, as the evaluated capacity for coordinated users in the subcarrier allocation process takes into account the coordinated link. Therefore, the number of required subcarriers on the direct link as evaluated by subcarrier allocation is lower than in the no-coordination case. These issues show that separating resource allocation into two independent phases may lead to inconsistencies. However, this is required due to the constraint of distributed networks, and due to inter-cell interference. In practical implementation, resource allocation could be improved by considering the past power allocation values within actual subcarrier allocation, for RC users whose average channel conditions little vary during several TTIs. When considering per TTI resource allocation, power allocation could be emulated within each BS prior to subcarrier allocation to improve its accuracy. Such an emulation method is however quite complex and does not bring much performance gain compared to the equal power allocation assumption considered in our method. The main limitation at subcarrier allocation step is that the received inter-cell interference per subcarrier is hard to predict. Besides, the opportunistic power allocation methods that we have derived tend to increase the distribution's spread of inter-cell interference over the subcarriers. Consequently, and as the performance gains are still quite important, our proposed method is a first step for QoS-aware resource allocation with network coordination.

4.5 Conclusion

In this chapter, we have described a resource allocation method for users with different QoS requirements in downlink OFDMA. Our method uses network coordination to fulfill the QoS objectives of the RC users located at the border of cells. Thanks to the diversity and interference mitigation gains brought by performing power control in a coordinated way between the direct BS and the most potentially interfering neighboring BS, these users require less subcarriers and power. The resource allocation method aims at minimizing the sum power required for RC users, so that the sum data rate of BE users can be maximized. Subcarrier allocation and power control are performed separately and both favor RC users over BE users. Numerical results show that the proposed method increases the ratio of satisfied RC users, as well as the sum data rate of BE users. It is more efficient than Equal Power Allocation and than other QoS-differentiating methods. Users' prioritization, favoring RC users over BE users regarding resource access, already provides performance improvements. Network coordination renders the proposed prioritized method even more efficient, thanks to the data rate gain for RC users that would otherwise be in bad radio conditions.

If the network architecture allows network coordination to be implemented, then it will be an interesting feature to improve the performance in multi-cell networks. However, network coordination will not be feasible in ad hoc networks, or in cellular networks where the inter-BS link is not reliable enough. In these cases, relaying protocols with unperfect transmission between the BSs may be considered. We do not investigate these aspects in the thesis. On the contrary, in the following chapters, we will focus on a totally distributed network without network coordination. No inter-BS transmission is required. Similarly to Chapters 3 and 4, each BS is responsible for allocating radio resources to its served users, but we also add the constraint that BSs cannot cooperate regarding users' data transmission. Distributed resource allocation methods are determined under these constraints for RC users in Chapter 5 and 6, and for BE users in Chapter 7.

APPENDIX

4.A Sum data rate maximization for BE users: water-filling

We use the notation $a^l_{d,k} = \frac{G^l_{d,k}}{I^l_k}$. On each BS, the power control objective for BE users is:

$$\begin{aligned}
&\max_{\mathbf{P}_d} \; B_{\mathrm{SC}} \sum_{k\in\mathcal{S}_{\mathrm{BE}}} \sum_{l\in\Theta_k} \log_2\left(1 + a^l_{d,k} P^l_{d,k}\right) \\
&\text{s.t.} \; \sum_{k\in\mathcal{S}_{\mathrm{BE}}} \sum_{l\in\Theta_k} P^l_{d,k} \le P_{\mathrm{BE}} \\
&\text{s.t.} \; P^l_{d,k} \ge 0, \; \forall k \in \mathcal{S}_{\mathrm{BE}}, \; l \in \Theta_k
\end{aligned} \tag{4.7}$$

P_{BE} is the remaining power available for BE users after RC users have been served.
This is a convex optimization problem in $\mathbf{P}_d$. The Lagrangian (normalized by $\frac{\log(2)}{B_{\mathrm{SC}}}$) is:

$$\begin{aligned}
L(\mathbf{P}_d, \lambda, \boldsymbol{\upsilon}) = & \sum_{k\in\mathcal{S}_{\mathrm{BE}}} \sum_{l\in\Theta_k} \log\left(1 + a^l_{d,k} P^l_{d,k}\right) \\
& + \lambda \left(P_{\mathrm{BE}} - \sum_{k\in\mathcal{S}_{\mathrm{BE}}} \sum_{l\in\Theta_k} P^l_{d,k} \right) + \sum_{k\in\mathcal{S}_{\mathrm{BE}}} \sum_{l\in\Theta_k} \upsilon^l_k P^l_{d,k}
\end{aligned} \tag{4.8}$$

where λ and $\boldsymbol{\upsilon}$ are Lagrange multipliers.
Deriving the Lagrangian with regard to $P^l_{d,k}$ gives:

$$\frac{a^l_{d,k}}{1 + a^l_{d,k} P^l_{d,k}} - \lambda + \upsilon^l_k = 0 \tag{4.9}$$

The KKT conditions [24] impose that $\upsilon^l_k \ge 0$. Consequently

$$\lambda \ge \frac{a^l_{d,k}}{1 + a^l_{d,k} P^l_{d,k}} \tag{4.10}$$

Another KKT condition imposes that $\upsilon^l_k \; P^l_{d,k} = 0$. Therefore

$$P^l_{d,k} \left(\lambda - \frac{a^l_{d,k}}{1 + a^l_{d,k} P^l_{d,k}} \right) = 0 \tag{4.11}$$

If $\lambda > \frac{a^l_{d,k}}{1+a^l_{d,k}P^l_{d,k}}$, then this condition is only fulfilled if $P^l_{d,k} = 0$. Else $P^l_{d,k} = \frac{1}{\lambda} - \frac{1}{a^l_{d,k}}$.
Finally, the solution is:

$$P^l_{d,k} = \left[\frac{1}{\lambda} - \frac{1}{a^l_{d,k}} \right]^+ \tag{4.12}$$

where the constant λ is chosen so as to fulfill the sum power constraint.

4.B Power minimization for RC users

Let us introduce the notations $a^l_{d,k} = \frac{G^l_{d,k}}{I^l_k}$ and $b^l_{c,k} = \frac{G^l_{c,k}}{I^l_k}$. The optimization problem is convex in the set defined by $P^l_{c,k} \in \mathbb{R}_+$ and $P^l_{d,k} \in [\frac{1}{a^l_{d,k}}\left(\sqrt{b^l_{c,k}P^l_{c,k}} - 1\right), +\infty[$. The

Lagrangian is:

$$L(\mathbf{P}_{d,k}, \mathbf{P}_{c,k}, \lambda, \boldsymbol{\alpha}_d, \boldsymbol{\gamma}_c) = \sum_{l\in\Theta_k} P^l_{d,k} + \sum_{l\in\Theta_k} P^l_{c,k} + \frac{2\log(2)\lambda}{B_{\mathrm{SC}}}\left(R_{k,\mathrm{target}} - R_k\right) - \sum_{l\in\Theta_k} \alpha^l_d P^l_{d,k} - \sum_{l\in\Theta_k} \gamma^l_c P^l_{c,k} \tag{4.13}$$

where λ, $\boldsymbol{\alpha}_d$ and $\boldsymbol{\gamma}_c$ are Lagrange multipliers.
Setting the derivative of $L(\mathbf{P}_{d,k}, \mathbf{P}_{c,k}, \lambda, \boldsymbol{\alpha}_d, \boldsymbol{\gamma}_c)$ with regard to $P^l_{d,k}$ and $P^l_{c,k}$, leads to the following equations:

$$1 - \alpha^l_d - \lambda\, \frac{2a^l_{d,k}(1 + a^l_{d,k}P^l_{d,k})}{(1 + a^l_{d,k}P^l_{d,k})^2 + b^l_{c,k}P^l_{c,k}} = 0 \tag{4.14}$$

and

$$1 - \gamma^l_c - \lambda\, \frac{b^l_{c,k}}{(1 + a^l_{d,k}P^l_{d,k})^2 + b^l_{c,k}P^l_{c,k}} = 0 \tag{4.15}$$

The KKT conditions impose that $\alpha^l_d \geq 0$ and $\gamma^l_c \geq 0$. Therefore:

$$1 - \lambda\, \frac{2a^l_{d,k}(1 + a^l_{d,k}P^l_{d,k})}{(1 + a^l_{d,k}P^l_{d,k})^2 + b^l_{c,k}P^l_{c,k}} \geq 0 \tag{4.16}$$

and

$$1 - \lambda\, \frac{b^l_{c,k}}{(1 + a^l_{d,k}P^l_{d,k})^2 + b^l_{c,k}P^l_{c,k}} \geq 0 \tag{4.17}$$

The KKT conditions $\alpha^l_d P^l_{d,k} = 0$ and $\gamma^l_c P^l_{c,k} = 0$ provide the following equations:

$$P^l_{d,k}\left(1 - \lambda\, \frac{2a^l_{d,k}(1 + a^l_{d,k}P^l_{d,k})}{(1 + a^l_{d,k}P^l_{d,k})^2 + b^l_{c,k}P^l_{c,k}}\right) = 0 \tag{4.18}$$

$$P^l_{c,k}\left(1 - \lambda\, \frac{b^l_{c,k}}{(1 + a^l_{d,k}P^l_{d,k})^2 + b^l_{c,k}P^l_{c,k}}\right) = 0 \tag{4.19}$$

If $P^l_{d,k} > 0$ and $P^l_{c,k} > 0$ then we necessarily have:

$$1 = \lambda\, \frac{2a^l_{d,k}(1 + a^l_{d,k}P^l_{d,k})}{(1 + a^l_{d,k}P^l_{d,k,})^2 + b^l_{c,k}P^l_{c,k}} \tag{4.20}$$

and

$$1 = \lambda\, \frac{b^l_{c,k}}{(1 + a^l_{d,k}P^l_{d,k})^2 + b^l_{c,k}P^l_{c,k}} \tag{4.21}$$

This imposes that $\lambda > 0$ and that $b^l_{c,k} = 2a^l_{d,k}(1 + a^l_{d,k}P^l_{d,k})$. Consequently we must have:

$$P^l_{d,k} = \frac{b^l_{c,k} - 2a^l_{d,k}}{2(a^l_{d,k})^2} \tag{4.22}$$

and

$$P^l_{c,k} = \lambda - \frac{b^l_{c,k}}{4(a^l_{d,k})^2} \tag{4.23}$$

Thus, there exists a joint solution with $P^l_{d,k} > 0$ and $P^l_{c,k} > 0$ if the following conditions are fulfilled:

$$\begin{aligned} b^l_{c,k} - 2a^l_{d,k} &> 0 \\ \lambda &> \frac{b^l_{c,k}}{4(a^l_{d,k})^2} \end{aligned} \tag{4.24}$$

So the existence of a joint solution depends on the relative values of the direct and coordinated link in each subcarrier, and also on the sum data rate constraint, as λ must be defined so as to reach the target data rate.

We can notice that if $b^l_{c,k} = 2a^l_{d,k}$, then the solution obtained with the joint optimization is $P^l_{d,k} = 0$ and $P^l_{c,k} = \lambda - \frac{1}{b^l_{c,k}}$, which is equivalent to the water-filling solution for $P^l_{c,k}$ only (assuming its value is positive, which is imposed by the constraint $\lambda > \frac{b^l_{c,k}}{4(a^l_{d,k})^2}$). Similarly, if $\lambda = \frac{b^l_{c,k}}{4(a^l_{d,k})^2}$, then we get $P^l_{c,k} = 0$ and $P^l_{d,k} = 2\lambda - \frac{1}{a^l_{d,k}}$, which is equivalent to the water-filling solution for $P^l_{d,k}$ only (assuming its value is positive, which is imposed by the constraint $b^l_{c,k} - 2a^l_{d,k} > 0$). Consequently, the joint solutions are also valid if there is an equality on one of the two constraints.
The optimization problem is convex if $(1 + a^l_{d,k}P^l_{d,k})^2 \geq b^l_{c,k}P^l_{c,k}$. If a joint solution with $P^l_{d,k} > 0$ and $P^l_{c,k} > 0$ exists, this imposes that $\lambda \leq \frac{b^l_{c,k}}{2(a^l_{d,k})^2}$. In that case, the KKT conditions lead to the optimum solution of the primal problem.
We now study the cases where a joint positive solution is not feasible.
If $b^l_{c,k} < 2a^l_{d,k}$, then either $P^l_{d,k} = 0$ and $P^l_{c,k} > 0$, or $P^l_{d,k} > 0$ and $P^l_{c,k} = 0$. Suppose that $P^l_{d,k} = 0$ and $P^l_{c,k} > 0$. Equation (4.18) leads to

$$\lambda = \frac{1 + b^l_{d,k}P^l_{c,k}}{b^l_{d,k}}$$

and equation (4.17) imposes that:

$$1 - \lambda \frac{2a^l_{d,k}}{1 + b^l_{c,k}P^l_{c,k}} \geq 0$$

Therefore:

$$\lambda \leq \frac{1 + b^l_{c,k}P^l_{c,k}}{2a^l_{d,k}} = \frac{\lambda b^l_{c,k}}{2a^l_{d,k}}$$

As $b^l_{c,k} < 2a^l_{d,k}$, $\frac{\lambda b^l_{c,k}}{2a^l_{d,k}} < \lambda$. Consequently we get the following inequation: $\lambda < \lambda$. This is contradictory, so the case $P^l_{d,k} = 0$ and $P^l_{c,k} > 0$ is not feasible.
Finally, if $b^l_{c,k} < 2a^l_{d,k}$, then $P^l_{c,k} = 0$ and

$$P^l_{d,k} = \left[2\lambda - \frac{1}{a^l_{d,k}}\right]^+ \tag{4.25}$$

If $\lambda < \frac{b^l_{c,k}}{4a^{l2}_{c,k}}$, then either $P^l_{d,k} = 0$ and $P^l_{c,k} > 0$, or $P^l_{d,k} > 0$ and $P^l_{c,k} = 0$. Suppose that $P^l_{d,k} > 0$ and $P^l_{c,k} = 0$. Equation (4.19) leads to

$$\lambda = \frac{1 + a^l_{d,k} P^l_{d,k}}{2a^l_{d,k}}$$

and equation (4.16) imposes that:

$$1 - \lambda \frac{b^l_{c,k}}{\left(1 + a^l_{d,k} P^l_{d,k}\right)^2} \geq 0$$

Therefore:

$$\lambda \leq \frac{\left(1 + a^l_{d,k} P^l_{d,k}\right)^2}{b^l_{c,k}} = \frac{(2a^l_{d,k}\lambda)^2}{b^l_{c,k}} = \frac{4(a^l_{d,k})^2\lambda^2}{b^l_{c,k}}$$

As $\lambda < \frac{b^l_{c,k}}{4(a^l_{d,k})^2}$, $\frac{4(a^l_{d,k})^2\lambda^2}{b^l_{c,k}} < \lambda$. Consequently we get the following inequation: $\lambda < \lambda$. This is contradictory, so the case $P^l_{d,k} > 0$ and $P^l_{c,k} = 0$ is not feasible.
Finally, if $\lambda < \frac{b^l_{c,k}}{4a^{l2}_{d,k}}$, then $P^l_{d,k} = 0$ and

$$P^l_{c,k} = \left[\lambda - \frac{1}{b^l_{c,k}}\right]^+ \tag{4.26}$$

If $\left(1 + a^l_{d,k} P^l_{d,k}\right)^2 < b^l_{c,k} P^l_{c,k}$, the optimization problem is not convex with regard to variable $P^l_{d,k}$. The KKT conditions consequently do not necessarily lead to the optimum solution. We can notice that in that set, $P^l_{c,k} > 0$. If a solution with $P^l_{d,k} > 0$ exists, then equations (4.22) and (4.23) lead to $\lambda \geq \frac{b^l_{c,k}}{2(a^l_{d,k})^2}$. Two cases are possible: in the first case, we assume that the joint solution is the optimum even outside of the convex set. As a consequence we never obtain $P^l_{d,k} = 0$. In the second case, we assume that the joint solution is not the optimum outside of the convex set. Consequently $P^l_{d,k} = 0$ and $P^l_{c,k} = \left[\lambda - \frac{1}{b^l_{c,k}}\right]^+$. The same data rate, $\frac{B_{\mathrm{SC}}}{2}\log(\lambda b^l_{c,k})$, is obtained on subcarrier l in both cases, if $\lambda b^l_{c,k} > 1$. The sum power required on subcarrier l with the joint solution is: $S_1 = \frac{b^l_{c,k}}{2a^{l2}_{d,k}} - \frac{1}{2a^l_{d,k}} + \lambda - \frac{b^l_{c,k}}{4a^{l2}_{d,k}}$.
The sum power required on subcarrier l with $P^l_{d,k} = 0$ is: $S_2 = \lambda - \frac{1}{b^l_{c,k}}$.
The sum power difference is independent of λ:

$$S_1 - S_2 = \frac{(b^l_{c,k} - 2(a^l_{d,k})^2}{4(a^l_{d,k})^2 b^l_{c,k}} \geq 0$$

Consequently to minimize the sum power, the second case is more efficient. We can however notice that both methods have been tested via simulations, and the power differences observed are almost negligible.

To conclude, the power allocation algorithm as summarized as follows:

1. Set an initial value for λ.

2. For each subcarrier $l \in \Theta_k$:

   (a) If $b^l_{c,k} < 2a^l_{d,k}$ then set $P^l_{c,k} = 0$ and $P^l_{d,k}$ according to equation (4.25).

   (b) Else

      i. If $\frac{b^l_{c,k}}{2(a^l_{d,k})^2} \geq \lambda \geq \frac{b^l_{c,k}}{4a^{l2}_{d,k}}$, set $P^l_{d,k}$ and $P^l_{c,k}$ according to equations (4.22) and (4.23).

      ii. Else, set $P^l_{d,k} = 0$ and $P^l_{c,k}$ according to equation (4.26).

3. If R_{target} is reached with λ, stop. Else, update λ and keep searching with bisection search.

Chapter 5

Distributed resource allocation for Rate Constrained users

5.1 Introduction

In this chapter and in the remainder of the dissertation, we consider fully distributed networks corresponding to the flat architecture of Fig. 1.2, where network coordination can no longer be used. The resource allocation methods proposed and discussed in Chapters 5, 6 and 7 may consequently be applied not only to cellular networks, but also to ad hoc networks. In Chapters 5 and 6, resource allocation for Rate Constrained users is investigated. The QoS indicator of RC users is their data rate. Each RC user should reach its target data rate in each TTI. In multi-cell networks, this should be achieved at the cost of the least possible power, in order to mitigate inter-cell interference. This set of QoS objectives is written as a Margin Adaptive resource allocation problem, i.e., minimization of the sum power required to reach the target data rates of all users.

In this chapter, we consider the distributed MA resource allocation problem in SISO multi-cell OFDMA networks. As seen in the state of the art of Chapter 4, Section 4.2, the MA problem in the multi-cell, multi-carrier case has not yet been treated in a distributed way in the literature. The main issue is the feasibility of MA resource allocation, i.e., whether the target data rates are achievable under the constraints of orthogonal intra-cell subcarrier allocation and of positive power allocation. A feasibility criterion for power control in the single-channel case, when the target SINR per user is known, has been obtained by Zander [66] [67], Foschini and Miljanic [68], and Yates [69]. In that simplest case, the feasibility criterion requires full knowledge of the CSI on all direct and interfering links, and must consequently be determined by a central controller. However, if the problem is feasible, then distributed power control converges toward the globally optimal solution. Other studies have extended this method to cases where the target SINR per user can vary, in order to add a dynamic behavior and benefit from time and multi-user diversity [70]. The game-theoretic proposal from [63] includes the feasibility criterion within a global mediator, that coordinates resource allocation on all cells.

In multi-cell OFDMA, inter-cell interference should be considered when optimizing resource allocation within each cell. Subcarriers can however not be seen as independent interference channels, because each user may be allocated several subcarriers per cell. The data rate fulfillment objective implies a joint optimization over all the subcarriers assigned to the same user. The interactions and dependences between multi-user resource allocation per cell and multi-cell interference per subcarrier increase the complexity of the problem, that may become intractable. Therefore, it is necessary to decompose the original problem into several sub-problems. This may be performed through parallel decomposition over several physical entities, such as the BSs. In our case, the parallel decomposition involves that resource allocation be performed iteratively, so as to take into account inter-cell interference. The decomposition may also be logical. As in the previous chapters, we separate resource allocation into two successive steps: subcarrier allocation and power control. We add a convergence constraint on power control, that is valid when inter-cell interference is the limiting feature. It is defined per subcarrier and user, and it only requires distributed information relative to each BS. The convergence condition introduces a new step between subcarrier allocation and power control, that consists of setting the target SINR per user and subcarrier.

In Section 5.3, we investigate the power convergence criterion for RC users. This work is based on the results from [66] [67] [68] [69] on the interference channel. Then we determine a distributed criterion per subcarrier to ensure that distributed power control will necessarily converge. The distributed criterion is based upon an upper bound on the spectral radius of the interference matrix characterizing each subcarrier. The obtained upper bound is the infinity norm of the interference matrix, that gives a useful and simple criterion, which will be denoted in the remaining of the dissertation as E criterion.

Section 5.4 then details the proposed resource allocation method for RC users. Resource allocation consists of three steps: subcarrier allocation, which aims at maximizing the probability that power control will converge, by selecting for each user the subcarriers that maximize the convergence criterion; target SINRs setting, subcarrier per subcarrier, which aims at determining a target SINRs set per user that both achieves the target data rate, and enforces the convergence of distributed power control when inter-cell interference is significant enough; and finally distributed power control, that necessarily converges to the globally optimal solution. Each subcarrier can then be treated as an independent interference channel in the power control process, thanks to the efficient setting of the target SINRs. While subcarrier allocation is performed only once, power control, which is composed of target SINRs setting and power allocation, is performed iteratively over the BSs.

Finally, the proposed method is assessed via numerical results in Section 5.5. It is compared with iterative water-filling [61], with two different subcarrier allocation methods.

The main contributions of this chapter are:

- A simple convergence criterion for power control with inter-cell interference is obtained per user and subcarrier. This criterion is fully distributed. Besides, if the criterion is verified, then independent power control, performed iteratively for each user by only sensing its received inter-cell interference and adapting its power accordingly so as to reach its target SINR, will necessarily converge. Consequently, in a distributed network, if each BS fulfills this criterion for all of its users and subcarriers, then we can predict that distributed power control will not diverge. This is an important improvement compared to classical methods such as iterative water-filling, where divergence cannot be foreseen, and compared to other methods using a centralized mediator to avoid divergence [63], that are not fully distributed.

- An iterative resource allocation method is described for RC users. Subcarrier allocation is first performed with a classical iterative method, in which we propose to use criterion E instead of the classical channel gain criterion G. This leads to a wider range of variations for the target SINR. Power allocation is performed by first determining the target SINR per user and subcarrier, and then running power control over each subcarrier for these sets of target SINRs. Power allocation is an iterative process, as inter-cell interference must be considered within target SINR determination. For a given inter-cell interference level, each user computes its target SINRs set, with the aim to reach the target data rate, while fulfilling the power control convergence criterion on each subcarrier, if inter-cell interference is not negligible. A distributed criterion is used to determine whether the convergence criterion should be considered or not. Then on each subcarrier, power control is performed independently in each subcarrier.

- The proposed method is compared with iterative water-filling. It avoids the power divergence situations that occur with iterative water-filling at medium to high load. The proposed method efficiently adapts to the inter-cell interference level, using the E criterion in power allocation only when inter-cell interference becomes the limiting feature.

5.2 Margin Adaptive problem

We consider the MA resource allocation problem in downlink, in a multi-cell OFDMA network. This problem aims at determining the subcarrier and power allocations required to achieve the target data rates for all users, while minimizing the sum power.
Let $\mathcal{N}$ be a network composed of N_{BS} BSs and $K_{\mathcal{N}}$ users. Each BS transmits in L_{SC} orthogonal subcarriers. Θ_k is the set of subcarriers assigned to user k by its BS. Let $R_{k,\mathrm{target}}$ be the target data rate of user k, R_k^l the data rate of user k in subcarrier l, and P_k^l the power transmitted to user k by its serving BS in subcarrier l. To simplify the notations, we equivalently denote P_k^l as $P_{n_{\mathrm{BS}}}^l$ when BS n_{BS} serves user k in subcarrier l. The MA

optimization problem is:

$$\begin{aligned}
&\min_{\{\mathbf{P},\boldsymbol{\Theta}\}} \sum_{k=1}^{K_{\mathcal{N}}} \sum_{l\in\Theta_k} P_k^l \\
&\text{s. t.} \sum_{l\in\Theta_k} R_k^l \geq R_{k,\text{target}}, \forall k \in \{1,...,K_{\mathcal{N}}\} \\
&\text{s. t.} \sum_{l=1}^{L_{\text{SC}}} P_{n_{\text{BS}}}^l \leq P_{\max}, \forall n_{\text{BS}} \in \{1,...,N_{\text{BS}}\} \\
&\text{s. t.} \sum_{l=1}^{L_{\text{SC}}} P_{n_{\text{BS}}}^l \geq 0, \forall (n_{\text{BS}}, l) \in \{1,...,N_{\text{BS}}\} \times \{1,...,L_{\text{SC}}\} \\
&\text{s.t. } \Theta_k \cap \Theta_{k'} = \emptyset, \forall (k,k') \text{ served by the same BS}, k \neq k'
\end{aligned} \tag{5.1}$$

Several levels of decompositions are necessary to turn this NP-hard problem [71] into a set of sub-problems with reasonable complexity. First, subcarrier allocation and power control are performed in sequence [52] [54]. For implementation purposes in distributed networks, the original problem is also separated per BS. The MA problem is then solved iteratively in each BS, with fixed inter-cell interference levels, that have been computed in the previous iteration. Finally, the per cell sum power constraint $\sum_{l=1}^{L_{\text{SC}}} P_{n_{\text{BS}}}^l \leq P_{\max}, \forall n_{\text{BS}} \in \{1,...,N_{\text{BS}}\}$ is not directly considered within power allocation, but an admission control step is added afterward to take it into account. Thus, power allocation can be decomposed over the users of the same cell.

Before detailing the MA resource allocation method in Section 5.4, we first investigate the convergence of distributed power control in OFDMA.

5.3 Distributed power control convergence in OFDMA

5.3.1 Power control on the interference channel

Let us first consider the case when the target SINR per subcarrier and per user is known. We assume that there is no maximum transmit power constraint per subcarrier or per cell. The power allocation problem is then spread over independent interference channels. On each subcarrier, power control can be solved as a TDMA power allocation problem. We here remind of the main results on power control on the interference channel from [66] [67] [68] [69].

The SINR of user k in subchannel l is

$$\Gamma_k^l = \frac{P_k^l G_{k,k}^l}{\sum_{\{n=1,n\neq k\}}^{N_{\text{BS}}} P_n^l G_{n,k}^l + N_0} \tag{5.2}$$

where $G_{n,k}^l$ is the channel gain between BS n and user k, including path loss, shadowing and fast fading, and N_0 is the noise variance. To simplify notations, we denote the channel gain between user k and its serving BS in subcarrier l as $G_{k,k}^l$.

Let γ_k^l be the target SINR for user k in subcarrier l. A maximum of $(N_{\text{BS}}-1)$ links interfere user k in subcarrier l, as at most one user is allocated per subcarrier in each BS.

Let us define $\mathbf{v}^l = [v_1^l, ..., v_{N_{\text{BS}}}^l]'$ with $v_k^l = N_0\gamma_k^l / G_{k,k}^l$, $\mathbf{D}^l = \text{diag}\left\{\gamma_1^l, ..., \gamma_{N_{\text{BS}}}^l\right\}$, $(\mathbf{F}^l)_{(k,n)} =$

$\frac{G^l_{n,k}}{G^l_{k,k}}$ if $n \neq k$ and $(\mathbf{F}^l)_{(k,k)} = 0$. Power control aims at reaching the target SINR for each user. This corresponds to the following set of objectives:

$$\Gamma^l_k \geq \gamma^l_k, \forall (k,l) \tag{5.3}$$

It can be written as:

$$(\mathbf{I}_{N_{\mathrm{BS}}} - \mathbf{D}^l\mathbf{F}^l)\mathbf{P}^l \geq \mathbf{v}^l, \forall l \tag{5.4}$$

By Perron-Frobenius theorem [72], there exists a positive power allocation if and only if the maximum modulus of the eigenvalues of $\mathbf{D}^l\mathbf{F}^l$, i.e. the spectral radius $\rho(\mathbf{D}^l\mathbf{F}^l)$, is less than 1. If $\rho(\mathbf{D}^l\mathbf{F}^l) < 1$, the system is feasible, and the optimal power solution is

$$\mathbf{P}^l = (\mathbf{I}_{N_{\mathrm{BS}}} - \mathbf{D}^l\mathbf{F}^l)^{-1}\mathbf{v}^l \tag{5.5}$$

Besides, if $\rho(\mathbf{D}^l\mathbf{F}^l) < 1$, distributed iterative power control converges toward the optimal power solution. The proof is detailed in Appendix 5.A.
If the target SINR per user and subcarrier is unknown, (5.4) is a Bilinear Matrix Inequality (BMI), where the target SINRs are linearly constrained if power values are fixed, and vice versa. A BMI problem is non convex and can have multiple local optima [73]. Consequently, a means to reduce the complexity would be to separate target SINRs' determination from power control. In the following section, we identify an upper bound on the SINR that will be used in the OFDMA power control method detailed in Section 5.4.2.

5.3.2 An upper bound for OFDMA SINR determination

We now investigate power control in multi-cell OFDMA, when the target SINR per subcarrier is unknown and is correlated with other subcarriers' target SINR. We assume that subcarrier allocation has already been performed, and that there is no maximum transmit power constraint per subcarrier or per cell. Each user k is allocated $l_{\mathrm{SC},k}$ subcarriers by its serving cell, in set Θ_k. Let $\theta_k : \mathbb{N} \mapsto \mathbb{N}$ be the function that maps the l^{th} subcarrier allocated to user k on its absolute index, with regard to $\{1, ..., L_{\mathrm{SC}}\}$. Resource allocation aims at determining the set of target SINRs $\boldsymbol{\gamma}_k = \left[\gamma_k^{\theta_k(1)}, ..., \gamma_k^{\theta_k(l_{\mathrm{SC},k})}\right]'$ for user k. This set should be assigned so as to ensure that power control in each subcarrier is feasible, and that the target data rate of user k is reached. Moreover, resource allocation should be performed in a distributed way in each BS. Resource allocation is distributed if the BS, for any of its served users k, only requires information reported by user k and information from the medium access layer (which is, in our case, reduced to user k's target data rate). User k measures the channel gains with all surrounding BSs n in all subcarriers l, $G^l_{n,k}$, on pilot channels. The BS also has information on the background noise N_0, obtained by measurements during inactivity periods.
We have seen that, on subcarrier l, there exists a power control solution (5.4) that leads to the target SINR γ^l_k for all users k if and only if the interference matrix $\mathbf{D}^l\mathbf{F}^l$ fulfills the spectral radius property: $\rho(\mathbf{D}^l\mathbf{F}^l) < 1$.
The spectral radius of matrix $\mathbf{D}^l\mathbf{F}^l$ is lower than any submultiplicative matrix norm $\|.\|$ of $\mathbf{D}^l\mathbf{F}^l$ [72],

$$\rho(\mathbf{D}^l\mathbf{F}^l) \leq \left\|\mathbf{D}^l\mathbf{F}^l\right\| \tag{5.6}$$

We consider square matrices with real positive elements. As the set of submultiplicative matrix norms is infinite, we restrict our study to the set of induced matrix norms and to the Frobenius norm. Our aim is to determine a matrix norm that provides one criterion per BS, and this criterion should only consist of distributed information. In matrix $\mathbf{D}^l\mathbf{F}^l$, each row k is composed of $\left(\gamma_k^l \frac{G_{n,k}^l}{G_{k,k}^l}\right)$ if $n \neq k$, and 0 if $n = k$. A distributed criterion should only involve the target SINR of user k, γ_k^l. Therefore the distributed criterion must consider each row independently. The norm should consequently be: $\left\|\mathbf{D}^l\mathbf{F}^l\right\| = \max_{\{1\leq k\leq N_{\mathrm{BS}}\}} \Phi\left((\mathbf{D}^l\mathbf{F}^l)_k\right)$, where Φ is a $\mathbb{R}_+^{N_{\mathrm{BS}}} \mapsto \mathbb{R}$ function. The distributed criterion over each BS is then $\Phi\left((\mathbf{D}^l\mathbf{F}^l)_k\right) < 1$, $\forall k \in \{1, ..., N_{\mathrm{BS}}\}$.
As the Frobenius norm does not fulfill this criterion, the chosen norm is an induced norm.

Lemma 5.3.1. *An induced norm that verifies* $\left\|\mathbf{D}^l\mathbf{F}^l\right\| = max_{\{1\leq k\leq N_{BS}\}} \Phi\left((\mathbf{D}^l\mathbf{F}^l)_k\right)$, *where* Φ *is a* $\mathbb{R}_+^{N_{BS}} \mapsto \mathbb{R}$ *function, is the infinity norm, defined as:*

$$\left\|\mathbf{D}^l\mathbf{F}^l\right\|_\infty = max_{\{1\leq k\leq N_{BS}\}} \sum_{n=1}^{N_{BS}} (\mathbf{D}^l\mathbf{F}^l)_{(k,n)} \tag{5.7}$$

The proof is given in Appendix 5.B.
The infinity norm of matrix $\mathbf{D}^l\mathbf{F}^l$ is

$$\left\|\mathbf{D}^l\mathbf{F}^l\right\|_\infty = \max_{\{1\leq k\leq N_{\mathrm{BS}}\}} \left(\frac{\gamma_k^l \sum_{\{n=1,n\neq k\}}^{N_{\mathrm{BS}}} G_{n,k}^l}{G_{k,k}^l}\right) \tag{5.8}$$

Consequently, $\left\|\mathbf{D^lF^l}\right\|_\infty < 1$ if and only if

$$\left(\frac{\gamma_k^l \sum_{\{n=1,n\neq k\}}^{N_{\mathrm{BS}}} G_{n,k}^l}{G_{k,k}^l}\right) < 1, \forall k \in \{1, ..., N_{\mathrm{BS}}\} \tag{5.9}$$

If condition (5.9) is fulfilled, then $\rho(\mathbf{D}^l\mathbf{F}^l) < 1$, and distributed power control leads to the globally optimal solution. Condition (5.9) is a sufficient condition, but not a necessary condition for $\rho(\mathbf{D}^l\mathbf{F}^l) < 1$. However, the infinity norm is the only usual norm allowing distributed power control that can be used as an upper bound on the spectral radius. This distributed feature is very useful when multi-carrier transmission is considered. In that case, the target SINR per subcarrier, γ_k^l, is not known. Each user should determine its set of target SINRs, with the objective to minimize the sum transmit power. Let us define :

$$E_k^l = \frac{G_{k,k}^l}{\sum_{\{n=1,n\neq k\}}^{N_{\mathrm{BS}}} G_{n,k}^l} \tag{5.10}$$

On each subcarrier, the target SINR must fulfill condition $\gamma_k^l < E_k^l$. E_k^l is equal to the SIR if all BSs transmit with the same power. E_k^l is the most accurate information on the 'potential' SIR value available in a distributed way, prior to power control.

5.3.3 Example: 2 users, 2 subcarriers case

In this section, we detail the 2 BSs, 2 users and 2 subcarriers case. This example provides us some insight into the tightness of the previously defined upper bound. Over each subcarrier

$l = \{1,2\}$, the constraint $\Gamma_k^l \geq \gamma_k^l$ for each user $k = \{1,2\}$ is $(\mathbf{I}_2 - \mathbf{D}^l\mathbf{F}^l)\mathbf{P}^l \geq \mathbf{v}^l$, where $\mathbf{v}^l = [\frac{N_0\gamma_1^l}{G_{11}}, \frac{N_0\gamma_2^l}{G_{22}}]'$, $\mathbf{D}^l = \text{diag}\{\gamma_1^l, \gamma_2^l\}$ and $\mathbf{F}^l$ is equal to

$$\mathbf{F}^l = \begin{pmatrix} 0 & \frac{G_{21}^l}{G_{11}^l} \\ \frac{G_{12}^l}{G_{22}^l} & 0 \end{pmatrix} \tag{5.11}$$

The spectral radius of $\mathbf{D}^l\mathbf{F}^l$ is

$$\rho(\mathbf{D}^l\mathbf{F}^l) = \sqrt{\frac{\gamma_1^l\gamma_2^l G_{12}^l G_{21}^l}{G_{11}^l G_{22}^l}} \tag{5.12}$$

The infinity norm per subcarrier l is

$$\left\|\mathbf{D}^l\mathbf{F}^l\right\|_\infty = \max\left\{\frac{\gamma_1^l G_{21}^l}{G_{11}^l}, \frac{\gamma_2^l G_{12}^l}{G_{22}^l}\right\} \tag{5.13}$$

Consequently, the upper bound is equal to the spectral radius if $\frac{\gamma_1^l G_{21}^l}{G_{11}^l} = \frac{\gamma_2^l G_{12}^l}{G_{22}^l}$. The distance between the spectral radius and the infinity norm, for instance in the case $\frac{\gamma_1^l G_{21}^l}{G_{11}^l} \geq \frac{\gamma_2^l G_{12}^l}{G_{22}^l}$, is:

$$\text{d} = \frac{\gamma_1^l G_{21}^l}{G_{11}^l}\left(1 - \sqrt{\frac{\frac{\gamma_2^l G_{12}^l}{G_{22}^l}}{\frac{\gamma_1^l G_{21}^l}{G_{11}^l}}}\right)$$

The upper bound's tightness decreases when the distance between the different row elements increases. In cellular networks, this corresponds to asymmetric target SINR values (which may happen even if all users request the same data rate, due to OFDMA multi-carrier allocation), and to asymmetric radio conditions for the interfering users occupying the same subcarrier.

Remark: This result can be generalized to the N_{BS} links case. For a given subcarrier l, the following equivalence can be shown: $\rho(\mathbf{D}^l\mathbf{F}^l) = \left\|\mathbf{D}^l\mathbf{F}^l\right\|_\infty$ and there exists a positive eigenvector $\mathbf{X}$ associated with eigenvalue $\rho(\mathbf{D}^l\mathbf{F}^l)$ is equivalent to $\forall k$, $\frac{\gamma_k^l}{E_k^l} = \Delta$, where Δ is a fixed value.

Consequently, the upper bound gets closer to the spectral radius when ratio $\frac{\gamma_k^l}{E_k^l}$ has almost the same value for all users k.

5.4 Resource allocation for MA in OFDMA

In this section, we propose a resource allocation method to solve the MA problem (5.1). It uses the convergence criterion per subcarrier $\gamma_k^l < E_k^l$ when inter-cell interference is significant enough. The accuracy of using the convergence criterion is automatically assessed. Subcarrier allocation is separated from power allocation to reduce complexity. Besides, the sum power constraints $\sum_{l=1}^{L_{\text{SC}}} P_{n_{\text{BS}}}^l \leq P_{\max}, \forall n_{\text{BS}} \in \{1, ..., N_{\text{BS}}\}$ are not directly considered within power control, but an additional admission control step is used after power control. The resource allocation method per TTI consists of the following steps:

1. Subcarrier allocation: distributed in each cell.

2. Power control: composed of local power allocations within each cell, iterated over all cells in order to take inter-cell interference into account.

3. Admission control: distributed in each cell.

It is depicted on Fig. 5.1. The three steps are detailed hereunder.

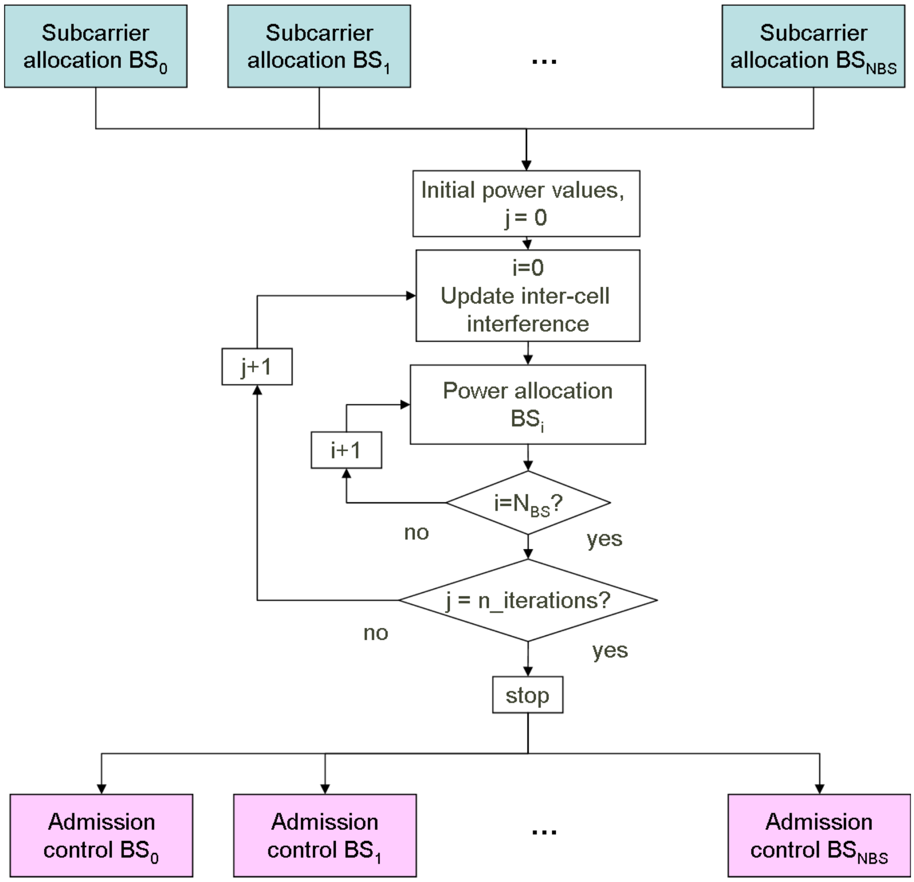


Figure 5.1: Resource allocation in multi-cell OFDMA

5.4.1 Subcarrier allocation

We propose a subcarrier allocation method derived from Rhee and Cioffi [52] for the MA problem. This method is similar to the one used for non-coordinated RC users in Chapter 4 (see Section 4.3.2). It assumes that all users request the same target data rate, $R_{k,\text{target}} = R_{\text{target}}, \forall k$. In case of various target data rates, this method can be performed per set of users with the same target data rate, the different sets being ordered depending on their priority. For instance, a possible strategy may be to assign subcarriers to users by increasing order of the target data rates.

In the method from [52] and from Section 4.3.2, the objective of subcarrier selection is to maximize the direct channel gain. We propose here to use a modified version where the objective is to maximize criterion E (5.10). In the following algorithm, we refer to the coefficient to be optimized as U_k^l. If the direct channel gain criterion is used, then $U_k^l = G_{k,k}^l$. If criterion E is used, then $U_k^l = E_k^l$.

1. Initialization:

(a) Set the approximate data rate to $\tilde{R}_k = 0$ for all users k.

(b) For each user k, find the subcarrier l that maximizes coefficient U_k^l, and assign it to user k. Update $\tilde{R}_k$ with the approximate data rate obtained on subcarrier l.

2. Iterative phase : as long as there are free subcarriers and as long as at least one user does not reach the target data rate R_{target},

(a) Find the user k that has the lowest data rate $\tilde{R}_k$.

(b) For that user k, find the free subcarrier l with the highest coefficient U_k^l, and assign it to user k.

(c) Update $\tilde{R}_k$ with the additional approximate data rate obtained on subcarrier l.

The approximate data rate per subcarrier, $\tilde{R}_k^l = B_{\text{SC}} \log_2 \left(1 + \frac{G_{k,k}^l P_{max}}{L_{\text{SC}} \tilde{I}}\right)$, is computed assuming equal power allocation and equal inter-cell interference level on all subcarriers. $\tilde{I}$ is an average value for the interference plus noise, that depends on the load.
We propose to use criterion E in subcarrier allocation, so as to maximize the probability that E_k^l has a high value for all (k, l). Indeed, this will then allow the target SINRs to have a higher range of variation, when the convergence criterion is used.

5.4.2 Distributed power control

Iterative water-filling [61] performs power allocation on each BS independently, considering inter-cell interference as noise. The process is iterated over all BSs. The main problem of iterative water-filling is that it does not necessarily converge. Indeed, power control over some subcarriers may not be feasible, which means that problem (5.4) may not have any positive solution. To overcome this limitation, we propose to adaptively use the convergence criterion $\gamma_k^l < E_k^l$ within power allocation.

Adaptive convergence criterion

The convergence criterion should not be used when the interference channel is noise-limited. It is thus necessary to evaluate the inter-cell interference level on each subcarrier. Indeed, criterion E may be too restrictive when inter-cell interference is insignificant. In a worst-case scenario, a user k could have its SINR on subcarrier l limited to E_k^l, although none of the interfering users are transmitting. As a consequence, we add a criterion to trigger the use of E_k^l per subcarrier, depending on the inter-cell interference level.
Let $I_k^l = N_0 + \sum_{\{n=1, n\neq k\}}^{N_{\text{BS}}} G_{n,k}^l P_n^l$ be the inter-cell interference plus noise received in a given power control iteration. The transmitter for user k computes

$$\tau_k^l = \left(\frac{I_k^l}{G_{k,k}^l P_{\max}}\right) - \left(\frac{N_0}{G_{k,k}^l P_{\max}}\right) = \frac{\sum_{\{n=1, n\neq k\}}^{N_{\text{BS}}} G_{n,k}^l P_n^l}{G_{k,k}^l P_{\max}} \tag{5.14}$$

τ_k^l provides an estimate for the inter-cell interference level. As $P_{\max} \geq P_n^l \geq 0, \forall n \in \{1, ..., N_{\mathrm{BS}}\}$, τ_k^l is upper-bounded by $\frac{1}{E_k^l}$,

$$\frac{1}{E_k^l} \geq \tau_k^l \geq 0 \tag{5.15}$$

The transmitter assumes that inter-cell interference can be neglected at the next iteration if τ_k^l is low with regard to $\frac{1}{E_k^l}$, i.e., if $E_k^l \tau_k^l < \delta$. δ is a parameter that should be set depending on the network characteristics.

Power allocation

Let us consider power allocation for BS n_{BS} at a given iteration of the power control process. The inter-cell interference levels are set to the values corresponding to the interfering powers from the previous iteration. As the sum power constraint is not directly considered within power allocation for n_{BS}, solving the MA problem with fixed inter-cell interference levels is equivalent to solving the power allocation problem for each user of n_{BS} independently. Let k be a user of n_{BS}. The objective for user k is to minimize the power required to reach its target data rate, summed over its $l_{\mathrm{SC},k}$ allocated subcarriers of set Θ_k. The power required to satisfy user k in subcarrier l is $P_k^l = \frac{\gamma_k^l}{G_{k,k}^l} I_k^l$.
Consequently, the problem of minimizing the sum power for user k is equivalent to the following target SINRs allocation problem:

$$\begin{aligned}
&\min_{\boldsymbol{\gamma}_k} \sum_{l \in \Theta_k} \gamma_k^l \left(\frac{I_k^l}{G_{k,k}^l} \right) \\
&\text{s. t. } B_{\mathrm{SC}} \sum_{l \in \Theta_k} \log_2(1 + \gamma_k^l) \geq R_{k,\mathrm{target}} \\
&\text{s. t. } \gamma_k^l \geq 0, \forall l \in \Theta_k \\
&\text{s. t. } E_k^l - \epsilon \geq \gamma_k^l \text{ if } E_k^l \tau_k^l \geq \delta, \forall l \in \Theta_k
\end{aligned} \tag{5.16}$$

ϵ is a small, strictly positive value that is used to avoid $E_k^l = \gamma_k^l$ in all cases. (5.16) is a convex optimization problem that can be solved with the KKT conditions. The solution of resource allocation (5.16) for user k is:

- If $E_k^l \tau_k^l \geq \delta$:

$$\gamma_k^l = \min \left\{ \left[\lambda \left(\frac{G_{k,k}^l}{I_k^l} \right) - 1 \right]^+ ; E_k^l - \epsilon \right\} \tag{5.17}$$

- If $E_k^l \tau_k^l < \delta$:

$$\gamma_k^l = \left[\lambda \left(\frac{G_{k,k}^l}{I_k^l} \right) - 1 \right]^+ \tag{5.18}$$

Where λ should fulfill the data rate condition $\sum_{l \in \Theta_k} B_{\mathrm{SC}} \log_2(1 + \gamma_k^l) = R_{k,\mathrm{target}}$. λ is obtained by bisection search. The details are provided in Appendix 5.C.

Iterative power control

Per BS power allocation considers the actual inter-cell interference in the determination of the target SINR values. Therefore, power control is performed according to the following iterative process:

1. Initialization: set all power values to 0.
2. Target SINRs determination iteration: for each user k, given the power values obtained in the previous iteration, compute the set of target SINRs γ_k on its allocated subcarriers.
3. On each subcarrier, determine the power level per user required to reach the target SINR.
4. Update interference values and go to the next target SINRs determination iteration.

It should be noted that if inter-cell interference is insignificant on all subcarriers, i.e., if condition $E_k^l \tau_k^l < \delta$ is fulfilled for all users, on all subcarriers and during all the iterations, then our proposed algorithm is simply iterative water-filling.

5.4.3 Admission control

In the previous algorithm, we did not consider the maximum sum power constraint per BS. Thanks to this simplification, each user can be independently handled in power allocation. However, at medium to high load, the sum power constraint may not hold. A simple admission control process is then used: if the sum power exceeds the maximum $P_{\max}$ on n_{BS}, the users served by n_{BS} are ordered by descending transmit power value. Then the users with highest power are rejected from the cell, until the sum power becomes less than $P_{\max}$. This is not a fair admission control process, but it is the most efficient to reject as little users as possible.

5.5 Numerical results

We consider a network $\mathcal{N}$ composed of two rings of interfering BSs with omnidirectional antennas and same cell radius. Numerical results are obtained via Monte-Carlo simulations. The inter-site distance is $\mathrm{d}_{\mathrm{is}} = 1.212$ km. The path loss model is Okumura-Hata with a log-normal shadowing with standard deviation 7 dB, and the fast fading is Rayleigh. The maximum transmit power for each BS is $P_{\max} = 43$ dBm. The total bandwidth is $B = 10$ MHz, the FFT size and the number of available subcarriers for data transmission is $L_{\mathrm{SC}} = 256$.
The proposed method is compared with iterative water-filling, with two subcarrier allocation criteria: the convergence criterion E, or the channel gain criterion G. Two scenarios, corresponding to different noise levels, are studied, so as to assess the accuracy of adaptive convergence criterion triggering.
In the iterative water-filling process used for comparisons with our method, an additional limitation on power control is used, in order to avoid the situations where the power values of some users importantly diverge. If the power of a user exceeds $P_{\max}$ after power allocation, then this user's power is set to zero.

5.5.1 High noise level

Table 5.1: Percentage of rejected users depending on the load (%), $R_{\text{Target}} = 64$ kbits/s, high noise level

N_{users}	< 128	128	160	192	224
Proposed method	0	1.88	7.58	14.82	20.02
Iterative water-filling, E	0	3.91	10.71	27.28	34.78
Iterative water-filling, G	0	4.78	18.42	30.76	38.43

In this section, the noise is equal $N_0 = -105$ dBm. It is composed of the thermal noise, and of an additional background noise. All users require a minimum data rate of 64 kbits/s. Users that do not reach their target data rate are rejected.

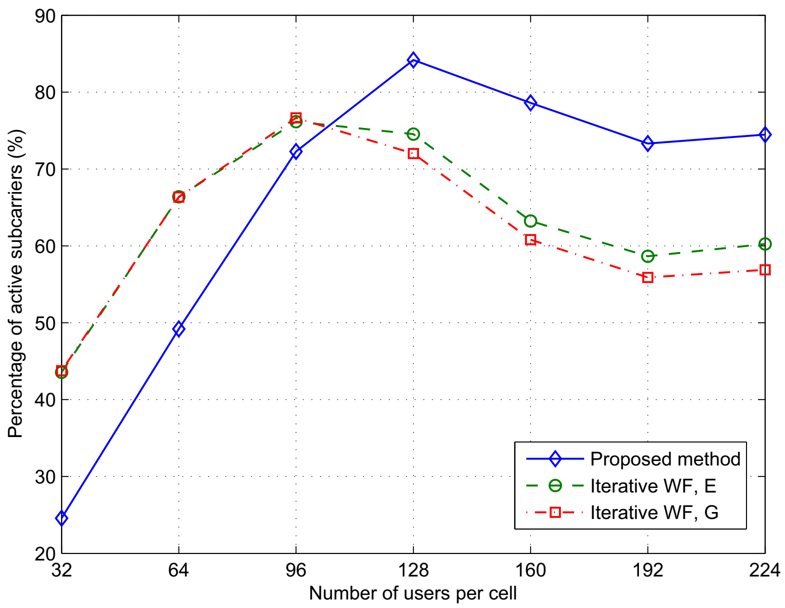


Figure 5.2: Percentage of active subcarriers, for varying load, $R_{\text{Target}} = 64$ kbits/s, high noise level

First, we should notice that in this high noise scenario, the percentage of users and subcarriers that do not use criterion E in the proposed method due to low inter-cell interference level is lower than 2% when the load exceeds 128 users per cell, with $\delta = 10^{-3}$. Inter-cell interference is always significant due to the high power values required to reach the target data rates. In the following, we represent the results obtained with $\delta = 0$, when criterion E is always used.

Fig. 5.2 shows the percentage of active subcarriers, defined as the percentage of allocated subcarriers where the power is different from zero. At low to medium load, the proposed

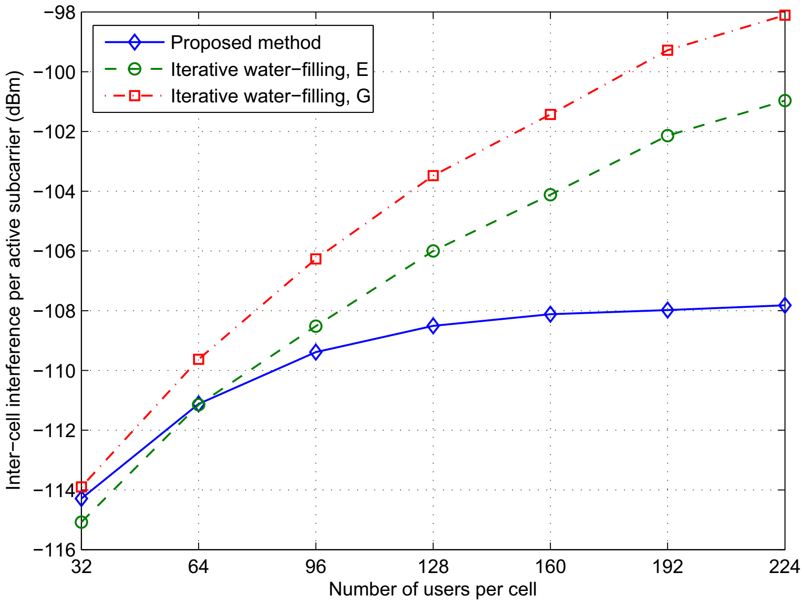


Figure 5.3: Interference per active subcarrier, for varying load, $R_{\text{Target}} = 64$ kbits/s, high noise level

method is more efficient than iterative water-filling regarding subcarriers consumption. Indeed, the proposed method determines the set of target SINRs iteratively, depending on the per subcarrier inter-cell interference plus noise from the previous iteration. Therefore, higher target SINRs are set in the least interfered subcarriers, whereas the most interfered subcarriers are not used. Then at medium to high load, iterative water-filling leads to power divergence situations that result in an increase of the percentage of rejected users (Table 5.1), and in a consequential decrease in the percentage of active subcarriers. The percentage of rejected users is almost divided by two with the proposed method. The average interference per active subcarrier is computed before users are rejected by admission control. It keeps increasing with the load (Fig. 5.3) with iterative water-filling, whereas the proposed method leads to an asymptote. The Cumulative Density Functions (CDF) of the inter-cell interference per active subcarrier are represented on Fig. 5.4, with 128 and 224 users per cell. The proposed method not only decreases inter-cell interference, but also improves the fairness among users, compared to iterative water-filling.

It should be noted that using the E criterion instead of the gain criterion for subcarrier allocation prior to iterative water-filling decreases the average inter-cell interference, but still leads to power divergence situations. Criterion E also increases fairness among users, especially at high load.

If δ is set to 10^{-3} instead of 0, the rejection rate slightly increases by up to 2% at high load, and the percentage of active subcarriers increases at low load. Indeed at low load, the proposed method is then equivalent to iterative water-filling in most cases, and thus does not lead to any resource consumption gain.

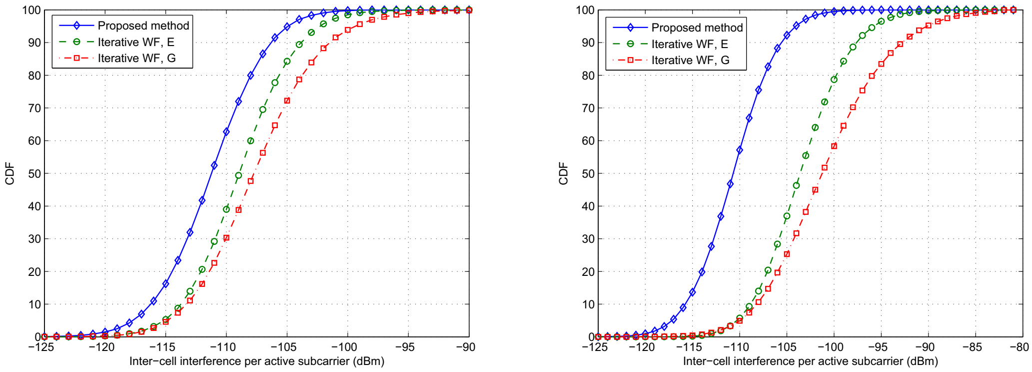


Figure 5.4: CDF of inter-cell interference per active subcarrier, $R_{\text{Target}} = 64$ kbits/s, 128 (left) and 224 (right) users per cell

5.5.2 Low noise level

In this section, only the thermal noise with spectral density equal to -174 dBm/Hz is considered. All users require a minimum data rate of 128 kbits/s. Users that do not reach their target data rate are rejected. We evaluate the influence of δ parameter, by testing $\delta \in \left\{10^{-3}, 5.10^{-3}, 10^{-2}\right\}$.

Table 5.2: Cases when E criterion is not relevant (%), $R_{\text{Target}} = 128$ kbits/s, low noise level

N_{users}	96	128	160	192	224
$\%E$ not used, $\delta = 10^{-3}$	99.6	95.3	63.8	48.9	48.9
$\%E$ not used, $\delta = 5.10^{-3}$	100	99.8	86.1	56.6	54.4
$\%E$ not used, $\delta = 10^{-2}$	100	100	93.0	58.5	55.9

The probability of not using criterion E is shown in Table 5.2. Criterion E is almost never triggered when the load is lower than 96 users per cell. It is used in 51% of the cases at most, at high load. The percentage of rejected users is represented on Fig. 5.5. With iterative water-filling, it increases very rapidly when the load exceeds 160 users per cell, with a maximum of 69% of rejected users when subcarrier allocation is performed with criterion G. The performance results of the proposed method depend on δ. If δ is set to 10^{-3}, criterion E is triggered at medium load before power divergence becomes an issue, leading to a rejection rate of 6.5% for 160 users per cell. At this load however, users are not rejected yet by iterative water-filling, if criterion E is used for subcarrier allocation. Consequently $\delta = 10^{-3}$ is too restrictive in that case, and leads to useless users' rejections. On the contrary, if δ is set to 10^{-2}, criterion E is not triggered often enough, and the rejection rate at high load reaches 34.3%. $\delta = 5.10^{-3}$ is a good trade-off, that avoids useless rejections at medium load, as well as power divergence situations at high load. With this value, the rejection rate is less than 25% at any load. It should be noted that we only need to consider criterion E in 45.5% of the cases, and yet a gain of 44% in rejection rate is

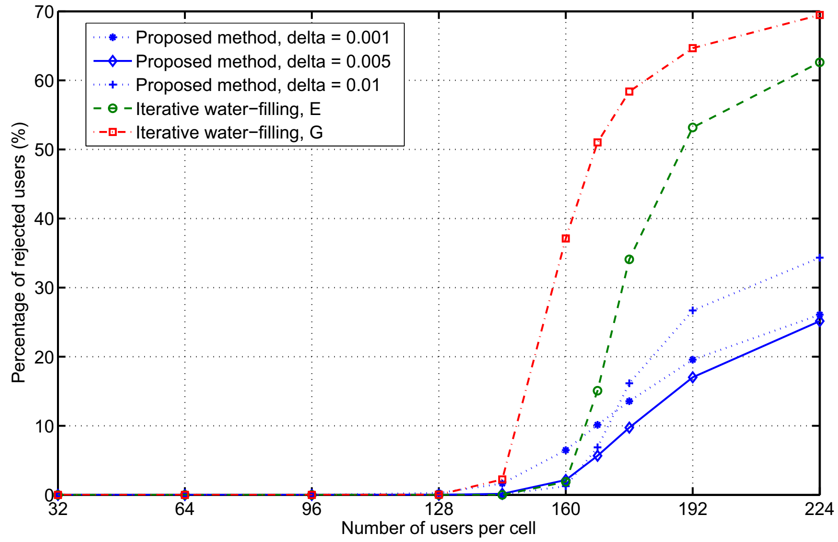


Figure 5.5: Percentage of rejected users, for varying load, $R_{\text{Target}} = 128$ kbits/s, low noise level

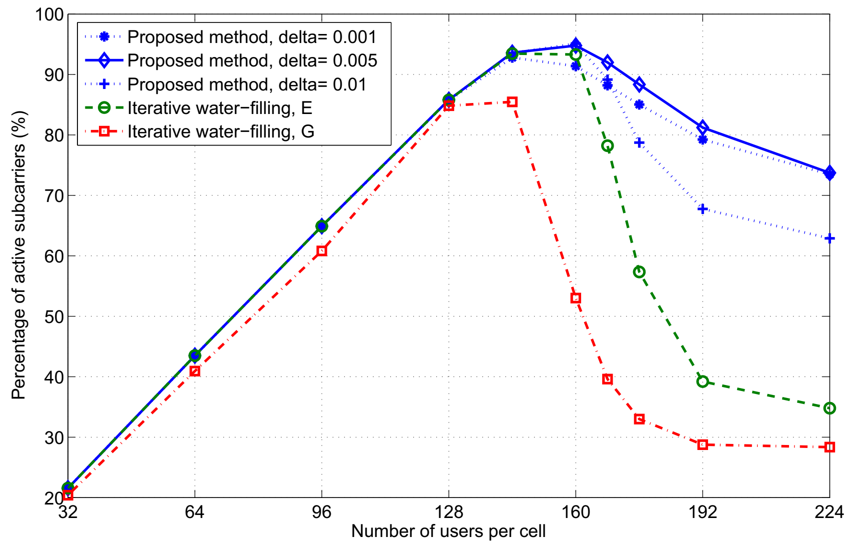


Figure 5.6: Percentage of active subcarriers, for varying load, $R_{\text{Target}} = 128$ kbits/s, low noise level

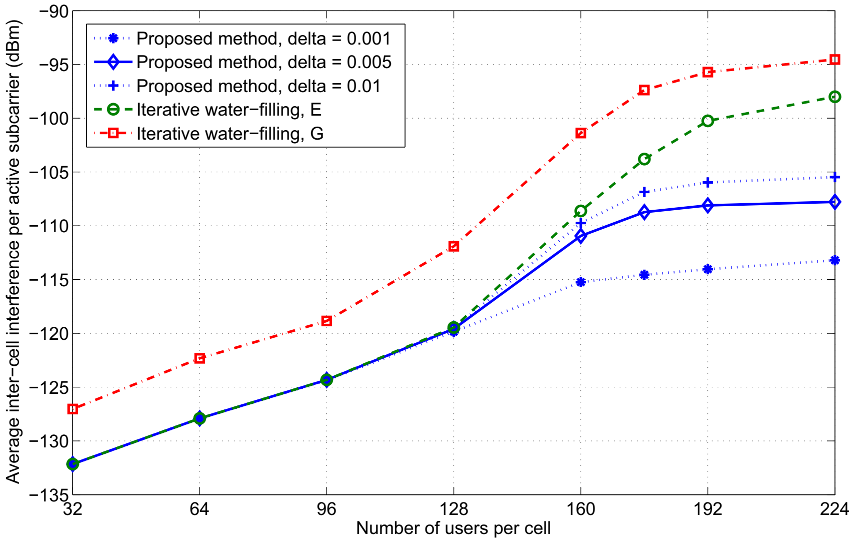


Figure 5.7: Interference per active subcarrier, for varying load, $R_{\text{Target}} = 128$ kbits/s, low noise level

obtained. This shows that the main issue is to avoid that the power of some users diverges, which results in many other users being affected -and in turn increasing their power levels without bound. The gain brought by our proposed method is higher here than in the high noise scenario, where the noise level mitigates power divergence's influence. The percentage of active subcarriers is represented on Fig. 5.6. At low load, the proposed method is equivalent to iterative water-filling with criterion E for subcarrier allocation. With iterative water-filling, the percentage of active subcarriers rapidly decreases, following the inter-cell interference increase that leads to users' rejection (see Fig. 5.7).
Finally, we can notice that the influence of the criterion chosen for subcarrier allocation remains quite important. Maximizing E instead of the direct channel gain G in subcarrier allocation decreases the rejection rate and the inter-cell interference level at any load.

5.6 Conclusion

In this chapter, we have studied resource allocation for RC users in SISO multi-cell OFDMA networks. A distributed convergence criterion for power control on the interference channel has been obtained. Based on this criterion, we have proposed a method that determines the target SINR per user and per subcarrier, which ensures that the distributed power control over each subcarrier is feasible and converges toward the global optimum, while fulfilling the data rate requirement of each user. The convergence criterion is used on the subcarriers where inter-cell interference is the limiting feature, in an adaptive way. This method requires an iterative implementation and may be performed within each TTI, assuming that iterations for a fixed channel state run over the time slots. Numerical results show that the proposed method avoids power divergence situations and is very efficient when compared to iterative water-filling, at medium to high load. It is also shown that maximizing the convergence coefficient E within subcarrier allocation is far more efficient than maximizing the channel gain coefficient. Finally, our proposed method accurately scales with the inter-cell interference level and is thus valid at any load. It is consequently a relevant alternative to iterative water-filling in distributed networks.

We have seen in this chapter that MA resource allocation in multi-cell, multi-carrier SISO communication systems, when full CSI is available at the transmitter, can be solved without making any simplifying assumption. However, this conclusion is no longer valid for MIMO point-to-point transmissions. Indeed, power allocation should then involve the different MIMO streams, but the streams' diagonalization bases differ from one interfering cell to the other. Besides, retrieving all CSI at the transmitter involves far more signaling than in SISO, so we should also consider the cases where full CSI is not available at the transmitter. We will investigate these issues in the following chapter.

APPENDIX

5.A Proof of power control convergence constraint on the interference channel

Let us consider distributed power control in subcarrier l, where the target SINR per user k, γ_k^l, is known. The following power value is set for user k by its serving BS:

$$P_k^l = \frac{\gamma_k^l}{G_{k,k}^l}\left(N_0 + \sum_{\{n=1,n\neq k\}}^{N_{\mathrm{BS}}} P_n^l G_{n,k}^l\right) \tag{5.19}$$

Let $\mathbf{P}_{(0)}^l = \left[P_{1,(0)}^l, ..., P_{N_{\mathrm{BS}},(0)}^l\right]'$ be the vector of initial power values. At step T, power is updated according to the inter-cell interference measured in the previous step, $T-1$. Therefore, $\mathbf{P}_{(T)}^l = \left[P_{1,(T)}^l, ..., P_{N_{\mathrm{BS}},(T)}^l\right]'$ can be written as:

$$\mathbf{P}_{(T)}^l = (\mathbf{D}^l\mathbf{F}^l)\mathbf{P}_{(T-1)}^l + \mathbf{v}^l \tag{5.20}$$

where $\mathbf{v}^l$ and $\mathbf{D}^l\mathbf{F}^l$ have been previously defined.

Iterating over $\mathbf{P}^l$ consequently leads to:

$$\mathbf{P}_{(T)}^l = (\mathbf{D}^l\mathbf{F}^l)^T\mathbf{P}_{(0)}^l + \left(\sum_{t=0}^{T-1}(\mathbf{D}^l\mathbf{F}^l)^t\right)\mathbf{v}^l \tag{5.21}$$

From [72], this Neuman series converges toward $(\mathbf{I}_{N_{\mathrm{BS}}} - \mathbf{D}^l\mathbf{F}^l)^{-1}\mathbf{v}^l$ if and only if $\rho(\mathbf{D}^l\mathbf{F}^l) < 1$, and, from the Perron-Frobenius theorem, if $\mathbf{D}^l\mathbf{F}^l$ is nonnegative and irreducible, then $(\mathbf{I}_{N_{\mathrm{BS}}} - \mathbf{D}^l\mathbf{F}^l)^{-1}\mathbf{v}^l > 0$.

As a consequence, if $\rho(\mathbf{D}^l\mathbf{F}^l) < 1$, then $\mathbf{P}_{(k)}^l$ has a strictly positive limit:

$$\lim_{T\to+\infty} \mathbf{P}_{(T)}^l = (\mathbf{I}_{N_{\mathrm{BS}}} - \mathbf{D}^l\mathbf{F}^l)^{-1}\mathbf{v}^l \tag{5.22}$$

Therefore, the iterative, distributed power control leads to the globally optimal solution if and only if the system is feasible.

5.B Proof of Lemma 5.3.1

Our aim is to obtain an induced norm whose original vector norm is defined as $\|\mathbf{x}\| = \max_{\{1\leq k\leq N_{\mathrm{BS}}\}} f(x_k)$, where f is a $\mathbb{R}_+ \mapsto \mathbb{R}$ function. The vector norm should be positive definite. Consequently $f(0) = 0$, and f must be an increasing function. The positive homogeneity criterion states that for any α in $\mathbb{R}_+$, and any x, $\|\alpha\mathbf{x}\| = \alpha\|\mathbf{x}\|$. Consequently $f(\alpha x_k) = \alpha f(x_k)$ for any α and x_k in $(\mathbb{R}_+, \mathbb{R}_+)$. Let us assume that f is twice differentiable in $\mathbb{R}_+$. This assumption is quite reasonable as most norm functions are polynomials. Deriving once with regard to x_k and then once with regard to α provides: $x_k f''(\alpha x_k) = 0$ for any (α, x_k). Consequently $f''(x_k) = 0, \forall x_k$, and f is a linear function. As $f(0) = 0$, f is: $f(x_k) = \beta x_k$ where $\beta \in \mathbb{R}_+$.

The induced matrix norm is equal to:

$$\left\|\mathbf{D}^l\mathbf{F}^l\right\| = \max\left\{\frac{\left\|\mathbf{D}^l\mathbf{F}^l\mathbf{x}\right\|}{\|\mathbf{x}\|} : \mathbf{x} \in \mathbb{R}_+^{N_{\mathrm{BS}}}, \mathbf{x} \neq \mathbf{0}\right\}$$

As constant β appears in both numerator and denominator, the induced matrix norm is:

$$\left\|\mathbf{D}^l\mathbf{F}^l\right\| = \max_{\{1\leq k\leq N_{\mathrm{BS}}\}} \sum_{n=1}^{N_{\mathrm{BS}}} (\mathbf{D}^l\mathbf{F}^l)_{(k,n)}$$

This norm is defined as the infinity norm for matrices in $\mathbb{R}_{+}^{N_{\mathrm{BS}}}$.

5.C Solution of power allocation (5.16)

The Lagrangian of problem (5.16) is:

$$\begin{aligned} L(\boldsymbol{\gamma}_k, \lambda, \boldsymbol{\alpha}, \boldsymbol{\beta}) = & \sum_{l\in\Theta_k} \gamma_k^l \left(\frac{I_k^l}{G_{k,k}^l}\right) + \frac{\lambda \log(2)}{B_{\mathrm{SC}}} \left(R_{k,\mathrm{target}} - \sum_{l\in\Theta_k} \log_2(1+\gamma_k^l)\right) \\ & - \sum_{l\in\Theta_k} \alpha^l \gamma_k^l + \sum_{l\in\Theta_k} \beta^l (\tau_k^l)^+ \left(\gamma_k^l - E_k^l + \epsilon\right) \end{aligned} \tag{5.23}$$

Where $\lambda, \boldsymbol{\alpha}, \boldsymbol{\beta}$ are Lagrange multipliers, and ϵ is a small value that is used to avoid $E_k^l = \gamma_k^l$ in all cases.
$\beta^l(\tau_k^l)^+$ is set to 0 if $E_k^l\tau_k^l < \delta$, and is a positive value if $E_k^l\tau_k^l \geq \delta$.

$$\frac{\partial L(\boldsymbol{\gamma}_k, \lambda, \boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \gamma_k^l} = \frac{I_k^l}{G_{k,k}^l} - \frac{\lambda}{(1+\gamma_k^l)} - \alpha^l + \beta^l(\tau_k^l)^+ \tag{5.24}$$

The KKT conditions impose that $\frac{\partial L(\boldsymbol{\gamma}_k, \lambda, \boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \gamma_k^l} = 0$. α^l is a slack variable that can be set to zero [24]. From the KKT condition $\beta^l(\tau_k^l)^+ \geq 0$ and equation (5.24), we get :

$$\lambda \geq \frac{(1+\gamma_k^l)I_k^l}{G_{k,k}^l} \tag{5.25}$$

Another KKT condition is $\beta^l(\tau_k^l)^+\left(\gamma_k^l - E_k^l + \epsilon\right) = 0$. If $\beta^l(\tau_k^l)^+ > 0$, then $\left(\gamma_k^l - E_k^l + \epsilon\right) = 0$. In that case, the solution is $\gamma_k^l = E_k^l - \epsilon$.
If $\beta^l(\tau_k^l)^+ = 0$, then $\lambda = \frac{(1+\gamma_k^l)I_k^l}{G_{k,k}^l}$, and the solution is $\gamma_k^l = \left[\lambda\left(\frac{G_{k,k}^l}{I_k^l}\right) - 1\right]^+$, where the positivity constraint comes from equation (5.25).
Consequently, the solution is:

- If $E_k^l\tau_k^l \geq \delta$:

$$\gamma_k^l = \min\left\{\left[\lambda\left(\frac{G_{k,k}^l}{I_k^l}\right) - 1\right]^+; E_k^l - \epsilon\right\} \tag{5.26}$$

- If $E_k^l\tau_k^l < \delta$:

$$\gamma_k^l = \left[\lambda\left(\frac{G_{k,k}^l}{I_k^l}\right) - 1\right]^+ \tag{5.27}$$

Where λ is defined so as to fulfill the data rate condition $\sum_{l\in\Theta_k} B_{\mathrm{SC}} \log_2(1+\gamma_k^l) = R_{k,\mathrm{target}}$.

Chapter 6

Distributed resource allocation in MIMO networks

6.1 Introduction

This chapter extends the results from Chapter 5 on distributed MA resource allocation for RC users to the zero-mean independent and identically distributed (i.i.d.) MIMO channel. Each transmitter is equipped with n_t transmit antennas, and each receiver is equipped with n_r receive antennas. MIMO transmission increases the point-to-point data rate thanks to the additional degrees of freedom of the channel, compared to SISO transmission. The target data rate of RC users can consequently be reached with fewer power consumption. Chapter 5 has shown that inter-cell interference can efficiently be mitigated for RC users in distributed networks, by considering a power control convergence constraint.

Resource allocation in MIMO OFDMA depends on whether CSI is available at the transmitter (CSIT). The transmitter has full CSI if it knows all fast fading coefficients from each transmit antenna to each receive antenna, on top of the statistical channel information, path loss and shadowing. In that case, it is necessary to determine the type of precoding and detection required to make the best use of the channel, while managing inter-cell interference, and limiting the complexity of the treatments. For that purpose, we only consider linear processing, at transmitter and receiver. The linear transmitter should result from an optimization of power allocation on both subcarriers and streams.

If only the statistical properties of the channel are known at transmitter, the outage probability and the corresponding outage capacity are used in the MA resource allocation problem. The main issue is then to obtain an approximate analytical expression of the outage capacity. This approximation should enable us to solve power allocation efficiently in each iteration. We therefore aim at obtaining the outage capacity as a concave function of the BS power, in order to solve the power allocation problem with convex optimization techniques.

The chapter first details the state of the art on both topics. Specifically, we review the methods proposed for resource allocation in the full CSIT case with linear processing, and detail the analytical expressions of the outage capacity, whether exact or approximate, that have been obtained in the literature up to now. The general framework for MA resource allocation is described in Section 6.3. The same decomposition method of the original MA problem is used as in the previous chapters, namely, separation of subcarrier and power allocations, and decomposition of resource allocation per BS, with iterative power control over BSs. The adapted method for power allocation, depending on the availability of CSIT, is then derived. In Section 6.4, the MA resource allocation problem with full CSIT is treated. Power optimization is performed on all subcarriers and streams. Similarly to the SISO case, a distributed convergence criterion E is determined per user and subcarrier. It is an approximate upper bound, as it is not possible to diagonalize all interfering channels in the same basis. Criterion E is included in power control, as an upper bound on the achievable data rate per subcarrier.

In Section 6.5, the MA resource allocation problem is studied, when only the statistical properties of the channel are known at transmission. An approximate analytical expression of the outage capacity is derived. It is tested with several antenna configurations, and is shown to be very close to the outage capacity obtained with Monte-Carlo simulations, at practical outage and SNR values. It expresses the outage capacity for a fixed outage probability as a concave function of the SNR, that can be used to solve the MA problem as a convex optimization problem. A convergence criterion E is derived for that problem, and included into iterative power control. The proposed resource allocation methods are assessed via numerical results and compared with iterative water-filling in Section 6.6.

The main contributions of this chapter are:

- A complete resource allocation method for the MA problem with MIMO full CSIT is determined. Power allocation performs iterative water-filling over the streams of the equivalent MIMO channel with a per-subcarrier convergence constraint E. The distributed convergence constraint involves the highest singular value of the direct channel, and is adaptively triggered, depending on the inter-cell interference level.

- An approximate analytical expression of the outage capacity, as a function of the outage probability and of the SNR, is provided. The arithmetic mean of an upper bound and of a lower bound on the outage capacity proves to be very close to the outage capacity obtained via Monte-Carlo simulations, for various values of the outage probability, and various antenna configurations. The analytical expression is concave in the SNR for a fixed outage probability, thus turning the MA problem with MIMO statistical CSIT into a convex optimization problem. A resource allocation method for the MA problem with MIMO statistical CSIT is derived, based on the analytical outage capacity expression. A distributed convergence criterion E is determined for power control.

- Both methods are shown to be far more efficient than iterative water-filling to solve the MA problem at medium to high load. Power divergence situations are avoided thanks to the defined E convergence criterion. These results complete the ones obtained in Chapter 5 for SISO, thus assessing the feasibility of distributed resource allocation for RC users in all studied cases.

6.2 State of the art

6.2.1 Resource allocation with full CSIT and linear processing

In the single-channel case, distributed joint power control and linear processing in MIMO has been studied with various optimization objectives and system assumptions. In many papers, an iterative method, that is built on network duality, is used. Network duality [74] is based on the fact that for each channel, there exists a reciprocal channel, where the roles of the transmitters and receivers are switched. The reciprocal channel matrix is equal to the Hermitian of the original channel matrix. Distributed resource allocation thus first optimizes the direct channel parameters with fixed reciprocal channel parameters, and then optimizes the reciprocal channel parameters with fixed direct channel parameters. At each stage, the optimal receive matrix is determined for a fixed power allocation, and then the optimal power allocation is updated, depending on the receive matrix. In the following, we review the most relevant papers dealing with resource allocation in MIMO. It should be noted that they all concern single-channel transmission.
The single-cell case with inter-stream interference is studied in [75] [76]. The authors consider the following problems: RA and MA in [75], maximization of the weighted rate, summed over all streams, under a total power constraint, and its dual, minimization of the total power under a weighted sum rate requirement, in [76]. In both papers, it is shown that the SINR per stream is maximized by using a MMSE receiver. Then the Minimum Mean Square Error is equal to the inverse of $(1 + \mathrm{SINR})$. Inserting the MMSE instead of the SINR in the rate expression, for fixed MMSE filters, turns the power allocation problem into a geometric programming problem, that has a unique global optimum. An iterative process updating power levels and the MMSE filters is used, at receiver's side, and then at transmitter's side. The iterative process is shown to converge to a local optimum with the four optimization objectives.
Multi-cell networks with beamforming are studied for the uplink transmission without user multiplexing in [77]. The signal obtained at the receiver is a weighted sum of the received signals at each antenna, and the objective is to determine the optimum weights to solve the RA and MA problems. For both problems, the authors propose an iterative, network duality-based method. The iterative process for the MA problem does not converge if the set of target SINRs is not feasible. These approaches are extended to the multi-cell, multi-stream case in [78].
Distributed interference alignment for MIMO is presented in [79]. It consists in building interference-free spatial channels for direct transmission via iterative linear processing at transmitter and receiver. At each iterative step, the receive vector that aligns the interference in the null space is determined. The useful signal is then received through a full rank matrix, while interference is completely suppressed. The rank of the matrix corresponding to the useful signal is equal to the generalized degrees of freedom for the user's data. This process is performed at receiver, and then at transmitter using network duality. It is shown to converge, possibly to a local minimum of the interference, due to the non-convexity of the problem. Interference alignment may be performed in space, frequency [80] or time [81]. Its main limitation is that, in order to offer a minimum generalized degree of freedom to each user, the required bandwidth must grow with the number of users. Our work cannot

directly be seen under the scope of interference alignment. However, the proposed power allocation methods opportunistically make use of the least interfered streams and subcarriers, and prevent transmission within highly interfered subcarriers.
Besides, we do not use iterative distributed linear processing based on network duality for complexity purposes. The receive filter is MMSE, and the optimization variables of the resource allocation problem are the transmit power and the precoder matrix.

6.2.2 Analytical expressions of the outage capacity

If only the statistical properties of the channel are known at transmitter, resource allocation is performed on the outage capacity. Numerical evaluations of the outage capacity and probability are feasible for a fixed SNR, either via Monte-Carlo snapshots, or via numerical approximations. It is far more complicated to extract an analytical expression, even based on an approximation, of the SNR required to reach a target outage capacity with a target outage probability. However, this problem has to be solved in order to perform resource allocation.
For the zero-mean i.i.d. MIMO channel, this problem has only been treated under high SNR assumption, and for specific cases: when the minimum number of transmit and receive antennas is equal to 1 [82], and for the per stream outage probability [83]. The distribution of the mutual information was shown to be Gaussian in asymptotic cases, when the number of antennas become large [84]. The Gaussian approximation is even quite close to the mutual information's distribution in the general case [85]. However, the mean and variance of the Gaussian distribution are integrals in the SNR, and their formulas cannot be analytically inverted. The details on the approximations that have been proposed in the literature are provided in Appendix 6.A.

6.3 General framework for Margin Adaptive resource allocation

Let $\mathcal{N}$ be a network composed of N_{BS} base stations equipped with n_t antennas, and $K_{\mathcal{N}}$ users equipped with n_r antennas. $n_{\min} = \min(n_t, n_r)$ is the minimum of the number of transmit and receive antennas. Each BS transmits over L_{SC} orthogonal subcarriers. Θ_k is the set of subcarriers assigned to user k by its BS. Let $R_{k,\mathrm{target}}$ be the target data rate of user k, R_k^l the data rate of user k in subcarrier l, and P_k^l the power transmitted to user k by its serving BS in subcarrier l. We consider the following MA optimization problem in multi-cell OFDMA :

$$
\begin{aligned}
&\min_{\{\mathbf{P},\boldsymbol{\Theta}\}} \sum_{k=1}^{K_{\mathcal{N}}} \sum_{l\in\Theta_k} P_k^l \\
&\text{s. t.} \sum_{l\in\Theta_k} R_k^l \geq R_{k,\text{target}}, \forall k \in \{1,...,K_{\mathcal{N}}\} \\
&\text{s. t.} \sum_{l=1}^{L_{\text{SC}}} P_{n_{\text{BS}}}^l \leq P_{\max}, \forall n_{\text{BS}} \in \{1,...,N_{\text{BS}}\} \\
&\text{s. t.} \sum_{l=1}^{L_{\text{SC}}} P_{n_{\text{BS}}}^l \geq 0, \forall (n_{\text{BS}},l) \in \{1,...,N_{\text{BS}}\} \times \{1,...,L_{\text{SC}}\} \\
&\text{s.t. } \Theta_k \cap \Theta_{k'} = \emptyset, \forall (k,k') \text{ served by the same BS}, k \neq k'
\end{aligned}
\tag{6.1}
$$

Problem (6.1) is similar to the SISO MA resource allocation problem (5.1), when the expression of the data rate is not clarified. The detailed MA power allocation problems in MIMO, in the full CSIT and in the statistical CSIT cases, are derived in the following sections.

Similarly to the SISO case, resource allocation consists of three separate steps (see Fig. 5.1): subcarrier allocation, power control and admission control. Subcarrier allocation is carried out before power control, independently in each BS. Its aim is to use the least power, assuming equal power allocation and equal interference level on all streams and subcarriers, and also to maximize a given criterion per subcarrier, denoted as E_k^l. It is similar to the method described in Section 5.4.1. In an initialization step, each user is assigned the subcarrier that maximizes E_k^l. It is followed by an iterative phase, that stops when all users have reached their target data rate, or when there are no subcarriers left. At each iteration, the user with lowest approximated data rate, k^*, is identified, and is assigned the free subcarrier that maximizes $E_{k^*}^l$. Whenever a new subcarrier is allocated to a user, its approximated data rate is updated, under the aforementioned assumptions on power and interference. In the full CSIT as well as in the statistical CSIT case, the chosen criterion is the convergence criterion E, which is given by equations (6.12) and (6.28).

Finally, a simple admission control method is used after power control. If the sum power exceeds the maximum $P_{\max}$ on a BS, the users served by this BS are ordered by descending sum transmit power value. Then the users with highest sum power are rejected from the cell, until the sum power becomes less than $P_{\max}$. A maximum power constraint $P_k^l < P_{\max}$ per user k and subcarrier l is also added in each iteration, in order to avoid too important power divergences.

In the following, we detail the convergence criterion and the associated power control in MIMO in two cases, corresponding to two levels of knowledge of the CSI at transmitter.

6.4 Margin Adaptive objective, full CSIT

6.4.1 Linear processing on each subcarrier

Let us first focus on subcarrier l. The received vector for user k served by BS k, $\mathbf{y}_k^l \in \mathbb{C}^{n_r \times 1}$, is

$$\mathbf{y}_k^l = \sqrt{\rho_k^l}\mathbf{H}_{k,k}^l \mathbf{x}_k^l + \sum_{\{n=1,n\neq k\}}^{N_{\mathrm{BS}}} \sqrt{\mu_{n,k}^l}\mathbf{H}_{n,k}^l \mathbf{x}_n^l + \mathbf{n}^l \tag{6.2}$$

where $\mathbf{x}_n^l \in \mathbb{C}^{n_t \times 1}$ is the vector transmitted by BS n, $\mathbf{H}_{n,k}^l \in \mathbb{C}^{n_r \times n_t}$ is the normalized zero-mean i.i.d. channel matrix for the transmission between BS n and user k, and $\mathbf{n}^l \in \mathbb{C}^{n_r \times 1}$ is the normalized AWGN noise vector.
$\rho_k^l = \frac{g_{k,k}P_k^l}{N_0}$ is the SNR of user k, and $\mu_{n,k} = \frac{g_{n,k}P_n^l}{N_0}$ is the INR corresponding to the interference received by user k from BS n in subcarrier l. $g_{n,k}$ is the channel gain between BS n and user k, including path loss and shadowing. All channel matrices are independent of each other and of the noise.
A linear precoder $\mathbf{V}_k^l \in \mathbb{C}^{n_t \times M}$ is used at transmission. The transmitted vector is equal to $\mathbf{x}_k^l = \mathbf{V}_k^l \mathbf{s}_k^l$, where $\mathbf{s}_k^l \in \mathbb{C}^{M \times 1}$ is an isotropic complex Gaussian vector that contains the M symbols to be transmitted. The precoder's covariance matrix with unitary trace is $\mathbf{\Phi}_k^l = \mathbf{V}_k^l (\mathbf{V}_k^l)^H$.
At the receiver, the estimated symbols vector is $\hat{\mathbf{s}}_k^l = (\mathbf{W}_k^l)^H \mathbf{y}_k^l$, where $(\mathbf{W}_k^l)^H \in \mathbb{C}^{M \times n_r}$ is a normalized linear processing matrix.
As explained in Section 2.1.3, we use the general denomination 'capacity' here to express the mutual information. The capacity of user k in subcarrier l is [12]

$$\begin{aligned} C_k^l &= \log_2\left(\det\left(\rho_k^l (\mathbf{W}_k^l)^H (\mathbf{H}_{k,k}^l \mathbf{\Phi}_k^l (\mathbf{H}_{k,k}^l)^H + \mathbf{Q}_k^l)\mathbf{W}_k^l\right)\right) - \log_2\left(\det\left((\mathbf{W}_k^l)^H \mathbf{Q}_k^l \mathbf{W}_k^l\right)^{-1}\right) \\ &= \log_2\left(\det\left(\mathbf{I}_M + \rho_k^l (\mathbf{W}_k^l)^H \mathbf{H}_{k,k}^l \mathbf{\Phi}_k^l (\mathbf{H}_{k,k}^l)^H \mathbf{W}_k^l \left((\mathbf{W}_k^l)^H \mathbf{Q}_k^l \mathbf{W}_k^l\right)^{-1}\right)\right) \end{aligned} \tag{6.3}$$

where $\mathbf{Q}_k^l$ is the covariance matrix of the interference plus noise for user k,

$$\mathbf{Q}_k^l = \sum_{\{n=1,n\neq k\}}^{N_{\mathrm{BS}}} \mu_{n,k}^l \mathbf{H}_{n,k}^l \mathbf{\Phi}_n^l (\mathbf{H}_{n,k}^l)^H + \mathbf{I}_{n_r}$$

We use a MMSE receiver, defined as

$$\mathbf{W}_k^l = \frac{(\mathbf{Q}_k^l)^{-1}\mathbf{H}_{k,k}^l \mathbf{V}_k^l}{\left\|(\mathbf{Q}_k^l)^{-1}\mathbf{H}_{k,k}^l \mathbf{V}_k^l\right\|_F} \tag{6.4}$$

By substituting (6.4) into (6.3) and using the equivalence $\det(\mathbf{I}+\mathbf{AB}) = \det(\mathbf{I}+\mathbf{BA})$ for square matrices, the capacity becomes [13] [86]

$$C_k^l = \log_2\left(\det\left(\mathbf{I}_{n_t} + \rho_k^l (\mathbf{H}_{k,k}^l)^H (\mathbf{Q}_k^l)^{-1} \mathbf{H}_{k,k}^l \mathbf{\Phi}_k^l\right)\right) \tag{6.5}$$

Equation (6.5) shows that the inter-cell interference plus noise is a colored Gaussian noise with covariance matrix $\mathbf{Q}_k^l$. The channel is equivalent to [87]

$$\mathbf{y}_k = \sqrt{\rho_k^l}(\mathbf{Q}_k^l)^{-1/2}\mathbf{H}_{k,k}^l \mathbf{x}_k^l + \check{\mathbf{n}}^l$$

where $\check{\mathbf{n}}^l$ is a normalized noise vector. The Singular Value Decomposition (SVD) of the equivalent channel is [88]

$$(\mathbf{Q}_k^l)^{-1/2}\mathbf{H}_{k,k}^l = \mathbf{U}_{k,1}^l \mathbf{\Lambda}_k^l (\mathbf{U}_{k,2}^l)^H \tag{6.6}$$

where $\mathbf{U}_{k,1}^l$ and $\mathbf{U}_{k,2}^l$ are unitary matrices, and $\mathbf{\Lambda}_k^l = \mathrm{diag}\left\{\lambda_{k,1}^l, ..., \lambda_{k,n_{\min}}^l\right\}$ is a diagonal matrix with real non-negative elements. $\mathbf{y}_k^l = \sqrt{\rho_k^l}(\mathbf{Q}_k^l)^{-1/2}\mathbf{H}_{k,k}^l\mathbf{x}_k^l + \check{\mathbf{n}}^l$ is equivalent to $\tilde{\mathbf{y}}_k^l = \sqrt{\rho_k^l}\mathbf{\Lambda}_k^l\tilde{\mathbf{x}}_k^l + \tilde{\mathbf{n}}^l$ where $\tilde{\mathbf{x}}_k^l = (\mathbf{U}_{k,2}^l)^H\mathbf{x}_k^l$, $\tilde{\mathbf{y}}_k^l = (\mathbf{U}_{k,1}^l)^H\mathbf{y}_k^l$ and $\tilde{\mathbf{n}}^l = (\mathbf{U}_{k,1}^l)^H\check{\mathbf{n}}^l$. The channel can consequently be represented as a parallel Gaussian channel composed of $n_{\min}$ subchannels. For $j = \{1, ..., n_{\min}\}$,

$$\tilde{y}_{k,j}^l = \sqrt{\rho_k^l}\lambda_{k,j}^l\tilde{x}_{k,j}^l + \tilde{n}_j^l$$

The SVD leads to $(\mathbf{H}_{k,k}^l)^H(\mathbf{Q}_k^l)^{-1}\mathbf{H}_{k,k}^l = \mathbf{U}_{k,2}^l\mathbf{\Delta}_k^l(\mathbf{U}_{k,2}^l)^H$ where $\mathbf{\Delta}_k^l = \mathrm{diag}\left\{(\lambda_{k,1}^l)^2, ..., (\lambda_{k,n_{\min}}^l)^2\right\}$. Consequently, the capacity is

$$\begin{aligned} C_k^l &= \log_2\left(\det\left(\mathbf{I}_{n_t} + \rho_k^l\mathbf{U}_{k,2}^l\mathbf{\Delta}_k^l((\mathbf{U}_{k,2}^l)^H\mathbf{\Phi}_k^l\right)\right) \\ &= \log_2\left(\det\left(\mathbf{I}_{n_t} + \rho_k^l(\mathbf{\Delta}_k^l)^{1/2}(\mathbf{U}_{k,2}^l)^H\mathbf{\Phi}_k^l\mathbf{U}_{k,2}^l(\mathbf{\Delta}_k^l)^{1/2}\right)\right) \end{aligned}$$

It is maximized when $(\mathbf{U}_{k,2}^l)^H\mathbf{\Phi}_k^l\mathbf{U}_{k,2}^l$ is diagonal, and when the per stream power allocation policy is water-filling over the singular values of the equivalent channel.

6.4.2 Resource allocation in multi-cell OFDMA

This section details MA resource allocation for user k when full CSI is available at the transmitter. Subcarrier allocation has been performed prior to power allocation with the method described in Section 6.3., and has assigned $l_{\mathrm{SC},k}$ subcarriers to user k in set Θ_k. $\theta_k : \mathbb{N} \mapsto \mathbb{N}$ is defined as the function that maps the l^{th} subcarrier allocated to user k on its absolute index, with regard to $\{1, ..., L_{\mathrm{SC}}\}$. The MIMO channel in each subcarrier is used in order to increase the transmission rate of one symbol, $M = 1$. The data rate per subcarrier is then

$$R_k^l = B_{\mathrm{SC}}\sum_{j=1}^{n_{\min}}\log_2\left(1 + \rho_k d_{k,j}^l(\lambda_{k,j}^l)^2\right) \tag{6.7}$$

where $d_{k,j}^l$ is the normalized power per subcarrier l and per stream j, $\rho_k = \frac{g_{k,k}P_k}{N_0}$ is the SNR of user k corresponding to its sum power on all subcarriers, and $\left\{\lambda_{k,1}^l, ..., \lambda_{k,n_{\min}}^l\right\}$ are the singular values of the equivalent channel $(\mathbf{Q}_k^l)^{-1/2}\mathbf{H}_{k,k}^l$. B_{SC} is the bandwidth per subcarrier. We consider the equivalent channel of user k aggregated on all its subcarriers. It is a diagonal block matrix, where the matrix in block l is equal to the equivalent channel matrix for user k in subcarrier $\theta_k(l)$:

$$\mathbf{H}_{k,k} = \begin{pmatrix} \mathbf{H}_{k,k}^{\theta_k(1)} & \mathbf{0}_{n_r \times n_t} & ... & \mathbf{0}_{n_r \times n_t} \\ \mathbf{0}_{n_r \times n_t} & \mathbf{H}_{k,k}^{\theta_k(2)} & ... & \mathbf{0}_{n_r \times n_t} \\ \mathbf{0}_{n_r \times n_t} & \mathbf{0}_{n_r \times n_t} & ... & \mathbf{H}_{k,k}^{\theta_k(l_{\mathrm{SC},k})} \end{pmatrix} \tag{6.8}$$

Then the sum data rate for user k is $R_k = \sum_{l\in\Theta_k} R_k^l$. The power allocation problem at each iteration is:

$$\begin{aligned}
&\min_{\left\{\mathbf{P_k},\mathbf{\Phi}_k^{\theta_k(1)},..,\mathbf{\Phi}_k^{\theta_k(l_{\mathrm{SC},k})}\right\}} \sum_{l\in\Theta_k} P_k^l \\
&\text{s.t.} \sum_{l\in\Theta_k} R_k^l \geq R_{k,\mathrm{target}} \\
&\text{s.t. } P_k^l \geq 0, \forall l \ \in \Theta_k
\end{aligned} \tag{6.9}$$

Its optimization variables are the power per subcarrier, $\mathbf{P_k}$, and the precoder matrix per subcarrier, $\left\{\mathbf{\Phi}_k^{\theta_k(1)},..,\mathbf{\Phi}_k^{\theta_k(l_{\mathrm{SC},k})}\right\}$. As the data rate is an increasing function of the power, the minimum sum power is obtained when $\sum_{l\in\Theta_k} R_k^l = R_{k,\mathrm{target}}$.

Power control convergence criterion

We first determine the criterion for power control convergence in subcarrier l. The target data rate of user k in subcarrier l is $R_{k,\mathrm{target}}^l$. All the streams are jointly used to achieve $R_{k,\mathrm{target}}^l$. Per stream power control is not feasible, as this would require a diagonalization of all interfering links in the same basis.
Our aim is to put the power control problem in the form $\left(\mathbf{I}_{N_{\mathrm{BS}}} - \mathbf{D}^l\mathbf{F}^l\right)\mathbf{P}^l \geq \mathbf{v}^l$, so that the reasoning of section 5.3 on distributed power convergence can be applied. For that purpose, it is assumed that only the path loss and shadowing gains $g_{n,k}$ of the interfering links are known at the transmitter, and that equal power allocation is used at transmission for all links. Under these assumptions, the covariance matrix becomes

$$\mathbf{Q}^l = \begin{pmatrix} \left(\sum_{\{n=1,n\neq k\}}^{N_{\mathrm{BS}}} \mu_{n,k}^l/n_t + 1\right)\mathbf{I}_{n_{\min}} & \mathbf{0}_{n_r-n_{\min}} \\ \mathbf{0}_{n_{\min}} & \mathbf{I}_{n_r-n_{\min}} \end{pmatrix}$$

We consider the case $n_r = n_{\min}$. Then the data rate of user k is

$$R_k^l = B_{\mathrm{SC}} \log_2\left(\det\left(\mathbf{I}_{n_t} + \frac{g_{k,k}P_k^l/n_t}{N_0 + \sum_{\{n=1,n\neq k\}}^{N_{\mathrm{BS}}} g_{n,k}P_n^l/n_t}\left(\mathbf{H}_{k,k}^l\right)^H\mathbf{H}_{k,k}^l\right)\right) \tag{6.10}$$

Let $\left(\beta_{k,1}^l\right)^2 \leq ... \leq \left(\beta_{k,n_{\min}}^l\right)^2$ be the ordered eigenvalues of $\left(\mathbf{H}_{k,k}^l\right)^H\mathbf{H}_{k,k}^l$. The data rate constraint, $\sum_{l\in\Theta_k} R_k^l \geq R_{k,\mathrm{target}}$, becomes

$$2^{\frac{R_{k,\mathrm{target}}^l}{B_{\mathrm{SC}}}} \leq \prod_{j=1}^{n_{\min}}\left(1 + \frac{g_{k,k}P_k^l/n_t}{N_0 + \sum_{\{n=1,n\neq k\}}^{N_{\mathrm{BS}}} g_{n,k}P_n^l/n_t}\left(\beta_{k,j}^l\right)^2\right)$$

Let $\tilde{\gamma_k^l} = 2^{\frac{R_{k,\mathrm{target}}^l}{B_{\mathrm{SC}} n_{\min}}} - 1$ be defined as the equivalent SINR for user k in subcarrier l. An upper bound is obtained by considering only the highest eigenvalue:

$$\tilde{\gamma_k^l} \leq \frac{g_{k,k}P_k^l/n_t}{N_0 + \sum_{\{n=1,n\neq k\}}^{N_{\mathrm{BS}}} g_{n,k}P_n^l/n_t}\left(\beta_{k,n_{\min}}^l\right)^2$$

It can be written as $\left(\mathbf{I}_{N_{\mathrm{BS}}} - \mathbf{D}^l\mathbf{F}^l\right)\mathbf{P}^l \geq \mathbf{v}^l$, where $\mathbf{v}^l = [v_1^l, ..., v_{N_{\mathrm{BS}}}^l]'$ with $v_k^l = \frac{\tilde{\gamma_k^l} N_0 n_t}{(\beta_{n_{\min}}^l)^2 g_{k,k}^l}$, $\mathbf{D}^l = \mathrm{diag}\left\{\tilde{\gamma_1^l}, ..., \tilde{\gamma_{N_{\mathrm{BS}}}^l}\right\}$, $(\mathbf{F}^l)_{(k,n)} = \frac{g_{n,k}^l}{(\beta_{k,n_{\min}}^l)^2 g_{k,k}^l}$ if $n \neq k$ and $(\mathbf{F}^l)_{(k,k)} = 0$.

As seen in Section 5.3.1, this inequality has a unique positive solution if $\rho(\mathbf{D}^l\mathbf{F}^l) < 1$. By using the infinity norm as an upper bound on the spectral radius, we obtain the following distributed criterion:

$$\tilde{\gamma}_k^l < \frac{\left(\beta_{k,n_{\min}}^l\right)^2 g_{k,k}}{\sum_{\{n=1,n\neq k\}}^{N_{\mathrm{BS}}} g_{n,k}}, \forall k \in \{1,...,N_{\mathrm{BS}}\} \tag{6.11}$$

The upper bound on $\tilde{\gamma}_k^l$ is denoted as E_k^l,

$$E_k^l = \frac{\left(\beta_{k,n_{\min}}^l\right)^2 g_{k,k}}{\sum_{\{n=1,n\neq k\}}^{N_{\mathrm{BS}}} g_{n,k}} \tag{6.12}$$

The convergence constraint (6.11) may be too restrictive if inter-cell interference is very low. We thus add a criterion to trigger the use of E_k^l on each subcarrier, depending on the inter-cell interference level. Let $I_k^l = N_0 + \sum_{\{n=1,n\neq k\}}^{N} \frac{g_{n,k}P_n^l}{n_t}$ be the inter-cell interference plus noise received by user k in subcarrier l in the previous iteration, considering only the path loss and shadowing gains. The transmitter for user k computes

$$\begin{aligned} \tau_k^l &= \frac{n_t}{\left(\beta_{n_{\min}}^l\right)^2}\left(\frac{I_k^l}{g_{k,k}P_{\max}} - \frac{N_0}{g_{k,k}P_{\max}}\right) \\ &= \frac{\sum_{\{n=1,n\neq k\}}^{N_{\mathrm{BS}}} g_{n,k}P_n^l}{\left(\beta_{n_{\min}}^l\right)^2 g_{k,k}P_{\max}} \end{aligned} \tag{6.13}$$

τ_k^l provides an estimate for the inter-cell interference level and is upper-bounded by $\frac{1}{E_k}$:

$$\frac{1}{E_k^l} \geq \tau_k^l \geq 0 \tag{6.14}$$

The transmitter assumes that inter-cell interference can be neglected if τ_k^l is low with regard to $\frac{1}{E_k^l}$, that is, if $E_k^l\tau_k^l < \delta$, where δ is a parameter that should be set depending on the network characteristics.

Power control

The power allocation problem for user k at each iteration is equal to problem (6.9) with the additional constraint

$$R_k^l \leq R_{k,\max}^l - \epsilon \;\text{ if }\; E_k^l\tau_k^l \geq \delta, \forall l \in \Theta_k \tag{6.15}$$

where $R_{k,\max}^l = B_{\mathrm{SC}} n_{\min} \log_2\left(1 + E_k^l\right)$ is the maximum data rate ensuring power convergence, and ϵ is a small positive value, accounting for the strict inequality in (6.11).
The power allocation problem is equivalently optimized over the set of SNRs per subcarrier and per stream, $\left\{\rho_k \mathbf{d}^{\theta_k(1)}, ..., \rho_k \mathbf{d}^{\theta_k(l_{\mathrm{SC},k})}\right\}$. It is a convex optimization problem, that is solved by using the KKT conditions.
If condition (6.15) is fulfilled on subcarrier l, the solution is:

$$\rho_k d_{k,j}^l = \left[\mu - \frac{1}{\left(\lambda_{k,j}^l\right)^2}\right]^+, \forall j \in \{1,...,n_{\min}\} \tag{6.16}$$

Else, the solution is:

$$\rho_k d_{k,j}^l = \left[\mu_l - \frac{1}{(\lambda_{k,j}^l)^2}\right]^+, \forall j \in \{1, ..., n_{\min}\} \tag{6.17}$$

In the first case, μ is a constant that applies to all subcarriers, and that is determined so as to fulfill the target data rate constraint from (6.9). In the second case, μ_l is a constant specific to subcarrier l, that is determined in order to fulfill the target data rate constraint on subcarrier l, $R_k^l = R_{k,\max}^l - \epsilon$. The proof is detailed in Appendix 6.B.
The convergence condition is tested during the bisection search that is used to obtain μ. For a given μ, on each subcarrier l, after having computed $\rho_k d_{k,j}^l$ with (6.16), the accuracy of using criterion E is first evaluated by computing τ_k^l with equation (6.13). If $E_k^l \tau_k^l \geq \delta$, condition (6.15) is tested. If (6.15) is not verified, the data rate on subcarrier l is set to $R_{k,\max}^l - \epsilon$, and the corresponding power values per stream $d_{k,j}^l$ are given by (6.17).
Finally, once μ has been obtained by bisection search, the power of user k in subcarrier l is computed as $P_k^l = \sum_{j=1}^{n_{\min}} \rho_k d_{k,j}^l$. The power per stream j is $\tilde{d}_{k,j}^l = d_{k,j}^l / \left(\sum_{j=1}^{n_{\min}} d_{k,j}^k\right)$, and the precoder's covariance matrix for user k in subcarrier l is $\mathbf{\Phi}_k^l = \mathbf{U}_{k,2}^l \text{diag}\left\{\tilde{d}_{k,j}^l, .., \tilde{d}_{k,\text{nmin}}^l\right\} (\mathbf{U}_{k,2}^l)^H$, where $\mathbf{U}_{k,2}^l$ is given by the SVD of the equivalent channel (6.6).

6.5 Margin Adaptive objective, statistical CSIT

6.5.1 Approximation of the outage capacity

In this section, our aim is to determine an approximate expression of the SNR required to reach a target outage capacity with a target outage probability on the zero-mean i.i.d. MIMO channel, $\mathbf{y} = \mathbf{Hx} + \mathbf{n}$. ρ is the SNR, and $\{\lambda_i\}_{\{1 \leq i \leq n_{\min}\}}$ are the singular values of $\mathbf{H}$. Equal power allocation over the transmit antennas is assumed. The outage probability, for a given target rate C in bits per channel use (bits/c.u.), is then

$$P_{\text{out}} = \mathcal{P}\left(\prod_{i=1}^{n_{\min}} \left(1 + \frac{\rho}{n_t}{\lambda_i}^2\right) < 2^C\right)$$

The outage capacity for a target outage probability $P_{\text{out, target}}$ is defined as the maximum rate C leading to an outage probability $P_{\text{out}} \leq P_{\text{out, target}}$. The approximate expression should be concave in the SNR, in order to turn the MA problem into a convex optimization problem. For that purpose, two bounds on the outage capacity are first determined.

An upper bound on the outage capacity

The inequality of arithmetic and geometric means gives

$$\prod_{i=1}^{n_{\min}} \left(1 + \frac{\rho}{n_t}{\lambda_i}^2\right) \leq \left(\frac{1}{n_{\min}} \sum_{i=1}^{n_{\min}} \left(1 + \frac{\rho}{n_t}{\lambda_i}^2\right)\right)^{n_{\min}}$$

The two functions are equal when all ${\lambda_i}^2$ have the same value. Consequently, this formula gives us the best upper bound linear in $\sum_{i=1}^{n_{\min}} \lambda_i^2$. A lower bound on the outage probability,

$P_{\text{out, min}}$, is deduced from this inequality:

$$
\begin{aligned}
P_{\text{out, min}} &= \mathcal{P}\left(n_{\min}^{-n_{\min}} \left(\sum_{i=1}^{n_{\min}} \left(1 + \frac{\rho}{n_t} {\lambda_i}^2\right) \right)^{n_{\min}} < 2^C \right) \\
&= \mathcal{P}\left({\|\mathbf{H}\|_F}^2 < \frac{n_{\min} n_t}{\rho} \left(2^{\frac{C}{n_{\min}}} - 1\right) \right)
\end{aligned}
$$

${\|\mathbf{H}\|_F}^2 = \sum_{i=1}^{n_{\min}} {\lambda_i}^2$ follows a chi-square law with $2n_t n_r$ degrees of freedom. The lower bound on the outage probability is then

$$P_{\text{out, min}} = \frac{\gamma\left(n_r n_t, \frac{n_{\min} n_t}{\rho}\left(2^{\frac{C}{n_{\min}}} - 1\right)\right)}{\Gamma(n_r n_t)} \tag{6.18}$$

where $\gamma : (n,u) \mapsto \int_0^u e^{-x} x^{n-1} dx$ is the incomplete Gamma function and $\Gamma : n \mapsto \int_0^\infty e^{-x} x^{n-1} dx$ is the complete Gamma function. This is equivalent to determining an upper bound on the outage capacity $C_{\max}$, for a given outage probability P_{out}:

$$P_{\text{out}} = \frac{\gamma\left(n_r n_t, \frac{n_{\min} n_t}{\rho}\left(2^{\frac{C_{\max}}{n_{\min}}} - 1\right)\right)}{\Gamma(n_r n_t)}$$

Let us introduce $u = \frac{n_{\min} n_t}{\rho}\left(2^{\frac{C_{\max}}{n_{\min}}} - 1\right)$ and $\Phi(u) = P_{\text{out}} = \frac{\gamma(n_r n_t, u)}{\Gamma(n_r n_t)}$. As $\Phi(0) = 0$, u as a function of P_{out} near 0 is given by the Taylor series of Φ^{-1} in $\zeta = (P_{\text{out}})^{\frac{1}{n_r n_t}}$,

$$u = f(P_{\text{out}}) = f(\zeta^{n_r n_t}) = \sum_{k=1}^{n} \frac{(\Phi^{-1})^{(k)}(0)}{k!} \zeta^k + o(\zeta^n) \tag{6.19}$$

where $(\Phi^{-1})^{(k)}$ is the k^{th} derivative of function Φ^{-1}. For instance, with $n_r = n_t = 2$, when stopping the Taylor series at $n = n_r n_t$, $f(P_{\text{out}})$ is equal to :

$$f(P_{\text{out}}) = f(\zeta^{n_r n_t}) = 2\sqrt[4]{\frac{3}{2}}\zeta + \frac{2\sqrt{6}\zeta^2}{5} + \frac{34\zeta^3}{252^{3/4}\sqrt[4]{3}} + \frac{388\zeta^4}{875} + o(\zeta^4)$$

The values of $f(P_{\text{out}})$ corresponding to practical P_{out} and antenna configurations are given in Table 6.1.

Table 6.1: $f(P_{\text{out}})$ for practical P_{out} values and antenna configurations

P_{out}	0.1	0.01	0.001
$n_t = 2$, $n_r = 2$	1.67148	0.81526	0.42732
$n_t = 2$, $n_r = 4$ and $n_t = 4$, $n_r = 2$	5.17513	3.04051	2.00321
$n_t = 3$, $n_r = 3$	5.36784	3.50424	2.45223
$n_t = 4$, $n_r = 4$	11.0458	8.17635	6.40502

An upper bound on the outage capacity as a function of the outage probability and of the SNR is then:

$$C_{\max} = n_{\min} \log_2\left(1 + \frac{f(P_{\text{out}})}{n_{\min} n_t}\rho\right) \tag{6.20}$$

A lower bound on the outage capacity

As ${\lambda_i}^2 > 0, \forall i \in \{1, ..., n_{\min}\}$, the following inequality stands

$$\prod_{i=1}^{n_{\min}} \left(1 + \frac{\rho}{n_t} {\lambda_i}^2\right) \geq 1 + \frac{\rho}{n_t} \sum_{i=1}^{n_{\min}} {\lambda_i}^2$$

The two functions are equal when only one ${\lambda_i}^2$ is different from zero. Consequently, this formula gives us the best lower bound linear in $\sum_{i=1}^{n_{\min}} \lambda_i^2$. An upper bound on the outage probability is then:

$$\begin{aligned} P_{\text{out, max}} &= \mathcal{P}\left(\left(1 + \frac{\rho}{n_t} \sum_{i=1}^{n_{\min}} {\lambda_i}^2\right) < 2^C\right) \\ &= \frac{\gamma\left(n_r n_t, \frac{n_t(2^C - 1)}{\rho}\right)}{\Gamma(n_r n_t)} \end{aligned} \tag{6.21}$$

This is equivalent to determining a lower bound on the outage capacity $C_{\min}$, for a given outage probability P_{out}. Let $u = \frac{n_t}{\rho}\left(2^{C_{\min}} - 1\right)$ and $\Phi(u) = P_{\text{out}} = \frac{\gamma(n_r n_t, u)}{\Gamma(n_r n_t)}$. $u = f(P_{\text{out}})$ is given by equation (6.19). Thus, a lower bound on the outage capacity as a function of the outage probability and of the SNR is:

$$C_{\min} = \log_2\left(1 + \frac{f(P_{\text{out}})}{n_t} \rho\right) \tag{6.22}$$

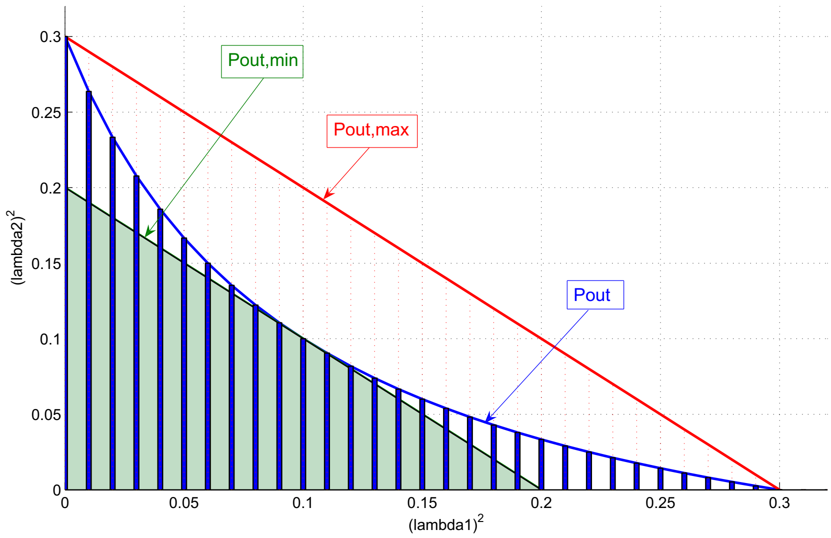


Figure 6.1: $n_{\min} = 2$, bounds on the outage probability

These two bounds on the outage probability are illustrated on Fig. 6.1 for $n_{\min} = 2$. The blue curve is $2^C = (1 + \frac{\rho}{n_t}{\lambda_1}^2)(1 + \frac{\rho}{n_t}{\lambda_2}^2)$, and the blue area is P_{out}. The green curve is obtained with the arithmetic-geometric means inequality, $2\sqrt{2^C} = 2 + \frac{\rho}{n_t}{\lambda_1}^2 + \frac{\rho}{n_t}{\lambda_2}^2$, and the green area is the lower bound, $P_{\text{out,min}}$. Finally, the red curve is obtained by only considering the trace, $2^C = 1 + \frac{\rho}{n_t}{\lambda_1}^2 + \frac{\rho}{n_t}{\lambda_2}^2$, and the red area is the upper bound, $P_{\text{out,max}}$. We used $\frac{\rho}{n_t} = 10$ and $C = 2$ bits/c.u.

Accuracy of the mean of the upper and lower bounds

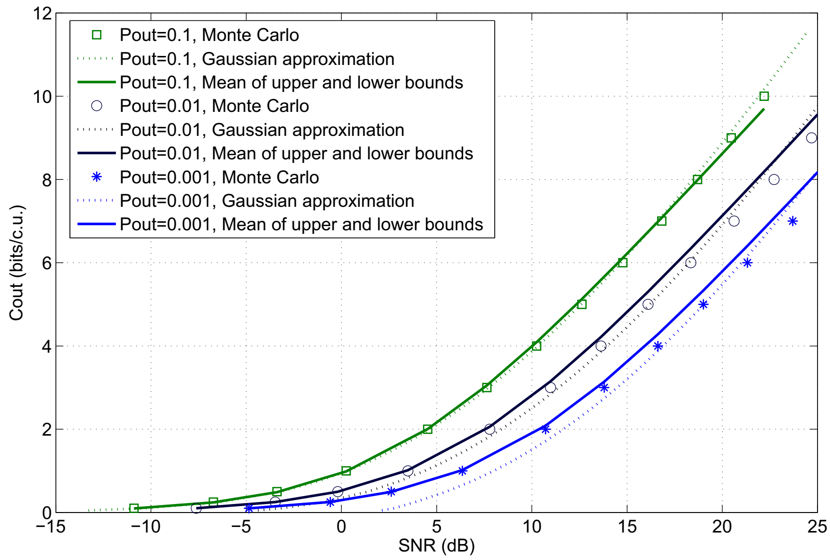


Figure 6.2: $n_t = n_r = 2$, accuracy of the approximation

The arithmetic mean of the upper and lower bounds is a concave function of the SNR, defined as:

$$\tilde{C} = \frac{1}{2}\left(n_{\min}\log_2\left(1+\frac{f(P_{\text{out}})}{n_{\min}n_t}\rho\right)+\log_2\left(1+\frac{f(P_{\text{out}})}{n_t}\rho\right)\right) \tag{6.23}$$

This approximation is compared with the outage capacity obtained with Monte-Carlo simulations, with 100000 snapshots. Fig. 6.2 and Fig. 6.3 show that $\tilde{C}$ is very close to the results obtained via Monte-Carlo simulations for several antenna configurations. It is even closer from the Monte-Carlo simulations than the Gaussian approximation from [85] at low SNR when $n_t = n_r = 2$. The absolute relative distance between $\tilde{C}$ and Monte-Carlo simulations increases with the SNR. If the outage capacity is lower than 10 bits/c.u., it remains lower than 10% with $n_r = n_t \in \{2, 3, 4\}$, and lower than 12% with $(n_t, n_r) = (2, 4)$ and $(4, 2)$, when $P_{\text{out}} \in \{10^{-1}, 10^{-2}, 10^{-3}\}$.

6.5.2 Resource allocation in multi-cell OFDMA

The approximate value (6.23) is used to model the outage capacity as a function of the outage probability and of the SNR in the downlink OFDMA MA problem, when only statistical CSI is available at the transmitter.

Let us consider power allocation for user k. To simplify notations, we introduce $a_k^l = \frac{f(P_{\text{out}})}{n_t}\frac{g_{k,k}}{I_k^l}$, where I_k^l is the inter-cell interference plus noise received by user k in subcarrier l. Using the approximate outage capacity expression (6.23), the data rate per subcarrier becomes

$$R_k^l = \frac{B_{\text{SC}}}{2}\left(n_{\min}\log_2\left(1+\frac{a_k^l}{n_{\min}}P_k^l\right)+\log_2\left(1+a_k^l P_k^l\right)\right) \tag{6.24}$$

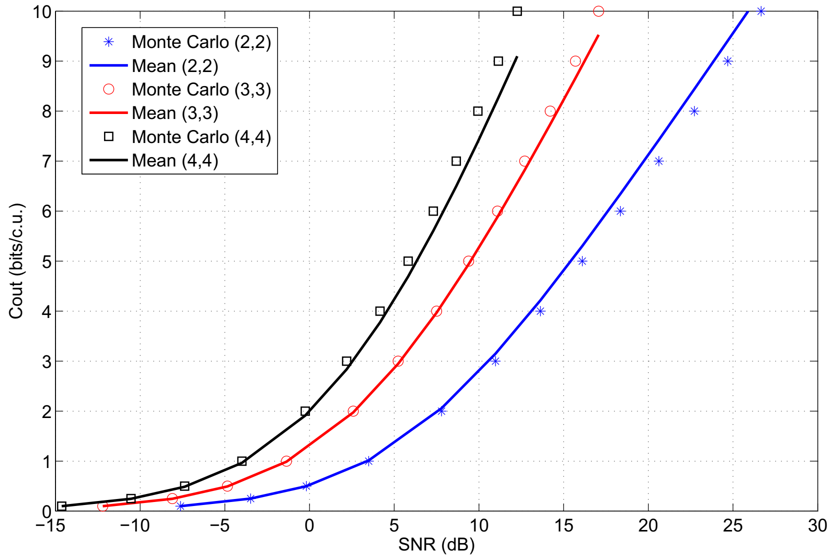


Figure 6.3: $P_{\text{out}} = 10^{-2}$, various number of antennas, accuracy of the approximations

The power allocation problem is the following:

$$\begin{aligned} &\min_{\mathbf{P}_k} \sum_{l\in\Theta_k} P_k^l \\ &\text{s. t. } \sum_{l\in\Theta_k} R_k^l = R_{k,\text{target}} \\ &\text{s. t. } P_k^l \geq 0, \forall l \in \Theta_k \end{aligned} \tag{6.25}$$

Power control convergence criterion

The outage capacity approximated by formula (6.23) is not linear in the power. In order to put power control in the required form $\left(\mathbf{I}_{N_{\text{BS}}} - \mathbf{D}^l\mathbf{F}^l\right)\mathbf{P}^l \geq \mathbf{v}^l$, the upper bound on the outage capacity (6.20) is considered. The data rate constraint per subcarrier is consequently

$$2^{\frac{R^l_{k,\text{target}}}{B_{\text{SC}} n_{\min}}} \leq \left(1 + \frac{f(P_{\text{out}})}{n_{\min} n_t} \frac{g_{k,k}P_k^l}{N_0 + \sum_{\{n=1,n\neq k\}}^{N_{\text{BS}}} g_{n,k}P_n^l/n_t}\right) \tag{6.26}$$

Let $\tilde{\gamma}_k^l = 2^{\frac{R^l_{k,\text{target}}}{B_{\text{SC}} n_{\min}}} - 1$. The data rate constraint can be written as $\left(\mathbf{I}_{N_{\text{BS}}} - \mathbf{D}^l\mathbf{F}^l\right)\mathbf{P}^l \geq \mathbf{v}^l$, where $\mathbf{v}^l = [v_1^l, ..., v_{N_{\text{BS}}}^l]'$ with $v_k^l = \frac{\tilde{\gamma}_k^l N_0 n_t n_{\min}}{f(P_{\text{out}})g^l_{k,k}}$, $\mathbf{D}^l = \text{diag}\left\{\tilde{\gamma}_1^l, ..., \tilde{\gamma_{N_{\text{BS}}}^l}\right\}$, $(\mathbf{F}^l)_{(k,n)} = \frac{n_{\min} g^l_{n,k}}{f(P_{\text{out}})g^l_{k,k}}$ if $n \neq k$ and $(\mathbf{F}^l)_{(k,k)} = 0$.

This problem has a unique positive solution if $\rho(\mathbf{D}^l\mathbf{F}^l) < 1$, where $\rho(\mathbf{D}^l\mathbf{F}^l)$ is the spectral radius of $\mathbf{D}^l\mathbf{F}^l$. By using the infinity norm as an upper bound on the spectral radius, we obtain the following distributed criterion:

$$\tilde{\gamma}_k^l < \frac{f(P_{\text{out}})g_{k,k}}{n_{\min}\sum_{\{n=1,n\neq k\}}^{N_{\text{BS}}} g_{n,k}}, \forall k \in \{1, ..., N_{\text{BS}}\} \tag{6.27}$$

The upper bound on $\tilde{\gamma}_k^l$ is denoted as E_k,

$$E_k = \frac{f(P_{\text{out}}) g_{k,k}}{n_{\min} \sum_{\{n=1,n\neq k\}}^{N_{\text{BS}}} g_{n,k}} \tag{6.28}$$

Similarly to the full CSIT case, the power control convergence criterion is tested whenever $E_k \tau_k^l \geq \delta$, where τ_k^l is defined as:

$$\begin{aligned}
\tau_k^l &= \frac{n_{\min} n_t}{f(P_{\text{out}})} \left(\frac{I_k^l}{g_{k,k} P_{\max}} - \frac{N_0}{g_{k,k} P_{\max}} \right) \\
&= \frac{n_{\min} n_t}{f(P_{\text{out}})} \left(\frac{\sum_{\{n=1,n\neq k\}}^{N_{\text{BS}}} g_{n,k} P_n^l}{g_{k,k} P_{\max}} \right)
\end{aligned} \tag{6.29}$$

Power control

The power allocation problem for user k in each iteration is equal to problem (6.25) with the additional constraint

$$R_k^l \leq R_{k,\max}^l - \epsilon \quad \text{if} \quad E_k^l \tau_k^l \geq \delta, \forall l \in \Theta_k \tag{6.30}$$

where $R_{k,\max}^l = B_{\text{SC}} n_{\min} \log_2 (1 + E_k)$ is the maximum allowed data rate for power convergence, and $\epsilon > 0$.

This problem is convex in $\mathbf{P}_k$. It is solved with the KKT conditions. The details are given in Appendix 6.C. When condition (6.30) is fulfilled, P_k^l is obtained by solving $g(P_k^l) = 0$, where g is defined as

$$g(P_k^l) = 1 - \nu \left(\frac{a_k^l}{2} \left(\frac{1}{1 + a_k^l P_k^l} + \frac{1}{1 + \frac{a_k^l}{n_{\min}} P_k^l} \right) \right)$$

This is equivalent to finding the roots of a polynomial of second order. If none of the roots is real positive, then P_k^l is set to zero. Else, P_k^l is equal to the minimum, strictly positive real solution. g is computed for a given Lagrange multiplier ν, that accounts for the target outage data rate constraint.

The convergence criterion is tested during the bisection search that is used to obtain ν. After having computed P_k^l for a given ν value by solving $g(P_k^l) = 0$, the accuracy of using criterion E is evaluated by computing τ with equation (6.29). If $E_k \tau_k^l \geq \delta$, condition (6.30) is tested. If it is verified, P_k^l is unchanged. Else, P_k^l is computed so as to exactly reach $R_k^l = R_{k,\max}^l - \epsilon$.

It should be noted that even if criterion E has the same value on all subcarriers, one test per subcarrier is required, as the τ_k^l values differ from one subcarrier to the other.

6.6 Performance results

6.6.1 MIMO with full CSIT

The studied network is composed of two rings of omnidirectional cells with same inter-site distance equal to 1.212 km. The path-loss model is Okumura-Hata, and the shadowing is log-normal with standard deviation equal to 7 dB. The thermal noise spectral density is

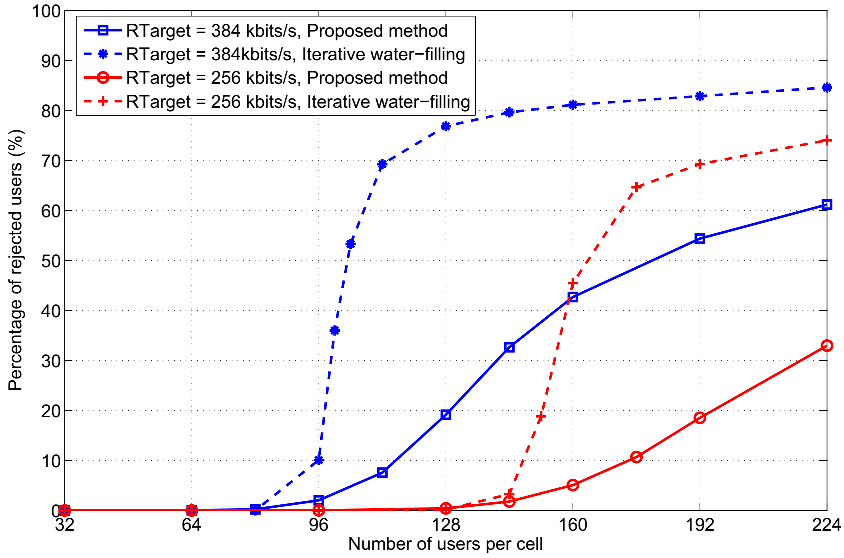


Figure 6.4: Full CSIT, percentage of rejected users depending on the load

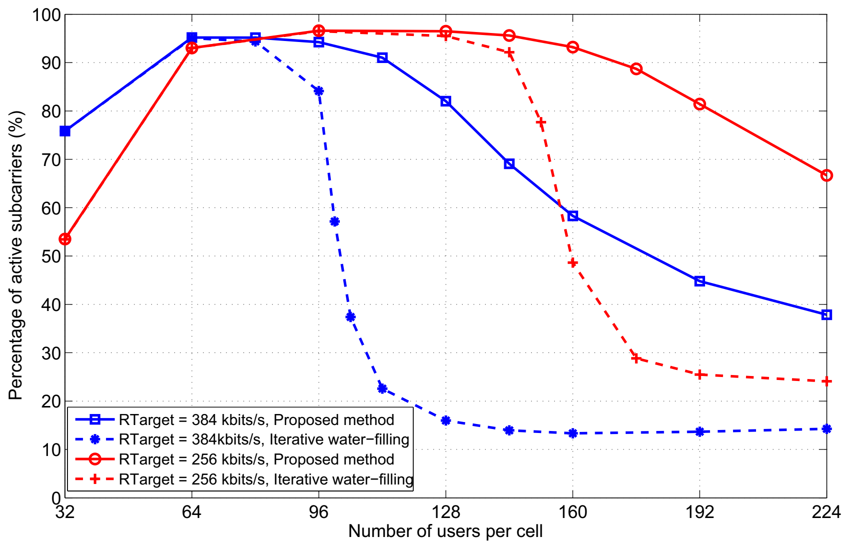


Figure 6.5: Full CSIT, percentage of active subcarriers depending on the load

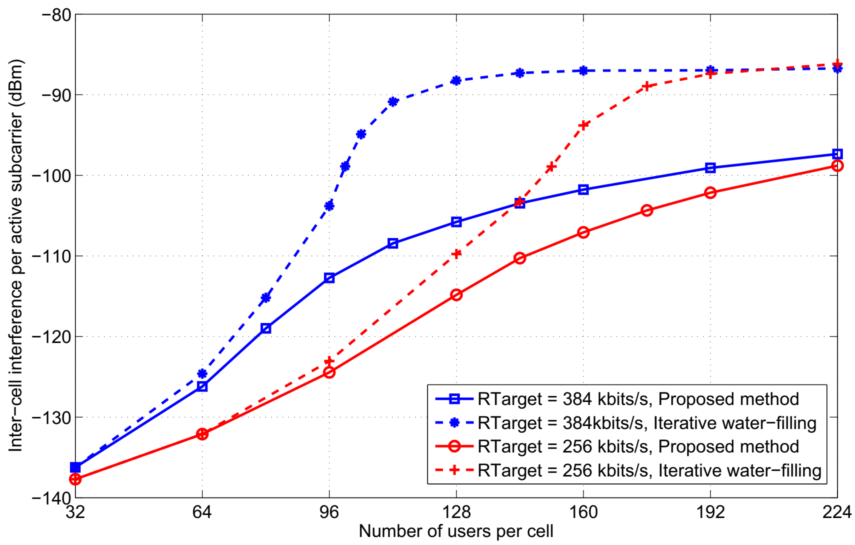


Figure 6.6: Full CSIT, inter-cell interference per active subcarriers, depending on the load

-174 dBm/Hz. The maximum transmit power for each BS is $P_{\max} = 43$ dBm. Each cell transmits in OFDMA with $L_{\mathrm{SC}} = 256$ subcarriers available for data transmission, and the network bandwidth is 10 MHz. The number of antennas is $n_t = n_r = 2$. The parameter for using criterion E is $\delta = 10^{-3}$. Our proposed method is compared with iterative water-filling [61]. In both cases, subcarrier allocation aims at maximizing criterion E (which is, here, equivalent to maximizing the maximum singular value of the direct channel). We consider two scenarios where all users have the same target data rate, $R_{\mathrm{target}} = 256$ or 384 kbits/s. Users that do not reach the target data rate are rejected.

In both scenarios, iterative water-filling leads to an abrupt increase of the percentage of rejected users, whereas our proposed method avoids this behavior (Fig. 6.4). The percentage of active subcarriers also decreases very rapidly with iterative water-filling when power divergence occurs (see Fig. 6.5 and Fig. 6.6). The convergence criterion for full CSIT MIMO only considers the statistical CSI from the interfering cells. As a consequence, we cannot theoretically guarantee that there will not be any power divergence situation with this criterion. However, the numerical results show that it is quite efficient compared to iterative water-filling, both at low and high load. At low load, criterion E is not too restrictive, thanks to the consideration of the highest singular value $\beta^{l}_{k,n_{\min}}$; whereas at high load, it identifies the subcarriers with potential power divergences and efficiently limits the allowed power levels on these subcarriers.

6.6.2 MIMO with statistical CSIT

The network parameters and assumptions are the same as in the full CSIT case, apart from $\delta = 10^{-2}$, and $R_{k,\mathrm{target}} = 128$ kbits/s. The proposed method is compared with iterative water-filling. In both cases, subcarrier allocation considers all subcarriers with the same priority, in the absence of per-subcarrier CSI.

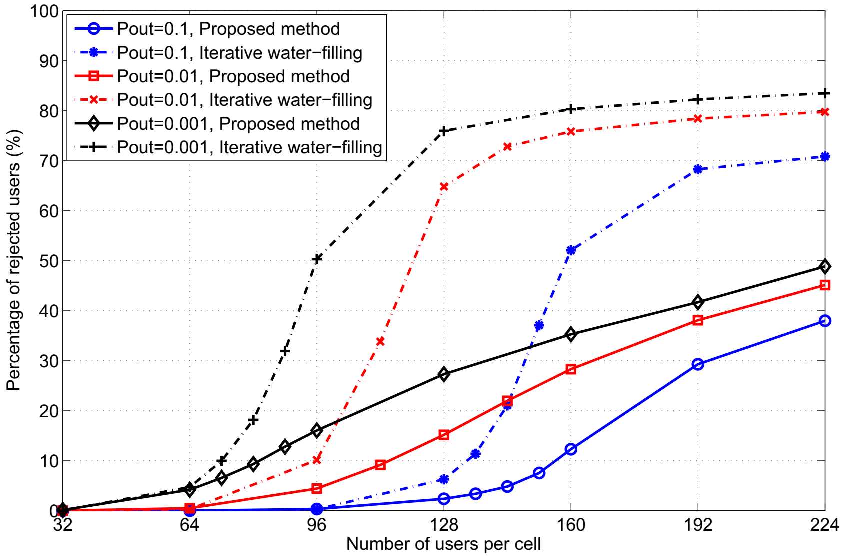


Figure 6.7: Statistical CSIT, percentage of rejected users depending on the load and on P_{out}

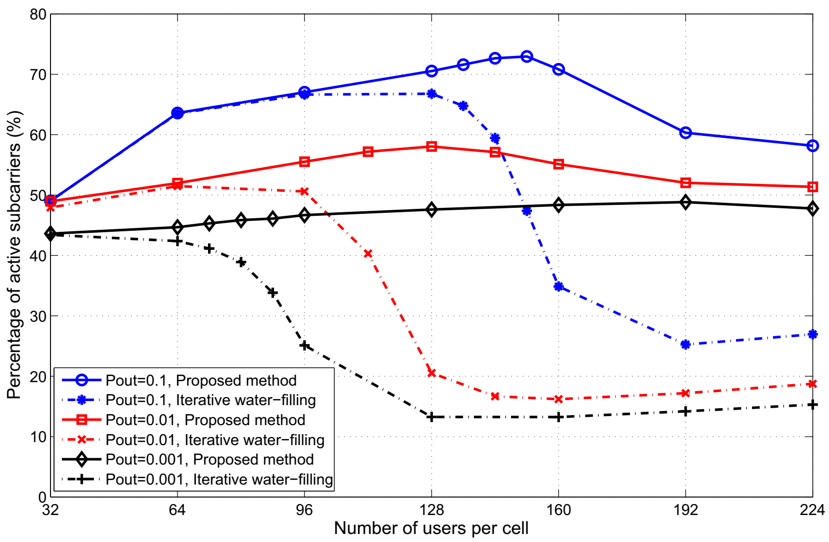


Figure 6.8: Statistical CSIT, percentage of active subcarriers, depending on the load and on P_{out}

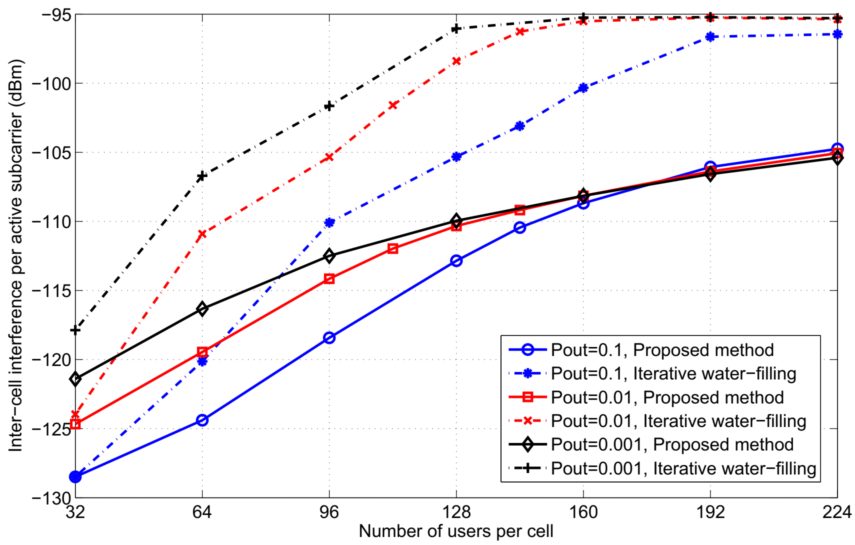


Figure 6.9: Statistical CSIT, inter-cell interference per active subcarrier, depending on the load and on P_{out}

Fig. 6.7 shows that our proposed method is more efficient in terms of rejection rate than iterative water-filling, whatever the outage probability value is. The maximum decrease in rejection rate compared to iterative water-filling is between 33 and 35% in the three cases. The proposed method avoids power divergence situations (see Fig. 6.9) that lead to many users being rejected by admission control. The rejection rate and the percentage of active subcarriers (Fig. 6.8) are not step-like functions as in the full CSIT case, emphasizing that we are dealing with outage data rates on the statistical channel.

6.7 Conclusion

In this chapter, we have determined distributed resource allocation methods to solve the MA problem in MIMO, when full CSI is available at transmission, and when only the statistical properties of the channel are available at transmission. The convergence criterion E has been derived in both cases, and included within subcarrier allocation and power control. In the full CSIT case, power control performs water-filling over the subcarriers and streams, with a per subcarrier convergence constraint E. In the statistical CSIT case, we have obtained an approximate analytical expression of the outage capacity as a function of the SNR and of the outage probability. This expression is concave in the power value, for a fixed inter-cell interference level, thus the MA problem is solved via convex optimization. The convergence criterion is triggered per subcarrier, depending on the interference level. In both cases, the proposed method avoids power divergence situations, compared to iterative water-filling. The percentage of rejected users and the inter-cell interference levels are consequently lower with our proposed methods than with iterative water-filling, at any load. These results complete the SISO results from Chapter 5, and show that the MA problem can be effectively solved in a distributed way.

Besides, the analytical expression of the outage capacity as a function of the SNR and of the outage probability for MIMO may be used in various resource allocation problems. Its concavity in the SNR is a very useful feature to solve resource allocation problems via convex optimization. It may also be used for virtual MIMO, in order to study the causal network coordination method from Chapters 3 and 4 with statistical CSIT. We do not investigate this topic in the dissertation. The next chapter still deals with fully distributed networks. As distributed resource allocation for RC users has been fully characterized in this chapter and the previous one, we are now interested in BE users.

APPENDIX

6.A Analytical expressions for the outage capacity

We consider the AWGN MIMO channel model $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$, where $\mathbf{H}$ is zero-mean i.i.d. ρ is the SNR, and C is the target rate, in bits/c.u.

6.A.1 Exact anatytical expressions for $n_t = 1$ or $n_r = 1$

In the SISO case, the outage probability is [14]

$$P_{\text{out}}(C) = 1 - \exp\left(-\frac{(2^C - 1)}{\rho}\right)$$

The outage capacity corresponding to a given outage probability P_{out} is then

$$C(P_{\text{out}}) = \log_2\left(1 - \ln\left(1 - P_{\text{out}}\right)\rho\right)$$

The Multiple-Input Single-Output (MISO) and Single-Input Multiple-Output (SIMO) cases have been treated in [12]. If $n_t = 1$, then

$$P_{\text{out}}(C, \rho) = \mathcal{P}\left(\log_2\left(\det\left(\mathbf{I} + \rho\mathbf{H}\mathbf{H}^H\right)\right) < C\right) = \mathcal{P}\left(\log_2\left(1 + \rho\mathbf{H}^H\mathbf{H}\right) < C\right)$$

$\mathbf{H}^H\mathbf{H} = \sum_{i=1}^{n_r} |h_i|^2$ and each h_i is a complex variable whose real and imaginary parts each follow an independent normal law. Consequently, $\mathbf{H}^H\mathbf{H}$ follows a chi-square random variable with $2n_r$ degrees of freedom. The pdf of a chi-square random variable with $2n$ degrees of freedom is

$$f_{\chi_n}(x) = \frac{e^{-x}x^{n-1}}{\Gamma(n)} = \frac{e^{-x}x^{n-1}}{[n-1]!} \text{ for } x > 0$$

where $\Gamma : n \mapsto \int_0^\infty e^{-x}x^{n-1}dx$ is the Gamma function (defined for $n \in \mathbb{C}$ and $Re(n) > 0$), and $\gamma : (n, u) \mapsto \int_0^u e^{-x}x^{n-1}dx$ is the incomplete Gamma function. Consequently,

$$P_{\text{out}}(C, \rho) = \frac{\gamma\left(n_r, \frac{(2^C - 1)}{\rho}\right)}{\Gamma(n_r)}$$

If $n_r = 1$, similarly, the outage probability is

$$P_{\text{out}}(C, \rho) = \frac{\gamma\left(n_t, \frac{n_t(2^C - 1)}{\rho}\right)}{\Gamma(n_t)}$$

Approximate expressions for the high SNR MISO and SIMO cases have been obtained in [82], by using the first order element of the Taylor series' approximation of the incomplete gamma function in 0. For $n_t = 1$, it is equal to:

$$P_{\text{out}}(C, \rho) = \frac{1}{\Gamma(n_r + 1)}\left(\frac{(2^C - 1)}{\rho}\right)^{n_r}$$

And for $n_r = 1$:

$$P_{\text{out}}(C, \rho) = \frac{1}{\Gamma(n_t + 1)}\left(\frac{n_t\left(2^C - 1\right)}{\rho}\right)^{n_t}$$

6.A.2 Asymptotic case

The distribution of the mutual information $\boldsymbol{I}$ is Gaussian in asymptotic cases, when the number of antennas becomes large [84].
For large n_r and fixed n_t, as $n_r \to \infty$,

$$\boldsymbol{I} \approx \mathcal{N}\left(n_t \log_2\left(1+\frac{n_r\rho}{n_t}\right), \frac{n_t \log_2(e)^2}{n_r}\right)$$

For large n_t and fixed n_r, as $n_t \to \infty$,

$$\boldsymbol{I} \approx \mathcal{N}\left(n_r \log_2(1+\rho), \frac{n_r\rho^2 \log_2(e)^2}{(1+\rho)^2}\right)$$

For large n_t and n_r, as $n_t \to \infty$ and $n_r \to \infty$,

$$\boldsymbol{I} \approx \mathcal{N}\left(n_r\rho \log_2(e), \frac{n_r}{n_t}\rho^2 \log_2(e)^2\right)$$

6.A.3 Gaussian approximation

The Gaussian approximation is even quite close to the mutual information distribution in the general case [85]. The outage probability is consequently well approximated by a Gaussian tail function:

$$\begin{aligned}P_{\text{out}}(C,\rho) &= \frac{1}{\sqrt{2\Pi}\sigma}\int_{-\infty}^{C} \exp\left(-\frac{(I-\mu(\rho))^2}{2(\sigma(\rho))^2}\right) dI \\ &= 1 - Q\left(\frac{C-\mu(\rho)}{\sigma(\rho)}\right)\end{aligned}$$

where $\mu(\rho)$ is the mean of the Gaussian distribution, $(\sigma(\rho))^2$ is its variance for a given SNR ρ, and $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-t^2/2}dt$.
The Gaussian approximation is computed with the following formulas:

$$\mu(\rho) = \int_0^\infty \log_2(1+\lambda\rho/n_t) K(\lambda,\lambda)d\lambda$$

$$\begin{aligned}\sigma(\rho)^2 = &\int_0^\infty (\log_2(1+\lambda\rho/n_t))^2 K(\lambda,\lambda)d\lambda \\ &-\int_0^\infty\int_0^\infty \log_2(1+\lambda_1\rho/n_t)\log_2(1+\lambda_2\rho/n_t) K^2(\lambda_1,\lambda_2)d\lambda_1 d\lambda_2\end{aligned}$$

where $n_{\min} = \min(n_t, n_r)$ and $d = \max(n_t, n_r) - n_{\min}$, and :

$$\begin{aligned}K(x,y) &= \sum_{i=0}^{n_{\min}-1} \Phi_i(x)\Phi_i(y) \\ \Phi_i(\lambda) &= \sqrt{\frac{i!}{(i+d)!}} L_i^d(\lambda)\lambda^{d/2}e^{-\lambda/2} \\ L_i^d(\lambda) &= \frac{1}{i!}e^\lambda\lambda^{-d}\frac{d}{d\lambda^i}(e^{-\lambda}\lambda^{d+i})\end{aligned}$$

L_i^d is the Laguerre polynomial of order i.

6.B Solution of power allocation, full CSIT case

The Lagrangian of the aggregate problem composed of (6.9) and (6.15) is:

$$\begin{aligned} L(\rho_k \mathbf{d}, \mu, \boldsymbol{\beta}, \boldsymbol{\alpha}) = & \sum_{l \in \Theta_k} \sum_{j=1}^{n_{\min}} \rho_k d_{k,j}^l + \mu \log(2) \left(\frac{R_{k,\text{target}}}{B_{\text{SC}}} - \sum_{j=1}^{n_{\min}} \log_2 \left(1 + \rho_k d_{k,j}^l (\lambda_{k,j}^l)^2 \right) \right) \\ & - \sum_{l \in \Theta_k} \sum_{j=1}^{n_{\min}} \alpha_j^l \rho_k d_{k,j}^l \\ & + \sum_{l \in \Theta_k} \beta^l (\tau_k^l)^+ \log(2) \left(\sum_{j=1}^{n_{\min}} \log_2 \left(1 + \rho_k d_{k,j}^l (\lambda_{k,j}^l)^2 \right) - \frac{(R_{k,\max}^l - \epsilon)}{B_{\text{SC}}} \right) \end{aligned} \tag{6.31}$$

where $\mu, \boldsymbol{\beta}, \boldsymbol{\alpha}$ are Lagrange multipliers. $\beta^l(\tau_k^l)^+$ is set to 0 if $E_k^l \tau_k^l < \delta$, and is a positive value if $E_k^l \tau_k^l \geq \delta$.

$$\frac{\partial L(\rho_k \mathbf{d}, \mu, \boldsymbol{\beta}, \boldsymbol{\alpha})}{\partial(\rho_k d_{k,j}^l)} = 1 - \frac{\mu (\lambda_{k,j}^l)^2}{1 + \rho_k d_{k,j}^l (\lambda_{k,j}^l)^2} - \alpha_j^l + \frac{\beta^l (\tau_k^l)^+ (\lambda_{k,j}^l)^2}{1 + \rho_k d_{k,j}^l (\lambda_{k,j}^l)^2}$$

The KKT conditions impose that $\frac{\partial L(\rho_k \mathbf{d}, \mu, \boldsymbol{\beta}, \boldsymbol{\alpha})}{\partial(\rho_k d_{k,j}^l)} = 0$. Condition $\alpha_j^l \geq 0$ leads to

$$1 - \frac{\mu (\lambda_{k,j}^l)^2}{1 + \rho_k d_{k,j}^l (\lambda_{k,j}^l)^2} + \frac{\beta^l (\tau_k^l)^+ (\lambda_{k,j}^l)^2}{1 + \rho_k d_{k,j}^l (\lambda_{k,j}^l)^2} \geq 0 \tag{6.32}$$

α_j^l is a slack variable that can be set to zero [24]. From $\frac{\partial L(\rho_k \mathbf{d}, \mu, \boldsymbol{\beta}, \boldsymbol{\alpha})}{\partial(\rho_k d_{k,j}^l)} = 0$ and the positivity constraint (6.32), the solution is:

$$\rho_k d_{k,j}^l = \left[\left(\mu - \beta^l (\tau_k^l)^+ \right) - \frac{1}{(\lambda_{k,j}^l)^2} \right]^+ \tag{6.33}$$

Let $R_k^l = B_{\text{SC}} \sum_{j=1}^{n_{\min}} \log_2 \left(1 + \rho_k d_{k,j}^l (\lambda_{k,j}^l)^2 \right)$ be the data rate on subcarrier l computed with (6.33). Due to the KKT condition $\beta^l (\tau_k^l)^+ \left(R_k^l - (R_{k,\max}^l - \epsilon) \right) = 0$, the data rate constraint per subcarrier $R_k^l = (R_{k,\max}^l - \epsilon)$ must be fulfilled whenever $\beta^l (\tau_k^l)^+ > 0$. On the contrary, if $R_k^l - (R_{k,\max}^l - \epsilon) < 0$, then $\beta^l (\tau_k^l)^+ = 0$. Therefore, the solution on subcarrier l is:

- If $R_k^l - (R_{k,\max}^l - \epsilon) < 0$:

$$\rho_k d_{k,j}^l = \left[\mu - \frac{1}{(\lambda_{k,j}^l)^2} \right]^+ , \forall j \in \{1, ..., n_{\min}\} \tag{6.34}$$

- Else:

$$\rho_k d_{k,j}^l = \left[\mu_l - \frac{1}{(\lambda_{k,j}^l)^2} \right]^+ , \forall j \in \{1, ..., n_{\min}\} \tag{6.35}$$

  where $\mu_l = (\mu - \beta^l (\tau_k^l)^+)$.

The constants μ and μ_l are determined so as to fulfill the data rate constraint, summed over all subcarriers in the first case, and restricted to subcarrier l in the second case.

6.C Solution of power allocation, statistical CSIT case

The aggregate problem composed of (6.25) and (6.30) is convex in $\mathbf{P}_k$. Its Lagrangian is:

$$
\begin{aligned}
L(\mathbf{P}_k, \nu, \boldsymbol{\beta}, \boldsymbol{\alpha}) = & \sum_{l \in \Theta_k} P_k^l - \sum_{l \in \Theta_k} \alpha^l P_k^l \\
& + \nu \log(2) \left(\frac{R_{k,\text{target}}}{B_{\text{SC}}} - \sum_{l \in \Theta_k} \frac{1}{2} \left(n_{\min} \log_2 \left(1 + \frac{a_k^l}{n_{\min}} P_k^l \right) + \log_2 \left(1 + a_k^l P_k^l \right) \right) \right) \\
& + \sum_{l \in \Theta_k} \beta^l (\tau_k^l)^+ \log(2) \left(\frac{1}{2} \left(n_{\min} \log_2 \left(1 + \frac{a_k^l}{n_{\min}} P_k^l \right) + \log_2 \left(1 + a_k^l P_k^l \right) \right) - \frac{(R_{k,\max}^l - \epsilon)}{B_{\text{SC}}} \right)
\end{aligned}
$$

where $\nu, \boldsymbol{\beta}, \boldsymbol{\alpha}$ are Lagrange multipliers. By the KKT conditions, the solution of (6.25) must verify $\frac{\partial L(\mathbf{P}_k, \nu, \boldsymbol{\beta}, \boldsymbol{\alpha})}{\partial P_k^l} = 0$ and $\alpha^l P_k^l = 0, \forall l \in \{1, ..., l_{\text{SC},k}\}$. The derivative of the Lagrangian is

$$
\begin{aligned}
\frac{\partial L(\mathbf{P}_k, \nu, \boldsymbol{\alpha})}{\partial P_k^l} &= 1 - \left(\nu - \beta^l (\tau_k^l)^+\right) \left(\frac{a_k^l}{2} \left(\frac{1}{1 + a_k^l P_k^l} + \frac{1}{1 + \frac{a_k^l}{n_{\min}} P_k^l} \right) \right) - \alpha^l \\
&= g(P_k^l) - \alpha^l
\end{aligned} \tag{6.36}
$$

$\alpha^l = 0$ if $P_k^l > 0$. In that case, the optimum P_k^l is a real positive solution is $g(P_k^l) = 0$. Solving $g(P_k^l) = 0$ is equivalent to finding the roots of a polynomial of second order.
Let R_k^l be the data rate on subcarrier l computed with the chosen solution of $g(P_k^l) = 0$. The KKT condition $\beta^l (\tau_k^l)^+ \left(R_k^l - (R_{k,\max}^l - \epsilon) \right) = 0$ imposes that the data rate constraint per subcarrier $R_k^l = (R_{k,\max}^l - \epsilon)$ be fulfilled whenever $\beta^l (\tau_k^l)^+ > 0$. On the contrary, if $R_k^l < (R_{k,\max}^l - \epsilon)$, then $\beta^l (\tau_k^l)^+ = 0$ Therefore, the solution on subcarrier l is:

- If $R_k^l < (R_{k,\max}^l - \epsilon)$, P_k^l is a real positive solution of $g(P_k^l) = 0$, where

  $$
  g(P_k^l) = 1 - \nu \left(\frac{a_k^l}{2} \left(\frac{1}{1 + a_k^l P_k^l} + \frac{1}{1 + \frac{a_k^l}{n_{\min}} P_k^l} \right) \right)
  $$

  If none of the solutions is real positive, then $P_k^l = 0$. Else, P_k^l is equal to the minimum, strictly positive real solution.

- Else, P_k^l is computed so as to exactly reach $R_k^l = R_{k,\max}^l - \epsilon$.

The constant ν is determined so as to reach the target data rate, summed over all subcarriers.

Chapter 7

Multi-cell weighted sum throughput maximization

7.1 Introduction

In the previous two chapters, we have determined resource allocation methods for RC users. The QoS objectives of these users are directly turned into the MA optimization problem, that should be solved per TTI. In this chapter, we consider BE users, that, unlike RC users, do not have any QoS constraint per TTI with regard to their data rate. However, when the queue lengths are taken into account, we can argue that a possible QoS objective for BE users is to avoid buffer overflows and their consequential data losses. In order to fulfill this requirement, cross-layer optimization between the medium access layer and the physical layer is needed. Radio resource allocation should therefore depend on the users' queue state. Consequently, the results of resource allocations in previous TTIs have an influence on the resource allocation objective in the present TTI. In this chapter, the studied optimization objective is Weighted Sum Throughput Maximization (WSTM). When the weight of each user is proportional to its queue length, WSTM may be seen as a cross-layer resource allocation problem for BE users. We will limit our study to SISO transmissions.
The chapter first reviews the literature relevant to our problem. Cross-layer resource allocation with the objective to avoid buffer overflows has been studied under the scope of information theory for the multi-access channel and for the broadcast channel. It was shown that maximizing the weighted sum of data rates, where the weight of each user is proportional to its queue length, leads to buffers' stability. Cross-layer optimization on the interference channel has not yet been studied. Based on the intuition that WSTM is a logical optimization problem to manage users' queue lengths, we study this problem in multi-cell OFDMA networks. The different papers focusing on this problem all make a high SINR assumption, that allows them to assure the convergence of their iterative methods. Indeed,

the high SINR assumptions turns the originally non-convex optimization problem into a geometric program. Our aim is to avoid performing this simplifying assumption, which is untrue in cellular networks with high levels of interference. For that purpose, we investigate two open issues:

- First, can we use subcarrier allocation to maximize the weighted sum data rate, based on channel coefficients and weights' knowledge?
- Second, how can we distinguish which users will eventually be in high SINR, prior to power control?

The first item is studied in Section 7.4. In order to determine an efficient subcarrier allocation algorithm, we study the capacity region of interfering users when inter-cell interference is treated as noise, and when each cell has a maximum power constraint. The simple case of two cells, one subcarrier and one user per cell, provides an analytical solution for the WSTM problem, and a criterion to evaluate if the joint transmission of both users leads to the maximization of the weighted sum throughput. This criterion is deduced from the convexity study of the capacity region, and from the relative links' weights. It can be extended to multi-cell OFDMA, thus indicating whether each couple of users should be transmitting on the same subcarrier or not. These binary conditions can be turned into an interference graph. The proposed subcarrier allocation then forms, via graph coloring, groups of users with allowed joint transmission.
The second item, studied in Section 7.5, deals with power control. Our aim is to determine a distributed power control that is valid in all SINR regimes. We propose a two-stage method. In the first step, power control is run for all users and subcarriers. Then a condition is tested on the SINR per subcarrier, to reject the users from the subcarriers where the SINR is too low. Finally, the second step uses a high SINR assumption for the remaining users and subcarriers. The optimization problem is thus a geometric program, and converges in a distributed way toward the global optimum. This method is compared with the optimum, centralized power control in Section 7.5.3. Then, Section 7.6 evaluates the performance results of the subcarrier allocation and power control methods, used jointly and separately, and compares them with binary power allocation, both in static and dynamic scenarios.
The main contributions of this chapter are:

- The capacity region of interfering users, when interference is considered as noise, is studied. A simple criterion for maximizing the weighted sum throughput per pair of links is deduced. It depends on the convexity of their capacity region, and on their relative links' weights. This criterion is extended to the multi-cell case by considering pairwise WSTM. Consequently, it can be used to build an interference graph for the network. A graph-based subcarrier allocation algorithm for OFDMA cellular networks using graph coloring is then deduced. It can be combined with a distributed power allocation algorithm operating in high SINR regime.
- A distributed power control algorithm is described for the WSTM problem. Unlike previous work, it is suitable for any SINR regime, as it first determines which users and subcarriers should be set to zero, and then operates in high SINR regime with the remaining ones. A comparison of the performance results of this algorithm with

the optimum, centralized power control shows that the proposed method leads to very low degradations in terms of data rates and in terms of power values.

- The proposed methods, used jointly or separately, are assessed via dynamic numerical simulations, where the weight of each user is proportional to its queue length. At medium to high load, they lead to higher weighted sum throughput and fairer queue length management than binary power allocation. Besides, graph-based subcarrier allocation followed by power control importantly decreases the power and subcarriers consumption.

7.2 State of the art

7.2.1 Information-theoretic results for multi-access channel and broadcast channel

Cross-layer resource allocation for the flat fading Gaussian multi-access channel (MAC) has been studied in [89]. The system is assumed feasible if for none of the users, the buffered queue is unbounded. A resource allocation policy that can stabilize a feasible system without knowledge of the queues' arrival rates is defined as throughput optimal. In order to determine such a policy, the authors refer to Tse and Hanly [90]. In this paper, all transmitters in a multi-access system have infinite backlogs of bits to send. Tse and Hanly solve the problem max $\boldsymbol{\mu}\mathbf{R}$, where $\boldsymbol{\mu}$ is a vector of non-negative weights. In [89], the authors show that a throughput optimal resource allocation policy for the MAC with random packet arrivals is given by the Tse-Hanly solution, where the direction $\boldsymbol{\mu}$ is equal to the queue length $\mathbf{u}$. The optimal rate allocation is obtained by performing successive decoding, from the shortest queue to the largest queue. Resource allocation for the broadcast channel is also investigated in [89]. Similarly to the MAC, it is shown that a throughput optimal resource allocation policy with random packet arrivals is given by the Tse, Li and Goldsmith [91] [92] solution (obtained with an infinite backlog of bits), where the direction $\boldsymbol{\mu}$ is chosen equal to the queue length $\mathbf{u}$.

7.2.2 Application to single-cell OFDMA

The capacity region of a single-cell OFDMA system with orthogonal subcarrier allocation is a sub-set of the capacity region of a single-cell OFDM system [93]. An OFDM cell with L_{SC} subcarriers is composed of L_{SC} multi-access channels in uplink, that are linked with one sum power constraint per user. Consequently, the stability properties defined in [89] for the MAC channel are valid for uplink OFDM. A throughput optimal strategy for the MAC channel is thus also valid for uplink OFDM and OFDMA. A similar result can be obtained for the downlink due to the duality of the MAC and broadcast channels [15].
Few papers deal with cross-layer queuing management and resource allocation in OFDMA. We here review the papers that consider the WSTM problem, where the weight of each user is proportional to its queue length. A general utility-based scheduling scheme for single-cell OFDMA is studied in [94]. It maximizes the projection of the data rates onto the gradient of a system's utility function. By choosing this gradient equal to the queue length, the mathematical framework from [94] can be used for solving WSTM. Resource allocation on

the long-term channel for users with heterogeneous services, namely dead-line sensitive and BE applications, is considered in [95]. The authors solve the problem by formulating it as a geometric program. This is only valid under a high SNR assumption. Several heuristics are proposed for subcarrier allocation with a WSTM objective in OFDMA in [96] [97], when the arrivals and channels are stochastic. They are compared with methods aiming at minimizing the long-term average packet delay. The main conclusion is that at low to medium load, balancing the queues between users is more critical than opportunistically taking advantage of the channel variations, while the opposite becomes true at high load.

7.2.3 WSTM resource allocation on the interference channel

It is not known yet if the information-theoretic results obtained on the MAC and broadcast channel extend to the interference channel. However, the WSTM resource allocation strategy intuitively seems well suited to efficiently balance the queues of users in any multi-user scenario. We will therefore consider the WSTM problem with the weight of each user proportional to its queue length as an example of possible application, and evaluate its influence on queue lengths management in the numerical assessments.
The interference channel is investigated in high SINR regime in [98]. In that case, the power control problem is formulated as a geometric program, that can be solved in a distributed way, by using Lagrange dual decomposition. In [99], Qiu and Chawla propose an iterative method for solving the sum throughput maximization problem in the general case and in the high SINR regime. In the general case, the obtained distributed solution may be a local optimum, whereas there is only one global optimum in the high SINR regime, because the objective function is then a standard interference function [69]. In [100], Chiang studies power control in conjunction with congestion control at TCP layer. The weight for each user in WSTM is proportional to the queuing delay. For a given weights' set, power control is then solved iteratively with the gradient method, under a high SINR assumption. Huang et al. propose a similar approach in [101], where supermodular game theory is used to show the convergence of the iterative process. In the same paper, the multi-channel case is solved for the high SINR regime, through a decomposition in the dual space.
A binary approach in multi-cell TDMA networks has been studied in [102], where each BS either transmits at full power, or does not transmit. The authors prove that for the two-cells case with one user per cell, the binary power allocation is optimum to solve the sum throughput maximization problem. This result does not extend to the cases with more than two interfering BSs. However, numerical results show that the sum throughput with binary allocation is very close to the optimum sum throughput. WSTM is not treated in [102]. Binary power allocation will serve a as reference for performance comparisons throughout this chapter.

7.3 WSTM problem

We consider the downlink transmission in a network $\mathcal{N}$ composed of $K_{\mathcal{N}}$ users and N_{BS} BSs using OFDMA. Each BS has L_{SC} subcarriers available for data transmission. The total available bandwidth is B, and the bandwidth per subcarrier is B_{SC}. The data rate of user

k served by BS n_{BS} is

$$R_k = \sum_{l \in \Theta_k} B_{\mathrm{SC}} \log_2 \left(1 + \frac{G^l_{k,k} P^l_k}{\left(N_0 + \sum_{n \neq k} G^l_{n,k} P^l_n \right)} \right) \tag{7.1}$$

where Θ_k is the set of subcarriers assigned to user k by n_{BS}, P^l_k is the power transmitted to user k by n_{BS} in subcarrier l, $G^l_{n,k}$ is the channel coefficient between BS n and user k in subcarrier l (including propagation loss, shadowing, and fast fading). N_0 is the variance of the AWGN noise. To simplify notations, we use $G^l_{k,k}$ for the channel coefficient between user k and n_{BS} in subcarrier l, and $P^l_k = P^l_{n_{\mathrm{BS}}}$ if subcarrier l is allocated to user k.
Let w_k be the weight of user k. The WSTM problem per TTI is:

$$\max_{\{\mathbf{P},\mathbf{\Theta}\}} \sum_{k=1}^{K_{\mathcal{N}}} w_k R_k \tag{7.2}$$

$$\text{s. t. } \sum_{l=1}^{L_{\mathrm{SC}}} P^l_{n_{\mathrm{BS}}} \leq P_{\max}, \forall n_{\mathrm{BS}} \in \{1, ..., N_{\mathrm{BS}}\}$$

$$\text{s. t. } P^l_{n_{\mathrm{BS}}} \geq 0, \forall (n_{\mathrm{BS}}, l) \in \{1, ..., N_{\mathrm{BS}}\} \times \{1, ..., L_{\mathrm{SC}}\}$$

$$\text{s.t. } \Theta_k \cap \Theta_{k'} = \emptyset, \forall (k, k') \text{ served by the same BS}, k \neq k'$$

This network-wide optimization problem is NP-hard due to the discrete subcarrier assignment, and due to the non-convexity of the power allocation function $S(\mathbf{P}, \mathbf{\Theta}) = \sum_{k=1}^{K_{\mathcal{N}}} w_k R_k$, even if subcarriers are already allocated [24]. Our aim is to determine sub-optimal solutions to that problem. Distributed algorithms per BS are preferred, in order to allow for a simple implementation in networks with flat architecture. We assume that subcarrier allocation is performed separately from power allocation. In the following, we equivalently write $S(\mathbf{P}, \mathbf{\Theta})$ as $S(\mathbf{P})$, by including the subcarrier allocation constraint as a binary power allocation constraint: on each cell, $\Theta_k \cap \Theta_{k'} = \emptyset, \forall (k, k')$ served by the same BS, $k \neq k'$, is indeed equivalent to $\forall l \in \{1, ..., L_{\mathrm{SC}}\}$, if $P^l_k > 0$ then $P^l_{k'} = 0, \forall (k, k')$ served by the same BS, $k \neq k'$.

7.4 WSTM subcarrier allocation

In this section, we propose two subcarrier allocation methods for WSTM. The first one, called graph-based subcarrier allocation, is deduced from the capacity region study. The second one, called WSTM distributed subcarrier allocation, is a simple distributed method that will be used in the numerical studies for comparison purposes.

7.4.1 Capacity region study : two users case

We first investigate the case with $N_{\mathrm{BS}} = 2$ cells, 1 user per cell, and 1 subcarrier per cell. In this section, we consider the normalized rates, with a bandwidth $B = B_{\mathrm{SC}} = 1$. Therefore, we remove B and the subcarrier index l to simplify notations.
We have seen in Section 2.1.4 that, on the interference channel, the interference may be jointly decoded by all receivers, when the transmitted information is composed of both private and common information [19]. Such methods lead to an increase of the generalized

degree of freedom of the interference channel [20], $d_{\text{sym}}(\alpha)$ (defined by equation (2.9)), when $\alpha = \frac{\log \text{INR}}{\log \text{SNR}}$ exceeds 0.5. Fig. 7.1 depicts the capacity-achieving generalized degree of freedom for the Gaussian symmetric interference channel with two users, in red. The green line represents the generalized degree of freedom when interference is considered as noise by both users. The blue line represents the generalized degree of freedom when users are never jointly transmitting.

In this dissertation, we do not investigate methods involving common information exchanges, and we restrict our study to either orthogonal transmission, or joint transmission when inter-cell interference is treated as noise. Our aim is then to determine in which cases the joint transmission of both users is more efficient than orthogonal transmission, with regard to the WSTM objective.

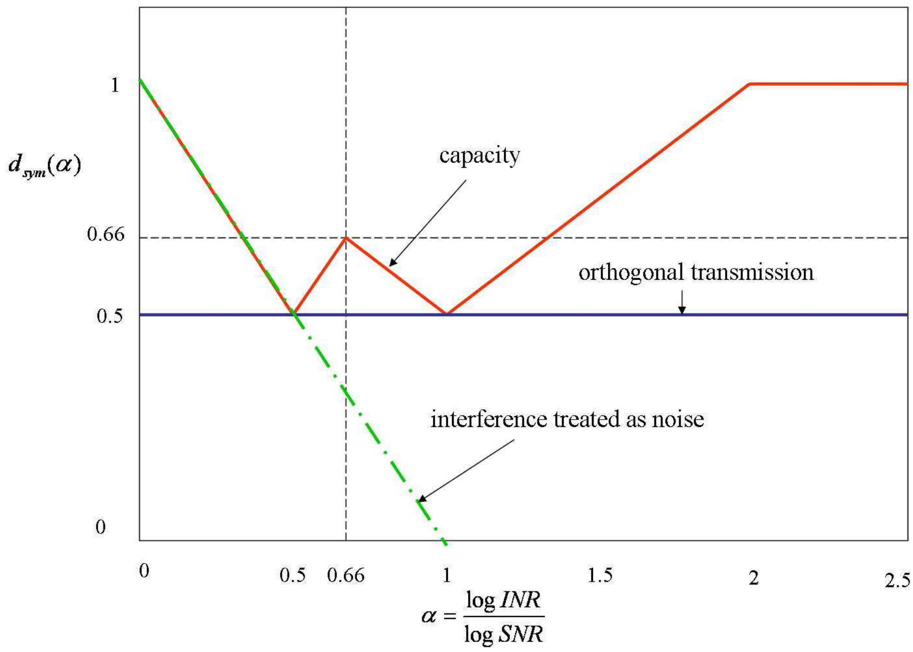


Figure 7.1: Generalized degree of freedom

The capacity region is defined as the set of all simultaneously achievable rate pairs (R_k, R_n), under the set of per cell power constraints. The data rate of user k interfered by user n is

$$R_k = \log_2\left(1 + \frac{G_{k,k}P_k}{N_0 + G_{n,k}P_n}\right)$$

with $0 \leq P_k \leq P_{\max}$ and $0 \leq P_n \leq P_{\max}$.

The rate of user n can be expressed as a function of its power, P_n, and of the interfering link's rate, R_k:

$$R_n = \tilde{f}_n(R_k, P_n) = \log_2\left(1 + \frac{G_{n,n}P_n}{N_0 + \frac{G_{k,n}}{G_{k,k}}(2^{R_k} - 1)(N_0 + G_{n,k}P_n)}\right) \tag{7.3}$$

Let us set $a = N_0 + \frac{G_{k,n}}{G_{k,k}}(2^{R_k} - 1)N_0$ and $b = \frac{G_{k,n}}{G_{k,k}}(2^{R_k} - 1)G_{n,k}$. Then, for any R_k,

$$\frac{\partial \tilde{f}_n(R_k, P_n)}{\partial P_n} = \frac{aG_{n,n}}{\log(2)(a + bP_n)(a + bP_n + G_{n,n}P_n)} > 0$$

Table 7.1: Parameters for the studied cases

	$G_{1,1}$	$G_{2,2}$	$G_{1,2}$	$G_{2,1}$
Case 1, Interference-limited	10^{-12}	10^{-12}	10^{-13}	10^{-13}
Case 2, Balanced case	10^{-12}	10^{-12}	10^{-14}	10^{-14}
Case 3, Power-limited	10^{-12}	10^{-12}	10^{-16}	10^{-16}
Case 4, Asymmetric case	10^{-12}	10^{-13}	10^{-13}	10^{-16}

Consequently for a fixed R_k, the maximum of $\tilde{f}_n(R_k, P_n)$ is obtained when $P_n = P_{\max}$. The capacity region is therefore determined by the following set of inequations:

$$R_n \leq f_n(R_k), \qquad \forall (n,k) \in \{1,2\}^2 \tag{7.4}$$

where $f_n(R_k) = \tilde{f}_n(R_k, P_{\max})$.
Four examples of capacity regions are represented on Fig. 7.2: an interference-limited case with concave capacity region, and three other cases with convex capacity regions. On these figures, the noise is $N_0 = -105$ dBm, and the maximum power is $P_{\max} = 43$ dBm. The gain parameters are given in Table 7.1 .

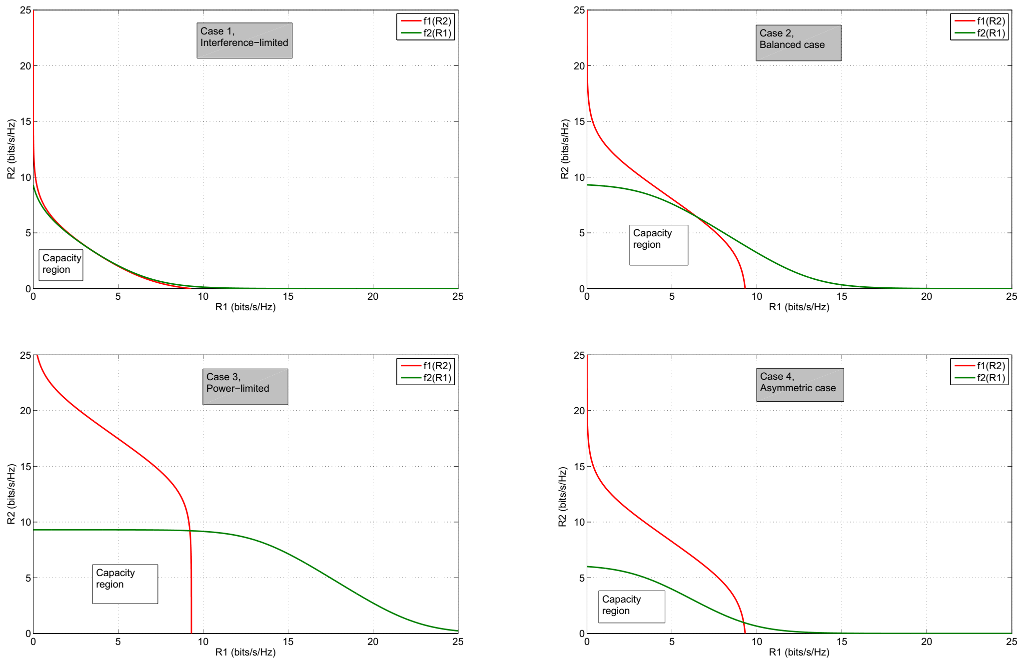


Figure 7.2: Capacity regions

The weighted sum data rate is $S(\mathbf{P}) = w_k R_k + w_n R_n$, with $w_k + w_n = 1$. If the solution that maximizes S is different from the extrema, then it is tangent to one of the functions f_n. Consequently, $f_n'(R_k^*) = -\frac{w_k}{w_n}$. We assume that $w_n \geq w_k$.
Two cases are possible regarding f_n curvature:

1. f_n is convex on $[0, +\infty[$. In that case, the solution to the weighted sum data rate maximization is $(P_{\max}, 0)$ or $(0, P_{\max})$.

2. f_n has only one inflexion point, R_k^{inflex}. In that case, f_n is first concave in $[0, R_k^{\text{inflex}}]$, and then convex in $[R_k^{\text{inflex}}, +\infty[$. We focus on this second case in the following.

The candidate solutions of $f_n'(R_k^*) = -\frac{w_k}{w_n}$ are $R_{k,a}^* = \log_2\left(\frac{\upsilon - \sqrt{\alpha}}{\chi}\right)$ and $R_{k,b}^* = \log_2\left(\frac{\upsilon + \sqrt{\alpha}}{\chi}\right)$, where

$$\begin{aligned}
\upsilon =& (w_n - w_k)G_{k,k}G_{n,n}P_{\max} - 2w_kG_{k,k}N_0 + 2w_kG_{k,n}\left(N_0 + G_{n,k}P_{\max}\right)\\
\alpha =& G_{k,k}G_{n,n}P_{\max}\big[\big((w_n - w_k)^2G_{k,k}G_{n,n} + 4w_kw_nG_{k,n}G_{n,k}\big)P_{\max}\\
&+ 4w_kw_nN_0(G_{k,n} - G_{k,k})\big]\\
\chi =& 2w_kG_{k,n}\left(N_0 + G_{n,k}P_{\max}\right)
\end{aligned}$$

In the concave and convex areas of f_n, f_n' varies in $]-1, 0[$. It is never equal to (-1). The sum data rate maximization can consequently not be achieved by studying only one function f_n. Its solution is binary, as shown in [102].

Let m_n be the minimum for $\frac{w_k}{w_n}$ in the concave area of f_n, and M_n its maximum. m_n is equal to

$$m_n = -f_n'(0) = \frac{G_{k,n}G_{n,n}P_{\max}\left(N_0 + G_{n,k}P_{\max}\right)}{G_{k,k}N_0\left(N_0 + G_{n,n}P_{\max}\right)} \tag{7.5}$$

and $M_n = -f_n'(R_k^{\text{inflex}})$. The expression of M_n cannot easily be simplified and is therefore not detailed here.

The solution R_k^* must lie in the concave area of f_n, so that it lies in the convex capacity region. Therefore only $R_{k,a}^*$ may be suitable. There is a solution to $f_n'(R_k) = -\frac{w_k}{w_n}$ in the concave area if $\frac{w_k}{w_n} \in [m_n, M_n]$. It is equal to $P_k = \left(\frac{N_0 + G_{n,k}P_{\max}}{G_{k,k}}\right)\left(1 - \frac{\upsilon - \sqrt{\alpha}}{\chi}\right)$. It belongs to the capacity region if $0 \leq P_k \leq P_{\max}$. Consequently,

$$P_k^* = \min\left\{\left[\left(\frac{N_0 + G_{n,k}P_{\max}}{G_{k,k}}\right)\left(1 - \frac{\upsilon - \sqrt{\alpha}}{\chi}\right)\right]^+, P_{\max}\right\} \tag{7.6}$$

and the solution to the maximum weighted sum data rate is $(P_k^*, P_{\max})$.

If $\frac{w_k}{w_n} \notin [m_n, M_n]$, then the solution is binary. To conclude, the optimum solution for maximizing S, if $w_n \geq w_k$, belongs to the following set:

$$\mathcal{P}_{\text{joint}} = \{(0, P_{\max}), (P_{\max}, 0), (P_{\max}, P_{\max}), (P_k^*, P_{\max})\} \tag{7.7}$$

We now study the convexity of the capacity region with regard to cell n through the sign of $f_n''(0)$. If $f_n''(0) > 0$, then f_n is convex, and the capacity region is concave. If $f_n''(0) \leq 0$, then f_n is concave between 0 and the inflexion point, and the capacity region is convex in this area.

$$f_n''(0) = -\frac{\log(2)G_{k,n}G_{n,n}P_{\max}\left(N_0 + G_{n,k}P_{\max}\right)\psi}{(G_{k,k})^2(N_0)^2\left(N_0 + G_{n,n}P_{\max}\right)^2}$$

where

$$\psi = G_{k,k}N_0\left(N_0 + G_{n,n}P_{\max}\right) - G_{k,n}\left(N_0 + G_{n,k}P_{\max}\right)\left(2N_0 + G_{n,n}P_{\max}\right)$$

Consequently, $f_n^{''}(0) \leq 0$ is equivalent to $A_{n,k} \geq 0$, where

$$A_{n,k} = \frac{G_{k,k}N_0}{N_0 + G_{n,k}P_{\max}} - G_{k,n}\frac{(2N_0 + G_{n,n}P_{\max})}{(N_0 + G_{n,n}P_{\max})} \tag{7.8}$$

In the following reasoning, we suppose that $w_n \log_2\left(1 + \frac{G_{n,n}P_{\max}}{N_0}\right) \geq w_k \log_2\left(1 + \frac{G_{k,k}P_{\max}}{N_0}\right)$. If the capacity region is not convex, the maximum weighted sum data rate is equal to the maximum data rate when only user n transmits, $S(\mathbf{P}) = w_n \log_2\left(1 + \frac{G_{n,n}P_{\max}}{N_0}\right)$. If the capacity region is convex in $\left[0, R_k^{\text{inflex}}\right]$, any tangent to f_n at a point in $\left[0, R_k^{\text{inflex}}\right]$ will have a higher ordinate $S(\mathbf{P}) = w_k R_k + w_n R_n$ than $f_n(0) = \log_2\left(1 + \frac{G_{n,n}P_{\max}}{N_0}\right) \geq w_n \log_2\left(1 + \frac{G_{n,n}P_{\max}}{N_0}\right)$. Consequently, if the capacity region is convex, and if a solution to $f_n'(R_k^*) = -\frac{w_k}{w_n}$ exists in $\left[0, R_n^{\text{inflex}}\right]$, then the weighted sum data rate at R_k^* is higher than the weighted sum data rate obtained when transmission is limited to user n.
Therefore, the weighted sum data rate can be increased by assigning subcarriers in order to avoid having $f_n^{''}(0) > 0$ or $f_k^{''}(0) > 0$, which results in avoiding that $A_{n,k} < 0$ or $A_{k,n} < 0$. It should be noted that criterion (7.8) does not depend on the weight values, as it only characterizes the convexity of the capacity region, which is independent of the weights. However, we have seen that if $\frac{w_k}{w_n} \in [0, m_n]$ or $\frac{w_n}{w_k} \in [0, m_k]$, the optimum solution is binary. As in that case, the weight of one of the users is very low, it is highly likely that the binary solution will be either $(0, P_{\max})$ or $(P_{\max}, 0)$. Consequently, we add a second condition, dependent on the weights, to evaluate if the joint transmission of users k and n leads to a maximization of the weighted sum data rate.
To conclude, the conditions are :

$$A_{k,n} \geq 0,\ A_{n,k} \geq 0,\ \frac{w_k}{w_n} \geq m_n \text{ and } \frac{w_n}{w_k} \geq m_k \tag{7.9}$$

7.4.2 Graph-based subcarrier allocation for OFDMA

Let us now consider the interference channel with $N_{\text{BS}} > 2$ links. The weighted sum data rate maximization problem on the whole network is too complex to be directly solved in the general case. Therefore, we propose a sub-optimal subcarrier allocation method that aims at maximizing the weighted sum data rate per couple of links (k, n). For each couple, the weighted sum data rate is higher with joint transmission than with separate transmission if conditions (7.9) are fulfilled.
We now extend our results to subcarrier allocation for OFDMA with L_{SC} subcarriers. The aim of the proposed subcarrier allocation is to assign on the same subcarrier only the users that verify conditions (7.9) on each pair. For that purpose, the coloration of the interference graph deduced from conditions (7.9) must be determined. The interference matrix $\mathbf{F}$ representing the interference graph is built as follows: $(\mathbf{F})_{(k,n)} = 1$ if $A_{k,n} < 0$, $A_{n,k} < 0$, $\frac{w_k}{w_n} < m_n$ or $\frac{w_n}{w_k} < m_k$ (pairwise interference condition), or if k and n are different users served by the same BS (OFDMA intra-cell orthogonality condition). Else, $(\mathbf{F})_{(k,n)} = 0$. If vertex k and n are adjacent in the interference graph (i.e., $(\mathbf{F})_{(k,n)} = 1$), then their simultaneous transmission in the same subcarrier is forbidden, either because this would decrease the weighted sum throughput, or because they are orthogonal users belonging to the same cell.
Two examples with $N_{\text{BS}} = 2$ cells and 2 users per cell are given in (7.10). The users are

ordered as $[\{\text{BS }1, \text{user }1\}, \{\text{BS }1, \text{user }2\}, \{\text{BS }2, \text{user }1\}, \{\text{BS }2, \text{user }2\}]$. In $\mathbf{F}^1$, user 1 of BS 2 is highly interfered by both users of BS 1, and user 2 of BS 2 is not highly interfered by the users of BS 1. In $\mathbf{F}^2$, user 1 of BS 1 highly interferes both users of BS 2, whereas user 2 of BS 1 does not generate high interference on the users of BS 2.

$$\mathbf{F}^1 = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \qquad \mathbf{F}^2 = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} \tag{7.10}$$

The users that are allowed simultaneous transmission in the same subcarriers are obtained through graph coloring. We use the greedy heuristic DSATUR [103] to solve the NP-hard graph coloring problem. In order to limit the complexity, we consider the interference graph on the average channel (including propagation loss and shadowing), so the same graph coloring is used in all subcarriers. The subcarrier allocation method is detailed hereunder. It determines a user i with higher priority, and then only allocates subcarriers to the users that have the same color as i. The algorithm runs in parallel for each subcarrier $l \in \{1, ..., L_{\mathrm{SC}}\}$:

1. On each subcarrier l, order all users k of the network by descending value of weighted rate $w_k \log_2\left(1 + \mathrm{SNR}_k^l\right)$, where $\mathrm{SNR}_k^l = \frac{G_{k,k}^l P_{\max}}{N_0 L_{\mathrm{SC}}}$ is the SNR of user k in subcarrier l with equal power allocation. Then allocate subcarrier l to the user that maximizes $w_k \log_2\left(1 + \mathrm{SNR}_k^l\right)$. Let BS^* be the base station that serves that user.

2. For all $\mathrm{BS} \neq \mathrm{BS}^*$, list the set of allowed users: these are the users of BS that have the same color as user k. If the set of allowed users is not empty, allocate subcarrier l to the allowed user that maximizes the weighted rate $w_n \log_2\left(1 + \mathrm{SNR}_n^l\right)$ in that subcarrier.

With DSATUR, the complexity of graph-based subcarrier allocation is polynomial in $\mathcal{O}((N_{\mathrm{BS}}K)^3)$, where K is the number of users per cell. It should be noted that distributed graph coloring methods could be used instead of DSATUR for implementation in distributed networks. A review of the complexity of distributed graph coloring is provided in [104].

7.4.3 WSTM distributed subcarrier allocation

We here briefly detail a simple WSTM distributed subcarrier allocation, that is used to maximize the weighted sum throughput on each OFDMA cell in an uncoordinated way. Inter-cell interference is therefore not considered. On each cell, each subcarrier $l \in \{1, ..., L_{\mathrm{SC}}\}$ is allocated to the user k that maximizes the weighted rate $w_k \log_2\left(1 + \mathrm{SNR}_k^l\right)$. This method is preferred to dynamic subcarrier allocation, where each cell iteratively updates subcarrier allocation by considering the SINR instead of the SNR. Indeed, in OFDMA, dynamic subcarrier allocation may not converge to a stable state.

7.5 WSTM power control

7.5.1 Proposed method

In this section, we detail a 2-Phases power control method for solving the WSTM problem (7.2) when subcarrier allocation is set. Contrary to previous work [99] [100] [101], our proposed method is suitable for distributed implementation in all SINR regimes, and is adapted to multi-user OFDMA. It consists of two phases: first, identify the links that should be set to zero, and then, operate in high SINR regime with the remaining links.

Phase I: all users and subcarriers

In Phase I, each subcarrier is considered as an independent interference channel. The aim of Phase I is to evaluate which users should be set to zero on the interference channel. For that purpose, all initial power values are set to $P_{\max}$, the maximum power per cell. In order to account for the number of allocated subcarriers, we consider the weight per user and subcarrier: $\alpha_k = \frac{w_k}{l_{\mathrm{SC},k}}$, where $l_{\mathrm{SC},k}$ is the number of subcarriers allocated to user k.
Then the iterative process from Qiu and Chawla [99], adapted for the WSTM problem, is used in parallel on each subcarrier. Let Ω_l be the set of interfering users that are active in subcarrier l. The optimization problem on subcarrier $l \in \{1, ..., L_{\mathrm{SC}}\}$ is:

$$
\begin{aligned}
&\max_{\mathbf{P}^l} \prod_{k \in \Omega_l} \left(1 + \frac{G_{k,k}^l P_k^l}{I_k^l}\right)^{\alpha_k} && (7.11)\\
&\text{s. t. } P_k^l \in [P_{\min}, P_{\max}],\ \forall k \in \Omega_l
\end{aligned}
$$

where I_k^l is the noise plus interference received by user k in subcarrier l, and $\gamma_k^l = \frac{G_{k,k}^l P_k^l}{I_k^l}$ is the SINR of user k in subcarrier l. $P_{\min} \geq 0$ is the minimum power per subcarrier.
Problem(7.11) is not convex due to inter-cell interference. It is solved with the following iterative method:

Initialization: at iteration $T = 0$, $\forall k \in \Omega_l$, set $P_k^l(0) = P_{\max}$, and compute the corresponding interference information: $\zeta_k^l(0) = \frac{\alpha_k \gamma_k^l(0)}{I_k^l(0)\left(1+\gamma_k^l(0)\right)}$

Iterative process:

1. Power update: Compute the power of each user $k \in \Omega_l$, depending on its channel state, weight, SINR, and on the interference information of the previous iteration, according to the following equations:

$$
\begin{aligned}
&P_k^l(T+1) = P_{\min},\ \text{if } X_k^l(\mathbf{P}^l(T)) \leq P_{\min}\\
&P_k^l(T+1) = P_{\max},\ \text{if } X_k^l(\mathbf{P}^l(T)) \geq P_{\max}\\
&P_k^l(T+1) = X_k^l(\mathbf{P}^l(T)),\ \text{otherwise} && (7.12)
\end{aligned}
$$

where

$$
X_k^l(\mathbf{P}^l(T)) = \frac{\alpha_k \gamma_k^l(T)}{\left(1+\gamma_k^l(T)\right)\left(\sum_{n \neq k} G_{k,n}^l \zeta_n^l(T)\right)} \tag{7.13}
$$

2. Interference information update: For each user $k \in \Omega_l$, deduce $\gamma_k^l(T+1)$ and $I_k^l(T+1)$ from the power values. Then compute the interference information depending on its weight, SINR, and received noise plus interference,

$$\zeta_k^l(T+1) = \frac{\alpha_k \gamma_k^l(T+1)}{I_k^l(T+1)\left(1+\gamma_k^l(T+1)\right)} \tag{7.14}$$

The function to be maximized in (7.11) is positive, continuous, differentiable, and defined over a compact set $\mathcal{S} = \left\{\mathbf{P}^l : \mathbf{P}_{\min} \leq \mathbf{P}^l \leq \mathbf{P}_{\max}\right\}$. Therefore, it has a global maximum on $\mathcal{S}$. Any fixed point of the iterative process converges to an optimum of this non convex optimization problem (7.11). The proof is detailed in [99], and comes from the continuity and differentiability of the studied function. As there may be several local optima, we cannot be sure that the iterative process will converge to a good solution. This will depend on the set of initial power values. However, numerical results show that, when starting from $P_{\max}$ for all users, if $P_{\min} = 0$, the users whose power tends to 0 reach this value in few iterations.

High SINR condition

At the end of Phase I, a test is used to check whether each subcarrier fulfills the high SINR condition, under a given precision. Only the subcarriers and users that fulfill that condition should be considered in Phase II.
In subcarrier l, the high SINR condition under precision β is fulfilled for user $k \in \Omega_l$ if and only if:

$$\left\|\alpha_k \log_2(1+\gamma_k^l) - \alpha_k \log_2(\gamma_k^l)\right\| \leq \beta$$

This is equivalent to

$$\frac{\alpha_k}{\log(2)} \log\left(1+\frac{1}{\gamma_k^l}\right) \leq \beta$$

$$\Leftrightarrow \gamma_k^l \geq \frac{1}{\left(2^{\frac{\beta}{\alpha_k}}-1\right)} \tag{7.15}$$

If this condition is fulfilled, P_k^l is set to $P_{\max}$. Else, user k is no longer considered active on subcarrier l and is removed from Ω_l.

Phase II: high SINR users and subcarriers

This phase is an adaptation of the Dual Asynchronous Distributed Pricing (DADP) algorithm for the multi-channel case from Huang et al. [101], where we introduce an additional condition on the power in the iterative process, in order to ensure that the high SINR condition is always verified. The optimization problem in high SINR regime is (with the approximation $\log(1+\mathrm{SINR}) \approx \log(\mathrm{SINR})$):

$$\begin{aligned}
&\max_{\mathbf{P}} \prod_{l=1}^{L_{\mathrm{SC}}} \prod_{k\in\Omega_l} \left(\frac{G_{k,k}^l P_k^l}{I_k^l}\right)^{\alpha_k} \\
&\text{s. t. } \sum_{l=1}^{L_{\mathrm{SC}}} P_{n_{\mathrm{BS}}}^l \leq P_{\max}, \forall n_{\mathrm{BS}} \in \{1, ..., N_{\mathrm{BS}}\} \\
&\text{and } P_k^l \geq P_{k,\min}^l, \forall l \in \{1, ..., L_{\mathrm{SC}}\}, \forall k \in \Omega_l
\end{aligned} \tag{7.16}$$

where $P^l_{k,\min} > 0$ is the minimum allowed power for user k in subcarrier l.
The sum power constraint is relaxed by introducing a dual price per BS, $\mu_{n_{\mathrm{BS}}}$. In the dual space, the initial problem is separated into L_{SC} problems, one per subcarrier. The iterative algorithm performs power control on each subcarrier independently, taking into account the dual prices. The dual prices are then updated, depending on whether the sum power constraint per BS is fulfilled or not. κ is the step for the dual price evaluation. The following iterative algorithm is used:

Initialization: at $T_d = 0$, set the initial power and price for all subcarriers l and all users $k \in \Omega_l$, and the initial dual price per BS $\mu_{n_{\mathrm{BS}}}(0) \geq 0$, $\forall n_{\mathrm{BS}} \in \{1, ..., N_{\mathrm{BS}}\}$.

Iterative process:

1. Dual price update: at each iteration T_d, each BS $n_{\mathrm{BS}} \in \{1, ..., N_{\mathrm{BS}}\}$ updates its dual price according to:

$$\mu_{n_{\mathrm{BS}}}(T_d) = \max \left\{ \mu_{n_{\mathrm{BS}}}(T_d - 1) + \kappa \left(\sum_{l=1}^{L_{\mathrm{SC}}} P^l_{n_{\mathrm{BS}}}(T_d - 1) - P_{\max} \right), 0 \right\} \tag{7.17}$$

2. Iterative power and interference information update: for a given dual price setting, an iterative process is used independently on each subcarrier $l \in \{1, ..., L_{\mathrm{SC}}\}$.
Initialization: at iteration $T = 0$, $\forall k \in \Omega_l$, set $P^l_k(0) = P_{\max}$, and compute the corresponding interference information: $\zeta^l_k(0) = \frac{\alpha_k}{I^l_k(0)}$
Iterative process:

   (a) Power update: Compute the power of each user $k \in \Omega_l$, depending on its channel state, weight, the interference information of the previous iteration, and on the dual price of its serving BS $n_{\mathrm{BS}}[k]$, $\mu_{n_{\mathrm{BS}}[k]}(T_d)$:

$$\begin{aligned} &P^l_k(T+1) = P^l_{k,\min}(T+1), \text{ if } Y^l_k(\mathbf{P}^l(T)) \leq P^l_{k,\min}(T+1) \\ &P^l_k(T+1) = P_{\max}, \text{ if } Y^l_k(\mathbf{P}^l(T)) \geq P_{\max} \\ &P^l_k(T+1) = Y^l_k(\mathbf{P}^l(T)), \text{ otherwise} \end{aligned} \tag{7.18}$$

   where

$$Y^l_k(\mathbf{P}^l(T)) = \frac{\alpha_k}{\left(\sum_{n \neq k} G^l_{k,n} \zeta^l_n(T) + \mu_{n_{\mathrm{BS}}[k]}(T_d) \right)} \tag{7.19}$$

   (b) Interference information update: Compute the interference information of each user $k \in \Omega_l$ depending on its weight and received noise plus interference,

$$\zeta^l_k(T+1) = \frac{\alpha_k}{I^l_k(T+1)} \tag{7.20}$$

At power update step, $P^l_{k,\min}(T+1)$ is set to a specific value that ensures that user k in subcarrier l will always fulfill the high SINR condition:

$$P^l_{k,\min}(T+1) = \frac{1}{\left(2^{\frac{\beta}{\alpha_k}} - 1\right)} \frac{I^l_k(T)}{G^l_{k,k}} \tag{7.21}$$

When the adaptive constraint $P_k^l \geq P_{k,\min}^l$ is not taken into consideration, and when $P_{\min}$ is fixed to a strictly positive value instead, this algorithm converges to the global optimum of problem (7.16). The convergence proof is similar to the one used for multi-channel DADP in [101]. It is based on the fact that the studied optimization problem belongs to the class of geometric programming [25], and can consequently be solved by Lagrangian relaxation. The dual price update can be viewed as a distributed gradient projection algorithm for solving the master problem. This algorithm converges for small enough step size κ. The proof of [101] applies to our case even if, contrary to Huang et al., we consider several users per cell, as the users served by the same BS are only differentiated by their channel gains, and the sum power constraint applies to the whole cell.
Introducing the adaptive constraint $P_k^l \geq P_{k,\min}^l$ into the optimization problem (7.16), instead of $P_k^l \geq P_{\min}$ as in [101], may question the convergence of the iterative process. However, numerical simulations have shown that, when the high SINR precision β is accurately set, $P_{k,\min}^l$ rapidly converges to a fixed value per user and subcarrier, and under this condition, the convergence proof from [101] remains valid.
The complexity of the 2-Phases power control depends on the second phase. It is polynomial in N_{BS}: $\mathcal{O}(L_{\mathrm{SC}}(N_{\mathrm{BS}})^2)$.

7.5.2 Graph-based subcarrier allocation and power control

The graph-based subcarrier allocation process determined in Section 7.4.2 may be followed by power control. As the users that are allowed to simultaneously transmit in the same subcarrier are not highly interfering each other, we can make the assumption that these users fulfill the high SINR condition. Consequently, we only use the second phase of the proposed power control to maximize the weighted sum throughput. In the numerical assessments, we will also evaluate the performance of graph-based subcarrier allocation followed by equal power allocation, where each active subcarrier gets an equal share of $P_{\max}$. In both cases, the complexity is in $\mathcal{O}((N_{\mathrm{BS}}K)^3)$.

7.5.3 Assessment of distributed power control

In order to evaluate the relevance of our proposed 2-Phases power control, we compare it with the optimum power control, and also with the binary power allocation from [102]. These two methods are centralized. The optimum power control is obtained by running the algorithm corresponding to the the first phase of our algorithm, with several different initial states. Indeed, this algorithm converges to a local optimum, that depends on its set of initial values. An exhaustive search through all possible local optima provides the global optimum. Similarly, the solution of binary power allocation is obtained through an exhaustive search, by testing all possible binary states.
We consider a network composed of two rings of interfering cells, with one subcarrier and one user per cell. The path loss model is Okumura-Hata with Rayleigh fast fading, and the shadowing's standard deviation is 7 dB. The thermal noise spectral density is $N_0 = -174$ dBm/Hz. The maximum transmit power per BS is $P_{\max} = 43$ dBm. The inter-site distance is $\mathrm{d_{is}} = 0.61$ km.
In this section, we set fixed values for the weight of each user. As there is only one user, we equivalently refer to it as the 'cell's weight'. Three cases are considered. In Case 1, all the

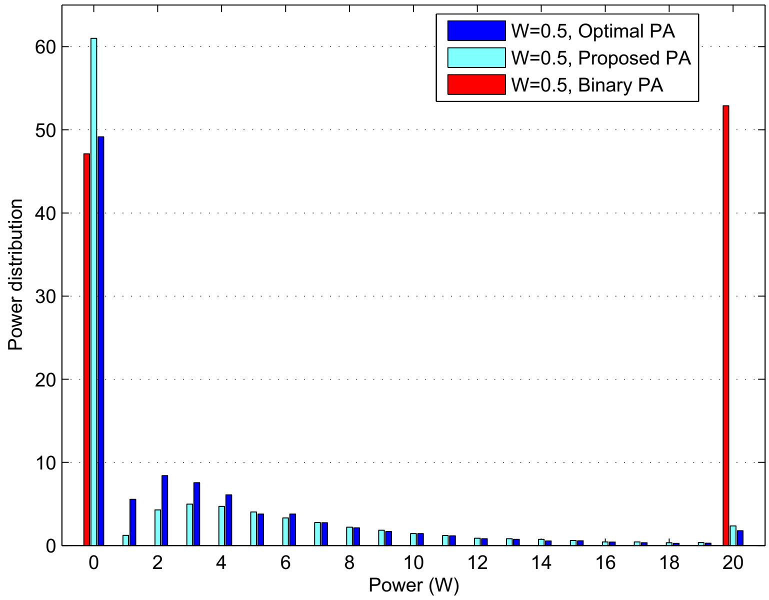


Figure 7.3: Distribution of power per cell, Case 1

cells have the same weight, $W = 0.5$. In Case 2, the even cells have a weight $W_1 = 0.67$, and the odd cells have a weight $W_2 = 0.33$. Finally in Case 3, the even cells have a weight $W_1 = 0.75$, and the odd cells have a weight $W_2 = 0.25$.
A first numerical assessment shows that the high SINR criterion $\beta = 0.15$ is a good compromise. Indeed, our aim is to minimize the Kullback-Leibler distance on the power, that measures the differences between two probability density functions, p_1 and p_2,

$$D(p_1, p_2) = \sum_k p_1(k) \log \left(\frac{p_1(k)}{p_2(k)} \right)$$

The power histograms for Cases 1 and 2 are represented on Fig. 7.3 and 7.4. The sub-optimal algorithm provides a good match with the optimal algorithm. Binary power allocation is

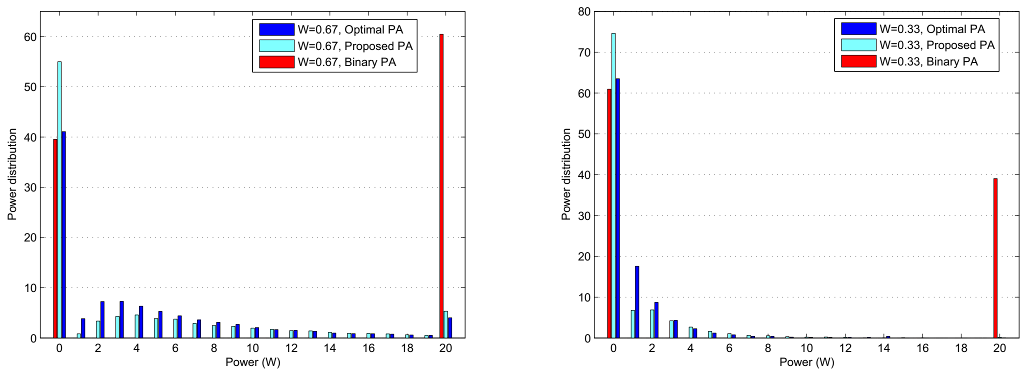


Figure 7.4: Distribution of power per cell, Case 2, $W_1 = 0.67$ (right) and $W_2 = 0.33$ (left)

far from the optimum. We then compare the performance results obtained with the optimal centralized algorithm and with binary allocation, with this value of β.

Table 7.2: Comparison of proposed power control with optimum power control and binary power allocation

	Case 1	Case 2 W=0.67	Case 2 W=0.33	Case 3 W=0.75	Case 3 W=0.25
Average capacity (bits/s/Hz)					
Optimum PC	2.61	3.37	1.45	3.56	1.05
Proposed PC	2.56	3.25	1.47	3.4	1.11
Relative decrease (%)	1.91	3.56	1.38	4.5	5.71
Binary PA	2.41	2.84	1.7	2.99	1.36
Relative decrease (%)	7.66	18.67	-17.24	16	-22.8
Average Power (W)					
Optimum PC	2.61	3.92	0.64	4.12	0.28
Proposed PC	2.61	3.79	0.71	3.94	0.35
Relative increase (%)	0	-3.32	10.9	-4.37	25
Binary PA	10.55	12.07	7.79	12.4	6.51
Relative increase (%)	304	218	1117	200	2225

Table 7.2 shows that the proposed power control algorithm leads to almost the same average capacity as the optimal algorithm. The average power required on each type of cell is also quite similar, with a small decrease in the power required for the users with highest weight, and an increase in the power required for the users with lowest weight. This is due to some inaccuracies at the end of Phase I: some users get to Phase II, although they should be rejected. We can notice that binary power allocation is less accurate than the proposed sub-optimal algorithm regarding capacity, and that it consumes too much power in all cases. Binary power allocation also increases the capacity of the users with lowest weight, which is useless with respect to the WSTM objective.

7.6 Numerical results

In this section, four methods are numerically compared: graph-based subcarrier allocation without power control (referred to as 'Graph-Based SC, NoPC'), graph-based subcarrier allocation with high SINR power control ('Graph-Based SC + PC'), distributed WSTM subcarrier allocation followed by power control in 2 phases ('2Phases PC'), and distributed WSTM subcarrier allocation followed by binary allocation ('Binary PA'). Binary PA selects the binary combination of powers $P_n^l = \left\{0, \frac{P_{\max}}{L_{\text{SC}}}\right\}$ for all cells $n \in \{1, ..., N_{\text{BS}}\}$, that maximizes the weighted sum data rate, on each subcarrier l. The complexity of this method is exponential in N_{BS} : $\mathcal{O}(L_{\mathrm{SC}} 2^{N_{\mathrm{BS}}})$.

7.6.1 Two cells, TDMA

We first consider the case of two interfering cells, one subcarrier and one user per cell. The weight of each user is set to a fixed value. We compare the optimum resource allocation (obtained with the analytical solution derived in Section 7.4.1) with the four methods, in terms of achieved weighted sum data rate, by using Monte-Carlo simulations. The sum weight is $w_1 + w_2 = 1$. The location of users is uniformly distributed within each cell. The other parameters are the same as in Section 7.5.3.
Fig. 7.5 represents the relative weighted sum data rate loss between the optimum weighted sum data rate and the four other methods. The optimum weighted sum data rate is between 70590 kbits/s for $w_1 = 0.5$ and 112958 kbits/s for $w_1 = 1$, and increases with w_1. The relative performance loss is always lower than 0.77% with Graph-Based SC + PC. This method is the closest to the optimum resource allocation.

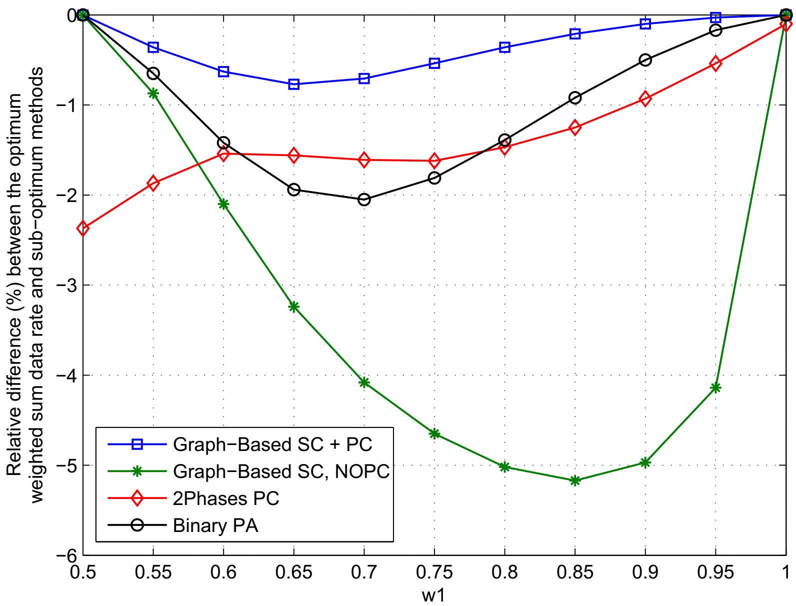


Figure 7.5: 2 cells, TDMA, relative difference to the optimum for weighted sum data rate

7.6.2 Dynamic simulations, 7 cells, multi-user OFDMA

We now consider a network composed of one central cell BS_0, and one ring of 6 interfering BSs, with two non-adjacent sets, $\mathcal{C}_1 = \{\mathrm{BS}_1, \mathrm{BS}_3, \mathrm{BS}_5\}$ and $\mathcal{C}_2 = \{\mathrm{BS}_2, \mathrm{BS}_4, \mathrm{BS}_6\}$. The FFT size and the number of available subcarriers per BS are $L_{\mathrm{SC}} = 256$. The location of users in each cell follows a uniform distribution and does not vary in a given Monte-Carlo snapshot. The other network characteristics are the same as in Section 7.5.3.
We study a dynamic scenario where the weight of each user is equal to its normalized queue

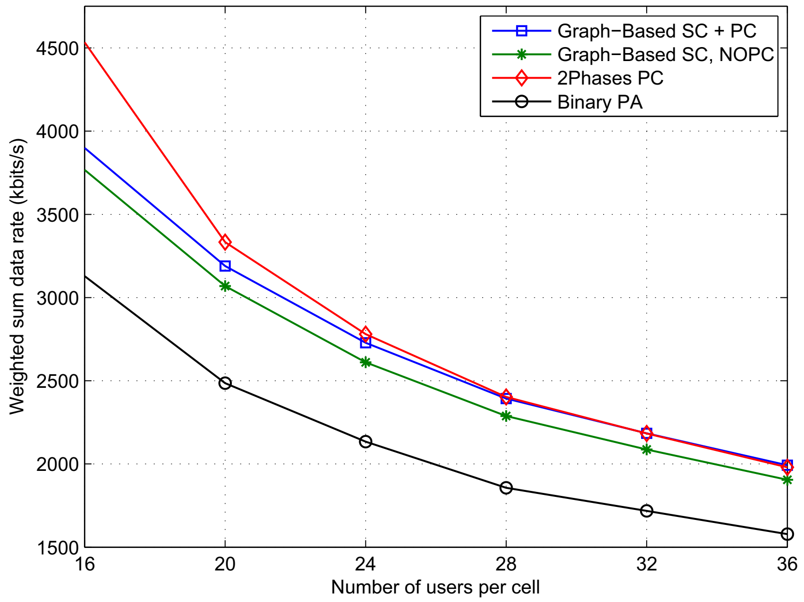


Figure 7.6: 7 cells, OFDMA, weighted sum data rate

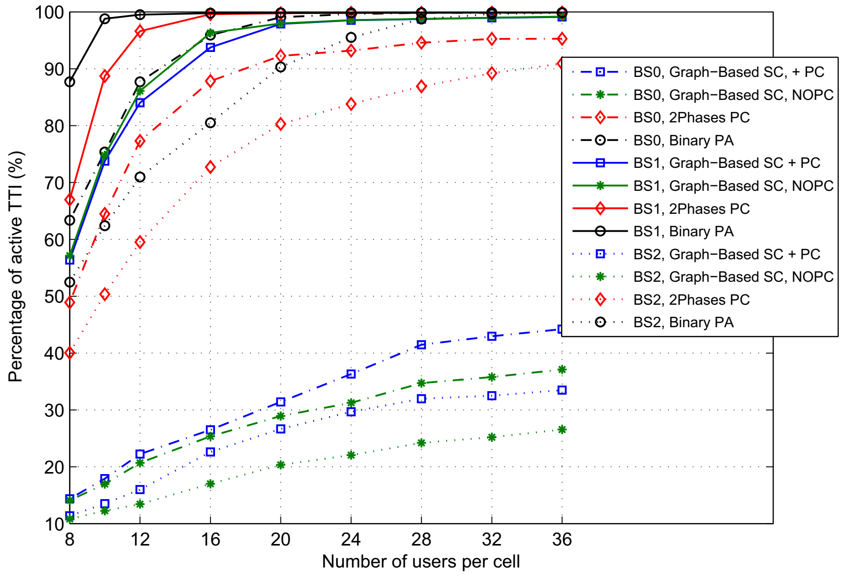


Figure 7.7: 7 cells, OFDMA, percentage of active TTIs

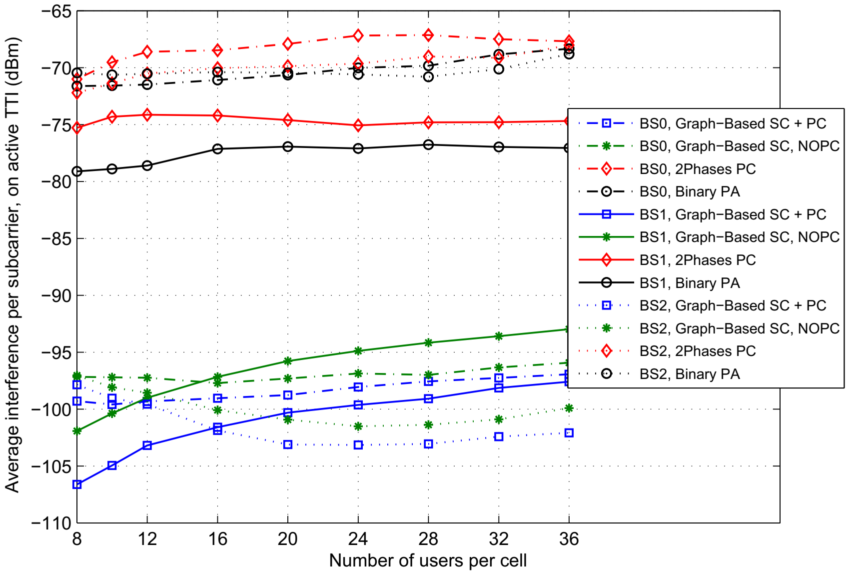


Figure 7.8: 7 cells, OFDMA, average inter-cell interference per subcarrier, per active TTI

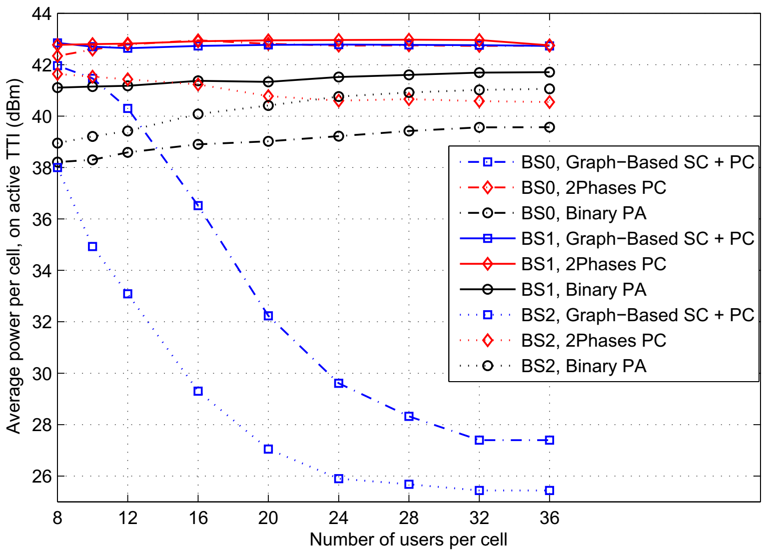


Figure 7.9: 7 cells, OFDMA, average power per cell, per active TTI

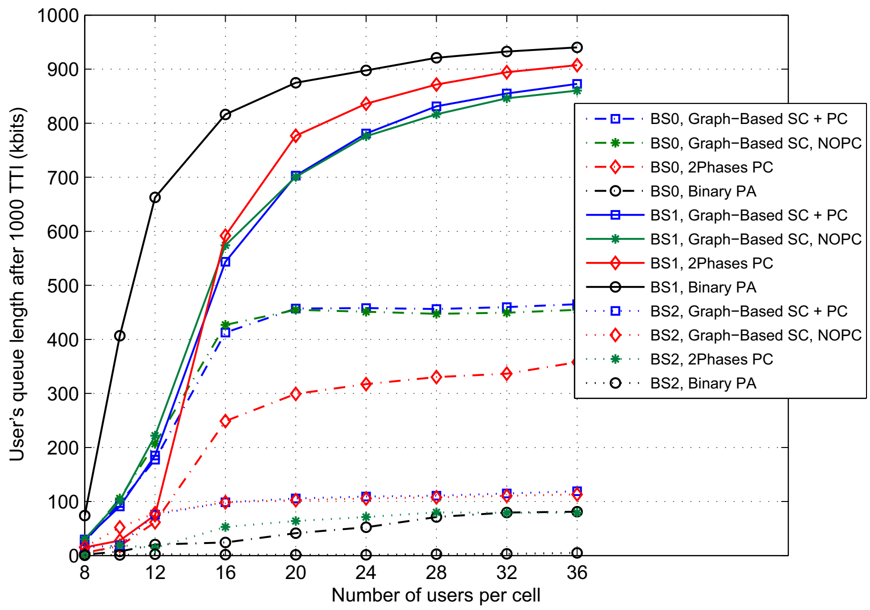


Figure 7.10: 7 cells, OFDMA, average queue length per user

length. Each user receives data in its queue with a Poisson traffic model, where the inter-arrival law follows an exponential law of average equal to $T_i = 20$ TTIs. The TTI duration is 2 ms. The packet size follows a log-normal law with average 2.5 kbits for the users of BS_0, 5 kbits for the users of $\mathcal{C}_1$, and 1.25 kbits for the users of $\mathcal{C}_2$. The packet size standard deviation is equal to 0.1 kbits. The queue length is limited to 1024 kbits. Users are in buffer overflow when their queue length reaches this value.

The weighted sum throughput on all cells is represented on Fig. 7.6. At low to medium load, 2Phases PC leads to higher weighted sum throughput than the other methods. At very low load (less than 8 users per cell), Binary PA is even better than Graph-Based methods. Indeed, in that case, it is not necessary to take into account the interfering users in the subcarrier allocation process, and it is more effective to assign all the subcarriers to the users with highest weighted SNR in each cell. However, at medium to high load, inter-cell interference limits the achievable data rate, and it is more efficient to allocate orthogonal subcarriers with graph-based subcarrier allocation, in order to avoid interference situations, rather than scheduling users simultaneously as in Binary PA. 2Phases PC determines which users should be set to zero, and consequently leads to high weighted sum throughput at any load.

The four methods are compared in terms of resource consumption on Fig. 7.7, 7.8 and 7.9. The percentage of active TTIs is equal to the ratio of TTIs where at least one user is active. The performance results for three BSs with different load levels are represented. With Graph-Based SC, the percentage of active TTIs is limited to 45% for the two least-loaded cells (BS_0 and BS_2), whereas it goes up to 90% with 2Phases PC, and to 95% with Binary PA. Graph-Based SC enables the users of these cells to efficiently empty their queues whenever they access to the resource, whereas the two other methods lead to useless resource access. This is due to the interference levels: with Graph-Based SC, the average inter-cell interference is between 20 and 30 dBm lower than with Binary PA and 2Phases PC. Therefore, 2Phases PC is less favorable than Graph-Based SC, although it leads to the same weighted sum throughput at high load. Power control after Graph-Based SC allows additional inter-cell interference limitations, as well as a power decrease. In that case, the power of the users with the lowest weights (users of BS_2) is decreased, in order to limit the interference generated to the users with highest weights (users of BS_1), that need higher data rates to empty their queues.

Graph-Based SC and 2Phases PC favor the users of BS_1 that have the highest queue lengths, at the cost of an increase in the queue lengths of the other users (Fig. 7.10). The lowest queue lengths for users of BS_1 (and also for all users in the network) is obtained with 2Phases PC at low load, and with Graph-Based SC at medium to high load. These results show that the WSTM objective is indeed valid to balance the users' queue lengths, and that Binary PA is not efficient to address the WSTM objective.

7.7 Conclusion

This chapter has investigated resource allocation for the WSTM problem in SISO multi-cell OFDMA networks. A graph-based subcarrier allocation and a distributed power control have been proposed. These methods may be combined, or used independently. In the two cells, TDMA case, graph-based subcarrier allocation followed by high SINR power control leads to the lowest relative performance loss, which is almost negligible compared to the optimum resource allocation. For OFDMA with 7 cells and where each user's weight is equal to its normalized queue length, the proposed methods are all more efficient in terms of WTSM than binary allocation, at medium to high load. Graph-based subcarrier allocation is very effective to decrease resource consumption, leading to lower power and interference levels.

The results from this chapter show that inter-cell interference needs to be treated for BE users in order to avoid buffer overflows. It may be treated directly through subcarrier allocation, by avoiding joint transmission of users that would be interfering too much. It should be noted that our proposed method can be viewed both as dynamic subcarrier allocation and as time-based scheduling. In both cases, the objective is to perform interference alignment whenever inter-cell interference becomes the limiting feature to achieve the optimization goal. This restriction of interference alignment to useful cases is a major difference with the work from [81], where time-based interference alignment is performed in all cases. A possible extension of our work would therefore be a theoretically study of dynamic interference alignment, not only for WSTM, but also for other optimization objectives.

Inter-cell interference may also be mitigated via power control. This method is particularly needed whenever several potentially highly interfering users request simultaneous transmission. In that case, we have seen that decreasing the transmit power in order to mitigate the interference generated to other users is globally efficient for all users. We may thus conclude that, for BE users as well as for RC users, the global network performance is improved when each cell not only considers the interference that it receives, but also the interference that it generates to the other cells. For RC users, this conclusion directly stems from the network-wide power convergence constraint on each subcarrier. For BE users, we cannot formalize the interdependences of power levels so directly, but the conclusions are quite close if the objective of resource allocation is to avoid buffer overflows for all users. In multi-cell networks, selfish resource allocation, performed within each BS irrespectively of the others, leads to network performance degradations. We have seen that distributed resource allocation with consideration of the other BSs is feasible, both for RC and BE users, at the cost of sub-optimality with respect to centralized resource allocation, and possibly of small information exchanges between the BSs.

Chapter 8

Conclusions and Perspectives

8.1 Conclusions

This thesis has studied distributed resource allocation in OFDMA multi-cell networks. Our aim was to determine methods, distributed per BS, that provide the required QoS to all users, whatever their location in the cell is. Subcarrier allocation and iterative power allocation have been investigated in several cases.
First, we have proposed a network coordination method adapted to distributed networks, and the complete resource allocation strategy for this type of causal coordination. Network coordination is triggered for the users located at the border of cells, that suffer the most from inter-cell interference. Four power control objectives with varying fairness levels have been tested. It has been shown that network coordination brings additional fairness at any load with all power allocation objectives, and increases the system throughput with fair objectives. Network coordination however decreases the peak data rate when the power allocation objective is unfair, as the data rate gain brought to users at cell edge is obtained at the expense of the users in the best radio conditions. The data rates of coordinated users are increased both thanks to the virtual MIMO diversity gain and to an inter-cell interference decrease, as the most potential interferer becomes the coordinated BS.
Based on these conclusions, we have then determined a resource allocation method using network coordination, that depends on the QoS constraint of each user. BE users follow an unfair, Globally Optimal resource allocation objective, whereas RC users follow a fair MA resource allocation objective. The proposed resource allocation algorithm aims at minimizing the sum power required for RC users to reach their target data rate, so that the sum data rate of BE users is maximized. It favors RC users over BE users, and adapts both subcarrier and power allocations to each QoS objective, including network coordination for the RC users located at the border of cells. The proposed method is far more efficient than the methods previously proposed in the literature. The conjugate effects of increasing the capacity via network coordination and prioritizing resource allocation between RC and

BE users lead to power and rejection rate decreases for RC users, and to a sum data rate increase for BE users.

In the remainder of the dissertation, we have focused on fully distributed networks. We first considered RC users, that follow a MA resource allocation problem, i.e., minimization of the sum power required to reach a target data rate per user. This problem may not be feasible due to inter-cell interference in multi-cell networks. We have determined a distributed convergence criterion for power control on the interference channel with SISO transmission. This criterion has been included into a distributed resource allocation method. It is used as a maximization objective in subcarrier allocation, and is considered within power allocation. Power control sets the SINR per user and subcarrier, in order to both guarantee power convergence, and reach the target data rate. The convergence criterion is triggered depending on the interference level. The proposed method is far more efficient than iterative water-filling, as it does not lead to any power divergence situation. At medium to high load, it thus increases the percentage of users fulfilling their QoS constraint, with far lower inter-cell interference levels. It is therefore a relevant alternative to iterative water-filling in distributed networks.

The MA resource allocation problem has then been studied in MIMO communication systems, both when full CSI is available at transmitter, and when only the statistical properties of the channel are available at transmitter. In both cases, a convergence criterion for power control has been determined. In the full CSIT case, water-filling is performed over the subcarriers and streams with the aim to reach users' target data rates. The convergence criterion corresponds to a maximum allowed data rate per subcarrier. In the statistical CSIT case, we have derived an approximate analytical expression of the outage capacity as a function of the SNR and of the outage probability. As this analytical expression is concave in the SNR, the MA problem is solved via convex optimization. The convergence criterion for power control is included within power control. In both cases, the proposed methods avoid power divergence situations at medium to high load, compared to iterative water-filling. They consequently lead to more RC users reaching their target data rate, and lower inter-cell interference levels.

Finally, we have investigated resource allocation for BE users in SISO. A possible QoS objective for these users is to avoid buffer overflows as much as possible. We have studied the WSTM problem. If the weight of each user is proportional to its queue length, then this objective is a trade-off between maximizing the sum data rate, and balancing the users' queue lengths. The WSTM problem has first been addressed via subcarrier allocation, by studying the capacity region in the two interfering links' case. A simple characterization of whether simultaneous transmission leads to higher weighted sum throughput than in-sequence transmission has been obtained. Based on this characterization, a graph-based subcarrier allocation method for multi-cell OFDMA has been derived. We have also proposed a distributed power control method for WSTM. At medium to high load, the proposed methods, whether used jointly or independently, lead to lower average queue lengths, and fairer queue lengths' distribution, than binary allocation. Graph-based subcarrier allocation is particularly useful to decrease power consumption.

8.2 Future work

This thesis has determined several distributed resource allocation methods for RC and BE users in OFDMA multi-cell networks. These are first steps that should be complemented by additional studies. We hereunder provide first insights into open issues that have not been addressed in the thesis, and that could be studied as future work:

- *Influence of the different simplifying assumptions taken throughout the dissertation*:
  The proposed methods assume that CSI is fully known at transmitter, that the channel remains stable during one TTI, and that all transmissions are synchronized. It would be of interest to adapt them to more realistic conditions. The influence of statistical CSIT has been studied for the MIMO channel in Chapter 6, and the analytical expression of the outage capacity as a function of the SNR obtained in this chapter could be adapted to the network coordination method of Chapters 3 and 4.

- *Extension of causal network coordination*:
  Network coordination involving more than two BSs and using other protocols than data forwarding could be investigated. Besides, in Chapters 3 and 4, we have considered that the inter-BS transmission was perfect in terms of SNR. This assumption is valid in the case of wired links, or directional radio links, but not for ad hoc networks. In that case, the transmission protocol should be based on DF. Besides, it would also be of interest to evalute the relevance of causal network coordination when the same air interface is used on the link between the coordinated transmitters as on the downlink.

- *Theoretical convergence of the proposed distributed methods*:
  The network-wide optimization objective has been decomposed into per BS subproblems, due to the constraint of distributed resource allocation. Inter-cell interference has been taken into consideration by iterating power allocation over the BSs. The theoretical convergence of the proposed algorithms has not been proved in all cases. Besides, even when the power control algorithms were shown to be convergent (as in Chapters 5, 7, and 6 for the statistical CSIT case), the final state may be a local optimum of the network-wide optimization problem. Most proposed methods belong to the set of alternating minimization algorithms. The convergence of alternating minimization algorithms has been studied in the static case in [105], and in the dynamic case, if the channel parameters are evolving during the iterative process, in [106]. These references may serve as bases to fully characterize the theoretical convergence conditions of our distributed methods.

- *Interference channel from an information-theoretic and signal processing point of view*:
  Many open issues remain on the interference channel, contrary to the multi-access and broadcast channels. In this thesis, the main focus was not on studying the interference channel, but rather on adapting resource allocation to it. Throughout the dissertation, we have considered inter-cell interference as noise. The methods determined in Chapters 5, 6 and 7 characterize the situations when neglecting inter-cell interference is optimal (thus allowing joint transmission of interfering links), and, on the contrary, when inter-cell interference should be removed. These two states could be complemented by additional transmission states, implying that inter-cell interference is not

treated as noise, as explained in [21].

The interference channel could also be studied from a cross-layer perspective, in order to evaluate whether WSTM with the weight of each user proportional to its queue length is throughput optimal. Finally, interference alignment could also be studied, both for the MIMO case, when interference alignment is performed on MIMO streams [79], and for the general case, when interference alignment is performed in time domain [81] or in frequency domain [80]. MIMO interference alignment could be compared to our results on MIMO transmission with full CSIT, so as to determine under which conditions interference alignment is more efficient than iterative water-filling with power control convergence constraints. More generally, we could formalize the various characterizations of joint transmission's limitations within the scope of interference alignment theory.

Bibliography

[1] IEEE STD 802.16e 2005, *IEEE Standard for Local and Metropolitan Area Networks, part 16: Air Interface for Fixed and Mobile Broadband Wireless systems*, Dec. 2005.

[2] H. Ekstrom, A. Furuskar, J. Karlsson, M. Meyer, S. Parkvall, J. Torsner, and M. Wahlqvist, "Technical solutions for the 3G Long-Term Evolution," *IEEE Commun. Magazine*, vol. 44, no. 3, pp. 38–45, Mar. 2006.

[3] 3GPP TS 25.331, *3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Radio Resource Control (RRC) protocol specification (Release 1999), v3.21.0*, Dec. 2004.

[4] H. Holma and A. Toskala, *WCDMA for UMTS*, Wiley, 2000.

[5] 3GPP TR 25.858, *High Speed Downlink Packet Access, Physical layer aspects (Release 5), v5.0.0*, Mar. 2002.

[6] 3GPP TR 25.899, *High Speed Downlink Packet Access (HSDPA) enhancements (Release 6), v6.1.0.*, Sept. 2004.

[7] 3GPP TS 36.300, *E-UTRA and E-UTRAN: Overall Description: Stage 2 (Release 8), v6.1.0*, Sept. 2004.

[8] T. Bonald, L. Massoulie, A. Proutiere, and J. Virtamo, "A queueing analysis of max-min fairness, proportional fairness and balanced fairness," *Queueing Systems*, vol. 53, no. 1, pp. 65–84, June 2006.

[9] J. Heyman, *Intercell interference management in OFDM based downlink*, Ph.D. thesis, Linköping Universitet, Linköping, 2006.

[10] A. Goldsmith and S-G. Chua, "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Commun.*, vol. 45, no. 10, pp. 1218–1230, Oct. 1997.

[11] G. Foschini and M.J. Gabs, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.

[12] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov.-Dec. 1999.

[13] T.M Cover and J.A. Thomas, *Elements of information theory*, Wiley, 1991.

[14] D. Tse and P. Viswanath, *Fundamentals of wireless communication*, Camdbridge University Press, 2005.

[15] N. Jindal, S. Vishwanath, and A. Goldsmith, "On the duality of Gaussian multiple access and broadcast channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 5, pp. 768–783, Jan. 2004.

[16] A.A.E. Gamal and M.H.M. Costa, "The capacity region of the discrete memoryless interference channel with strong interference," *IEEE Trans. Inf. Theory*, vol. 35, no. 5, pp. 710–711, Sept. 1987.

[17] M.H.M. Costa and A.A.E. Gamal, "The capacity region of a class of deterministic interference channel," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 49–60, Jan. 1982.

[18] R. Benzel, "The capacity region of a class of discrete additive degraded interference channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 2, pp. 228–231, March 1979.

[19] T.S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. 27, pp. 49–60, Jan. 1981.

[20] R. Etkin, D. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *Submitted to IEEE Trans. Inf. Theory, available on arXiv*, Feb. 2007.

[21] S.A. Jafar and S. Vishwanath, "Generalized degrees of freedom of the symmetric Gaussian K user interference channel," *Submitted, available on arXiv*, April 2008.

[22] Huawei, "Further analysis of soft frequency reuse scheme," in *3GPP TSG RAN WG1#42 R1-050841*, Sept. 2005.

[23] Ericsson, "Inter-cell interference handling for E-UTRA," in *3GPP TSG RAN WG1#42 R1-050764*, Sept. 2005.

[24] S. Boyd and L. Vanderbergue, *Convex Optimization*, Cambridge University Press, 2004.

[25] M. Chiang, *Geometric programming for communication systems*, Now Publishers (Fundations and trends in communications and information theory), 2005.

[26] M. Chiang, S.H. Low, A.R. Calderbank, and J.C. Doyle, "Layering as optimization decomposition: A mathematical theory of network architectures," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 255–312, Jan. 2007.

[27] M. Chiang, C.W. Tan, D.P. Palomar, and D. Julian, "Power control by geometric programming," *IEEE Trans. Wireless Commun.*, 2007.

[28] S. Barbarossa and G. Scutari, "Cooperative diversity through virtual arrays in multihop networks," in *Proc. of ICASSP*, Hong Kong, Apr. 2003.

[29] M. Dohler, *Virtual antenna arrays*, Ph.D. thesis, King's College London, London, 2003.

[30] A. Nosratinia, T. E. Hunter, and A. Hedayat, "Cooperative communications in wireless networks," *IEEE Commun. Magazine*, vol. 42, no. 10, pp. 74–80, Oct. 2004.

[31] A. Sendonaris, E. Erkip, and B. Aazhang, "Increasing uplink capacity via user cooperation diversity," in *Proc. IEEE Int. Symposium of Inf. Theory*, Cambridge, MA, USA, Aug. 1998.

[32] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity - Part I: System description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[33] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity - Part II: Implementation aspects and performance analysis," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1939–1948, Nov. 2003.

[34] J.N. Laneman, D.N. Tse, and G.W. Wornell, "An efficient protocol for realizing cooperation diversity in wireless networks," in *Proc. IEEE Int. Symposium of Inf. Theory*, Washington DC, USA, June 2000.

[35] J.N. Laneman, D.N. Tse, and G.W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[36] R.U. Nabar, H. Bölcskei, and F.W. Kneubühler, "Fading relay channels: performance limites and space-time signal design," *IEEE J. Select. Areas Commun.*, vol. 22, no. 6, pp. 1099–1109, Aug. 2004.

[37] B. Can, H. Yomo, and E. De Carvalho, "Hybrid forwarding scheme for cooperative relaying in OFDM based networks," in *Proc. IEEE Conf. on Commun.*, Istanbul, Turkey, June 2006.

[38] Q. Zhang, J. Zhang, C. Shao, Y. Wang, P. Zhang, and R. Hu, "Power allocation for regenerative relay channels with Rayleigh fading," in *Proc. IEEE Veh. Tech. Conf.*, Los Angeles, CA, USA, May 2004.

[39] I. Hammerstrom and A. Wittneben, "On the optimal power allocation for nonregenerative OFDM relay links," in *Proc. IEEE Conf. on Commun.*, Istanbul, Turkey, June 2006.

[40] Y. Li, B. Vucetic, Z. Zhendong, and M. Dohler, "Distributed adaptive power allocation for wireless relay networks," *IEEE Trans. Wireless Commun.*, vol. 6, no. 3, pp. 948–958, Mar. 2007.

[41] T. Irnich, D.C. Schultz, R. Pabst, and P. Wienert, "Capacity of relaying infrastructure for broadband radio coverage of urban areas," in *Proc. IEEE Veh. Tech. Conf*, Orlando, Fl, USA, Oct. 2003.

[42] J.G. Andrews, W. Choi, and R.W. Heath Jr., "Overcoming interference in spatial multiplexing MIMO cellular networks," *IEEE Wireless Communications*, vol. 14, no. 6, pp. 95–104, Dec. 2007.

[43] A. Goldsmith, S.A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Select. Areas Commun.*, vol. 21, no. 5, pp. 684–702, June 2003.

[44] M.K. Karakayali, G.J. Foschini, and R.A. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," *IEEE Wireless Communications*, vol. 13, no. 4, pp. 56–61, Aug. 2006.

[45] H. Zhang and H. Dai, "Co-channel interference mitigation and cooperative processing in downlink multicell multiuser MIMO networks," *Eur. Jour. Wireless Commun. and Networking*, vol. 2004, no. 2, pp. 222–235, Dec. 2007.

[46] S. Kaviani and W.A. Krzymien, "Sum rate maximization of MIMO broadcast channels with coordination of Base Stations," in *Proc. of Wireless Commun. and Networking Conf.*, Las Vegas, NV, USA, Mar. 2008.

[47] A. Papadogiannis, H. Bang, D. Gesbert, and E. Hardouin, "Downlink overhead reduction for multi-cell cooperative processing enabled wireless networks," in *Proc. of Personal, Indoor and Mobile Commun. Conf.*, Cannes, France, Sept. 2008.

[48] M. Vu and A. Paulraj, "Capacity optimization for Rician correlated MIMO wireless channels," in *Proc. Asilomar Conference*, Pacific Grove, DA, USA, Nov. 2005.

[49] H. Boche and M. Schubert, "Resource allocation in multiantenna systems- achieving max-min fairness by optimizing a sum of inverse SIR," *IEEE Trans. Signal Processing*, vol. 54, no. 6, pp. 1990–1997, June 2006.

[50] D. Julian, M. Chiang, D. O'Neill, and S. Boyd, "QoS and fairness constrained convex optimization and resource allocation for wireless cellular and ad hoc networks," in *Proc. IEEE Int. Conf. on Computer Communication*, New York, NY, USA, June 2002.

[51] M. Anas, K. Kim, S.Shin, and K. Kim, "QoS aware power allocation for combined guaranteed performance and best effort users in OFDMA systems," in *Proc. of IS-PACS*, Seoul, Korea, Nov. 2004.

[52] W. Rhee and J. M. Cioffi, "Increase in capacity of multiuser OFDM system using dynamic subchannel allocation," in *Proc. Veh. Tech. Conf.*, Tokyo, Japan, May 2000.

[53] J. Jang and KB. Lee, "Transmit power adaptation for multi-user OFDM systems," *IEEE J. Select. Areas Commun.*, vol. 21, no. 2, pp. 171–178, Feb. 2003.

[54] C.Y. Wong, R. S. Chen, K. Ben Letaief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE J. Select. Areas Commun.*, vol. 17, no. 10, pp. 1747–1785, Oct. 1999.

[55] I. Kim, I-S. Park, and Y.H.Lee, "Use of linear programming for dynamic subcarrier and bit allocation in multiuser OFDM," *IEEE Trans. Veh. Tech.*, vol. 55, no. 04, pp. 1195–1207, July 2006.

[56] I. Kim, H.L. Lee, B. Kim, and Y.H. Lee, "On the use of linear programming for dynamic subchannel and bit allocation in multi-user OFDM," in *Proc. of IEEE Globecom*, San Antonio, TX, USA, Nov. 2001.

[57] M. Ergen, S. Coleri, and P. Varaiya, "QoS aware adaptive resource allocation techniques for fair scheduling in OFDMA based broadcast wireless access systems," *IEEE Trans. on Broadcasting*, vol. 49, no. 4, pp. 362–370, Dec. 2003.

[58] K. Seong, M. Mohseni, and J.M.Cioffi, "Optimal resource allocation for OFDMA downlink systems," in *Proc. of IEEE Int. Symposium of Inf. Theory*, Seattle, WA, USA, July 2006.

[59] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. on Commun.*, vol. 54, no. 7, pp. 1310–1322, July 2006.

[60] Z. Shen, J.G. Andrews, and B.L. Evans, "Optimal power allocation in multiuser OFDM systems," in *Proc. of IEEE Globecom*, St Louis, MO, USA, Dec. 2003.

[61] D.D. Yu and J.M. Cioffi, "Iterative water-filling for optimal resource allocation in OFDM multiple access and broadcast channels," in *IEEE Proc. Globecom*, San Francisco, CA, USA, Nov. 2006.

[62] Z. Han, Z. Ji, and K.J.R. Liu, "Power minimization for multi-cell OFDM networks using distributed non-cooperative game approach," in *Proc. of IEEE Globecom*, Dallas, TX, USA, Dec. 2004.

[63] Z. Han, Z. Ji, and K.J.R. Liu, "Non-cooperative resource competition game by virtual referee in multi-cell OFDMA networks," *IEEE Jour. Select. Commun.*, vol. 25, no. 6, pp. 1079–1090, Aug. 2007.

[64] G. Kulkarni, S. Adladkha, and M. Srivastava, "Subcarrier allocation and bit loading algorithms for OFDMA-based wireless networks," *IEEE Trans. on Mobile Computing*, vol. 4, no. 6, pp. 652–662, Nov. 2005.

[65] S. Piertzyk and G.J.M. Janssen, "Radio Resource Allocation for cellular networks based on OFDMA with QoS guarantees," in *Proc. of IEEE Globecom*, Dallas, TX, USA, Dec. 2004.

[66] J. Zander, "Performance of optimum transmitter power control in cellular radio networks," *IEEE Trans. Veh. Technol.*, vol. 41, no. 1, pp. 57–62, Feb. 1992.

[67] J. Zander, "Distributed co-channel interference control in cellular radio systems," *IEEE Trans. Veh. Technol.*, vol. 41, no. 3, pp. 305–311, Aug. 1992.

[68] G.J. Foschini and Z. Miljanic, "A simple distributed autonomous power control algorithm and its convergence," *IEEE Trans. Veh. Technol.*, vol. 42, no. 4, pp. 641–646, Nov. 1993.

[69] R. Yates, "A framework for uplink power control in cellular radio systems," *IEEE Jour. Select. Commun.*, vol. 13, no. 7, pp. 1341–1348, Sept. 1995.

[70] Z. Han, Z. Ji, and K.J.R. Liu, "Joint link quality and power management over wireless networks with fairness constraint and space-time diversity," *IEEE Trans. Veh. Technol.*, vol. 53, no. 4, pp. 1138–1148, July 2004.

[71] H. Kelterer, V. Pferschy, and D. Pisinger, *Knapsack problems*, New York: Springer, 2004.

[72] C.D. Meyer, *Matrix analysis and applied linear algebra*, Society for Industrial and Applied Mathematics (SIAM), 2000.

[73] M. Mesbahi, M.G. Safonov, and G.P. Papavassilopoulos, "Bilinearity and complementarity in robust control," in *Proc. Advances in linear matrix inequality methods in control*, SIAM, Ed., 2000.

[74] B. Song, R.L. Cruz, and B.D. Rao, "Network duality for multiuser MIMO beamforming networks and applications," *IEEE Trans. Commun.*, vol. 55, no. 3, pp. 618–630, Mar. 2007.

[75] S. Shi, M. Schubert, and H. Boche, "Capacity Balancing for multiuser MIMO systems," in *Proc. Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Honolulu, Hawai, USA, Apr. 2007.

[76] S. Shi, M. Schubert, and H. Boche, "Weighted sum-rate optimization for multiuser MIMO systems," in *Proc. Conf. Inform., Sciences and Systems (CISS)*, Baltimore, MD, USA, Mar. 2007.

[77] J-H. Chang, L. Tassiulas, and F. Rashid-Farrokhi, "Joint transmitter receiver diversity for efficient space division multiaccess," *IEEE Trans. Wireless Commun.*, vol. 1, no. 1, pp. 16–27, Jan. 2002.

[78] D.P. Palomar, M.A. Lagunas, and J.M. Cioffi, "Optimum linear joint transmit-receive processing for MIMO channels with QoS constraints," *IEEE Trans. Signal Processing*, vol. 5, no. 5, pp. 1179–1197, May 2004.

[79] K. Gomadan, V.R. Cadambe, and S.A. Jafar, "Approaching the capacity of wireless network through distributed interference alignment," *Submitted, available on arXiv*, March 2008.

[80] V. Cadambe and S. Jafar, "Interference alignment and degrees of freedom for the K user interference channel," *IEEE Trans. Inf. Theory*, pp. 3245–3441, Aug. 2008.

[81] L. Grokop, D.N.C. Tse, and R.D. Yates, "Interference alignment for line-of-sight channels," *Submitted, available on arXiv*, Sept. 2008.

[82] M. Dohler and M. Arndt, "Inverse incomplete gamma function and its application," *Electronic Letters*, vol. 2, no. 1, pp. 35–36, Jan. 2006.

[83] L.G. Ordoez, D.P. Palomar, A. Pages-Zamora, and J. Rodriguez Fonollosa, "High-SNR analytical performance of spatial multiplexing MIMO systems with CSI," *IEEE Trans. Sig. Processing*, vol. 55, no. 11, pp. 5447 – 5463, Nov. 2007.

[84] B.M. Hochwald, T.L. Marzetta, and V. Tarokh, "Multiple-antenna channel hardening and its implications for rate feedback and scheduling," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1893–1909, Sept. 2004.

[85] Z. Wang and G.B. Giannakis, "Outage mutual information of space-time MIMO channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 4, pp. 657–662, Apr. 2004.

[86] M. Olfat, *Spatial processing, power control, and channel allocation for OFDM wireless communications*, Ph.D. thesis, University of Maryland, College Park, 2003.

[87] R.S. Blum, "MIMO capacity with interference," *IEEE Jour. Select. Commun.*, vol. 21, no. 5, pp. 793–801, June 2003.

[88] F.R. Farrokhi, G.J. Foschini, A. Lozano, and R.A. Valenzuala, "Link-optimal BLAST processing with multiple-access interference," in *Proc. IEEE Veh. Technol. Conf.*, Sept. 2000.

[89] E.M. Yeh and A. S. Cohen, "Information theory, queuing, and resource allocation in multi-user fading communications," in *Proc. Conf. Information Sciences and Systems*, Princeton, NJ, USA, Mar. 2004.

[90] D.N.C. Tse and S.V. Hanly, "Multi-access fading channels -part i: Polymatroïd structure, optimal resource allocation and throughout capacity," *IEEE Trans. Inform. Theory*, vol. 44, no. 7, pp. 2796–2815, Nov. 1998.

[91] D.N.C. Tse, "Optimal power allocation over parallel Gaussian broadcast channels," in *Proc. Int. Symp. Inform. Theory*, Ulm, Germany, July 1997.

[92] L. Li and A.J. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels -part I: Ergodic capacity," *IEEE Trans. Inform. Theory*, vol. 47, no. 3, pp. 1082–1102, Mar. 2001.

[93] S. Pfletschinger, "Achievable rate regions for OFDMA with link adaptation," in *Proc. of Personal, Indoor and Mobile Commun. Conf.*, Athens, Grece, Sept. 2007.

[94] R. Agrawal, R. Berry, J. Huang, and V. Subramanian, "Downlink scheduling and resource allocation for OFDM systems," in *Proc. Inf. Sciences and Systems Conf.*, Princeton, NJ, USA, Mar. 2006.

[95] R. Agrawal, V. Majjigi, R. Vannithamby, and J.M. Cioffi, "Efficient scheduling for heterogeneous services in OFDMA downlink," in *Proc. IEEE Globecom*, Washington DC, USA, Nov. 2007.

[96] S. Kittipiyajul and T. Javidi, "Resource allocation in OFDMA with time-varying channel and bursty arrivals," *IEEE Commun. Letters*, vol. 11, no. 9, pp. 708–710, Sept. 2007.

[97] S. Kittipiyajul and T. Javidi, "A fresh look at optimal subcarrier allocation in OFDMA systems," in *Proc. IEEE Conf. Decision and Control*, Atlantis, Bahamas, Dec. 2004.

[98] M. Chiang, C.W. Tan, D.P. Palomar, D. O'Neill, and D. Julian, "Power control by geometric programming," *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2640–2651, July 2007.

[99] X. Qiu and K. Chawla, "On the performance of adaptive modulation in cellular systems," *IEEE Trans. Commun.*, vol. 47, no. 6, pp. 884–895, June 1999.

[100] M. Chiang, “Balancing transport and physical layers in wireless multihop networks: jointly optimal congestion control and power control,” *IEEE Jour. Select. Commun.*, vol. 23, no. 1, pp. 104–116, Jan. 2005.

[101] J. Huang, R. Berry, and M.L. Honig, “Distributed interference compensation for wireless networks,” *IEEE Jour. Select. Commun.*, vol. 24, no. 5, pp. 1074–1084, May 2006.

[102] A. Gjendemjso, D. Gesbert, G.E. Oien, and S.G. Kiani, “Binary power control for sum rate maximization over multiple interfering links,” *IEEE Trans. Wireless Commun.*, vol. 7, no. 8, pp. 3164–3163, Aug. 2008.

[103] D. Brélaz, “New methods to color the vertices of a graph,” *Communications of the ACM*, vol. 22, no. 4, pp. 251–256, Apr. 1979.

[104] F. Kuhn and R. Wattenhofer, “On the complexity of distributed graph coloring,” in *Proc. of Symp. Princ. Distributed Computing*, Denver, Colorado, USA, July 2006.

[105] I. Csiszár and G. Tusnàdy, “Information geometry and alternating minimization procedures,” *Statistics and Decisions, Supplement issue*, pp. 205–237, 1984.

[106] U. Niesen, D. Shah, and G. Wornell, “Adaptive Alternating minimization algorithms,” *Submitted, available on arXiv*, Oct. 2008.