



Théorie des jeux et apprentissage pour les réseaux sans fil distribués

François Mériaux

► To cite this version:

François Mériaux. Théorie des jeux et apprentissage pour les réseaux sans fil distribués. Autre [cond-mat.other]. Université Paris Sud - Paris XI, 2013. Français. NNT : 2013PA112296 . tel-00952069

HAL Id: tel-00952069

<https://theses.hal.science/tel-00952069>

Submitted on 26 Feb 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ PARIS-SUD
ÉCOLE DOCTORALE STITS
Laboratoire des Signaux et Systèmes (L2S)

Discipline : PHYSIQUE

THÈSE DE DOCTORAT

soutenue le 26 Novembre 2013

par

François MÉRIAUX

Théorie des jeux et apprentissage pour les réseaux sans fil distribués

Thèse dirigée par Samson LASAULCE
et co-encadrée par Michel KIEFFER

Jury :

<i>Rapporteurs :</i>	Rachid EL-AZOUZI	-	Université d'Avignon
	David GESBERT	-	Eurecom
<i>Directeur :</i>	Samson LASAULCE	-	CNRS
<i>Co-encadrant :</i>	Michel KIEFFER	-	Université Paris-Sud
<i>Examineurs :</i>	Mohamad ASSAAD	-	Supélec
	Jean-Claude BELFIORE	-	Telecom ParisTech
	Romain COUILLET	-	Supélec
	Marceau COUPECHOUX	-	Telecom ParisTech

Remerciements

Mes remerciements vont tout d'abord aux professeurs qui ont eu une influence significative sur mes études jusqu'à cette thèse. Cela commence en classe préparatoire par Thomas Lafforgue et Gilles De Bouard qui m'ont véritablement fait découvrir, non sans douleur, les Mathématiques et de la Physique. Plus tard, en école, ce fut Sylvie Retailleau, Cécile Durieu et Stéphanie Cassan qui les premières, m'initèrent au plaisir d'enseigner. Puis, lorsqu'il fut temps de choisir un domaine de recherche, Michel Kieffer et son cours de *Théorie de l'information et codage de source* furent déterminants.

Au laboratoire des Signaux et Systèmes, j'ai eu la chance de rejoindre une équipe formée de Véronica, Samir et Maël, que je remercie infiniment pour leur accueil. Je ne peux bien sûr pas oublier tous les doctorants qui donnaient vie au laboratoire lors de mon arrivée : Francesca, Lana, Ziad, Nabile, Sofiane, Amadou, José, Thang et tous ceux qui sont arrivés en même temps ou après moi : Elsa, Zeina, Pierre, Vineeth, Benjamin, Anna, Olivier, Achal, Étienne, Diane, Chengfang.

Puisque cette thèse a aussi été une aventure personnelle, je remercie toutes les personnes qui ont eu confiance en moi et m'ont toujours permis d'avancer : Samson Lasaulce, Michel Kieffer, Silviu Niculescu, Gilles Fleury et Mérouane Debbah. Pour ce que nous sommes en train de construire, mes remerciements vont également à Jakob Hoydis, avec qui c'est un bonheur de travailler.

Enfin, je remercie mes parents et ma soeur pour leur amour et leur présence indéfectible.

Table des matières

1	Introduction	5
1.1	Contexte et motivation	5
1.2	Structure et contributions	7
1.3	Publications	8
2	Du jeu statique au jeu stochastique	11
2.1	Jeu statique de contrôle de puissance	14
2.1.1	Modèle du réseau sans fil	14
2.1.2	Définition du jeu de contrôle de puissance	15
2.1.3	Équilibre de Nash	17
2.2	Report stratégique du gain du canal	17
2.2.1	Modélisation du système	18
2.2.2	Jeu statique de report des gains de canaux	20
2.2.3	Résultats numériques	23
2.3	Jeu répété de contrôle de puissance	26
2.3.1	Définition du jeu répété	27
2.3.2	Stratégie d'équilibre du jeu répété	29
2.4	Jeu stochastique de contrôle de puissance	30
2.4.1	Définition du jeu stochastique	31
2.4.2	Résultats analytiques	32
2.4.3	Résultats numériques	35
2.5	Conclusion	39
3	Jeux à champ moyen	41
3.1	Formulation du jeu différentiel stochastique	43
3.1.1	Définition du jeu différentiel stochastique	43
3.1.2	Équilibre de Nash du jeu différentiel stochastique	45
3.1.3	Cas du jeu à un joueur	47
3.2	Analyse du jeu à champ moyen	47
3.2.1	Convergence vers le jeu à champ moyen	47
3.2.2	Solution du problème de réponse au champ moyen	49
3.2.3	Unicité de la solution	51
3.2.4	Résultats numériques pour le jeu à champ moyen	51
3.3	Conclusion	54
4	Apprentissage et dynamiques de meilleures réponses	59
4.1	Jeu de positionnement de points d'accès	61
4.1.1	Modélisation du système	62
4.1.2	Étude de l'équilibre de Nash	66
4.1.3	Convergence vers l'équilibre de Nash	67

4.2	Jeu de satisfaction de contraintes de qualité de service	71
4.2.1	Formulation du jeu de satisfaction	73
4.2.2	Convergence de la DMR vers un ESE	74
4.2.3	Convergence de la dynamique de réponse robuste aveugle (DRRA) vers un ESE	78
4.2.4	Résultats numériques	79
4.3	Conclusion	80
5	Conclusion et perspectives	83
A	Publications relatives aux jeux statiques et jeux stochastiques	85
A.1	IEEE-VALUETOOLS-2012	85
A.2	IEEE-DSP-2011	94
B	Publications relatives aux jeux à champ moyen	101
B.1	SPRINGER-DGAA-2013	101
B.2	IEEE-ASILOMAR-2012	124
C	Publications relatives à l'apprentissage et aux dynamiques de meilleures réponses	131
C.1	ICST-VALUETOOLS-2011	131
C.2	SPRINGER-GAMENETS-2012	141
	Bibliographie	157

Résumé :

Dans cette thèse, nous étudions des réseaux sans fil dans lesquels les terminaux mobiles sont autonomes dans le choix de leurs configurations de communication. Cette autonomie de décision peut notamment concerner le choix de la technologie d'accès au réseau, le choix du point d'accès, la modulation du signal, les bandes de fréquences occupées, la puissance du signal émis, etc. Typiquement, ces choix de configuration sont réalisés dans le but de maximiser des métriques de performances propres à chaque terminal. Sous l'hypothèse que les terminaux prennent leurs décisions de manière rationnelle afin de maximiser leurs performances, la théorie des jeux s'applique naturellement pour modéliser les interactions entre les décisions des différents terminaux.

Plus précisément, l'objectif principal de cette thèse est d'étudier des stratégies d'équilibre de contrôle de puissance d'émission afin de satisfaire des considérations d'efficacité énergétique. Le cadre des jeux stochastiques est particulièrement adapté à ce problème et nous permet notamment de caractériser la région de performance atteignable pour toutes les stratégies de contrôle de puissance qui mènent à un état d'équilibre. Lorsque le nombre de terminaux en jeu est grand, nous faisons appel à la théorie des jeux à champ moyen pour simplifier l'étude du système. Cette théorie nous permet d'étudier non pas les interactions individuelles entre les terminaux, mais l'interaction de chaque terminal avec un champ moyen qui représente l'état global des autres terminaux. Des stratégies de contrôle de puissance optimales du jeu à champ moyen sont étudiées.

Une autre partie de la thèse a été consacrée à des problématiques d'apprentissage de points d'équilibre dans les réseaux distribués. En particulier, après avoir caractérisé les positions d'équilibre d'un jeu de positionnement de points d'accès, nous montrons comment des dynamiques de meilleures réponses et d'apprentissage permettent de converger vers un équilibre. Enfin, pour un jeu de contrôle de puissance, la convergence des dynamiques de meilleures réponses vers des points d'équilibre a été étudiée. Il est notamment proposé un algorithme d'adaptation de puissance convergeant vers un équilibre avec une faible connaissance du réseau.

Game Theory and Learning for Wireless Distributed Networks

Abstract : In this thesis, we study wireless networks in which mobile terminals are free to choose their communication configuration. These configuration choices include access wireless technology, access point association, coding-modulation scheme, occupied bandwidth, power allocation, etc. Typically, these configuration choices are made to maximize some performance metrics associated to every terminals. Under the assumption that mobile terminals take their decisions in a rational manner, game theory can be applied to model the interactions between the terminals.

Precisely, the main objective of this thesis is to study energy-efficient power control policies from which no terminal has an interest to deviate. The framework of stochastic games is particularly suited to this problem and allows to characterize the achievable utility region for equilibrium power control strategies. When the number of terminals in the network is large, we invoke mean field game theory to simplify the study of the system. Indeed, in a mean field game, the interactions between a player and all the other players are not considered individually. Instead, one only studies the interactions between each player and a mean field, which is the distribution of the states of all the other players. Optimal power control strategies from the mean field formulation are studied.

Another part of this thesis has been focused on learning equilibria in distributed games. In particular, we show how best response dynamics and learning algorithms can converge to an equilibrium in a base station location game. For another scenario, namely a power control problem, we study the convergence of the best response dynamics. In this case, we propose a power control behavioral rule that converges to an equilibrium with very little information about the network.

Notations : Dans l'ensemble de ce manuscrit, les scalaires et les vecteurs sont respectivement représentés par des symboles en minuscule et des symboles en minuscule soulignés. La notation $\underline{a}_{-k} = (a_1, \dots, a_{k-1}, a_{k+1}, \dots, a_K)$ représente le vecteur obtenu en retirant la $k^{\text{ième}}$ composante du vecteur $\underline{a}$. Avec un léger abus de notation, nous pourrions écrire $\underline{a} = (a_k, \underline{a}_{-k})$ pour mettre en valeur l'influence de la $k^{\text{ième}}$ composante du vecteur $\underline{a}$. Les notations $\nabla_{\underline{x}} f$ et $\Delta_{\underline{x}} f$ représentent respectivement le gradient et le Laplacien de la fonction f par rapport au vecteur $\underline{x}$. L'opérateur de divergence par rapport à $\underline{x}$ est noté $\text{div}_{\underline{x}}$. Enfin, $\langle \cdot, \cdot \rangle_{\mathbb{E}}$ représente le produit scalaire dans l'espace $\mathbb{E}$.

Introduction

1.1 Contexte et motivation

La théorie des jeux est une branche des mathématiques consacrée à l'étude de systèmes composés d'un ensemble d'agents, nommés joueurs, choisissant leurs actions en fonction de leur anticipation des actions des autres joueurs. Chaque joueur prend ses décisions afin de maximiser une fonction d'utilité propre qui dépend généralement des actions de tous les joueurs. Dans l'étude d'un jeu, une question cruciale est l'existence d'un équilibre de Nash [Nash 1950]. Un tel équilibre est un profil d'actions dont aucun joueur n'a intérêt à dévier unilatéralement, c'est donc un profil dont la stabilité est assurée sous l'hypothèse que les joueurs sont rationnels. Particulièrement adaptée à des problèmes d'ordre économique, la théorie des jeux a été récompensée de plusieurs prix Nobel, le dernier en date ayant été attribué à L. Shapley et A. Roth en 2012. Avec l'essor des réseaux de communication sans fil, sont apparus des problèmes de partage de ressources (bandes de fréquences, puissance d'émission, ordonnancement, etc) ainsi que des configuration de réseaux donnant une autonomie de décision aux terminaux mobiles du système. C'est notamment le cas pour les réseaux de capteurs sans fil, les réseaux dans les bandes de fréquence ISM (Wifi, Bluetooth, ZigBee), les réseaux femto-cell et les réseaux ad hoc en général. Dans ce contexte, la théorie des jeux est donc un outil pleinement adapté à l'étude des états stables de ces réseaux. Un autre argument en faveur de la théorie des jeux qui est exploité dans cette thèse est que dans de tels réseaux, la configuration des communication n'est pas figée, elle évolue au cours du temps en fonction de paramètres tels que le nombre de terminaux dans le système et la qualité des canaux de transmission des terminaux. Au delà du jeu statique, la théorie des jeux propose justement des modèles de jeux répétés et jeux stochastiques pour tenir respectivement compte des implications stratégiques de la répétition d'un jeu et des variations des conditions de jeu au cours du temps.

En parallèle de l'autonomie de décision croissante des terminaux mobiles, les usages de ces terminaux sont aussi en train d'évoluer. En particulier, la consommation de données, qui était jusqu'à peu cloisonnée aux réseaux filaires, est en train d'exploser dans les réseaux mobiles. Ce constat pose des problèmes d'optimisation d'allocation des ressources, aussi bien au niveau du point d'accès sans fil que des terminaux mobiles. S'il est bien sûr important de travailler à augmenter le débit disponible pour chaque terminal, il est aussi nécessaire de veiller à ce que cela ne se fasse pas au détriment de la consommation énergétique. En effet, entre deux recharges de batteries, les terminaux mobiles disposent d'une quantité d'énergie

finie qui doit donc être employée de manière efficace. En d'autres termes, en raison de la quantité finie d'énergie contenue dans les terminaux mobiles, il n'est donc pas seulement important d'optimiser les débits des communications mais il est aussi fondamental d'optimiser la quantité d'information échangée relativement à la quantité d'énergie dépensée par le terminal mobile. Une telle fonction d'utilité a justement été introduite dans [Goodman 2000] pour un problème de contrôle de puissance sur un canal à accès multiple et une large partie de ce travail de thèse est basée sur ces travaux.

Classiquement dans un jeu, les actions d'un joueur ont une influence directe sur l'utilité des autres joueurs et réciproquement. Les décisions d'actions de chaque joueur dépendent donc de l'anticipation du comportement de chacun des autres joueurs. Ce phénomène peut rendre compliqué, voir impossible, la résolution d'équations de points fixes caractérisant les équilibres lorsque le nombre de joueurs est trop important. Dans cette thèse, c'est notamment le cas lorsque nous modélisons le problème de contrôle de puissance sur un canal à accès multiple par un jeu différentiel stochastique. Ceci dit, bien que dans un jeu avec un grand nombre de joueurs, les interactions sont plus complexes, l'importance relative d'un joueur est aussi moindre à l'échelle du système. Il se peut donc que les actions d'un seul joueur ne changent pas véritablement l'état global du système. C'est sur la base de ces idées que s'est développée la théorie des jeux à champ moyen [Lasry 2007]. Dans un jeu à champ moyen, l'ensemble des joueurs devient continu et chaque joueur n'est plus directement dépendant de chacun des autres joueurs, mais dépend d'un champ moyen qui est défini comme la distribution des états de l'ensemble des joueurs. Pour un joueur, la résolution d'un équilibre de ce jeu implique donc seulement son propre état ainsi que le champ moyen qui est un paramètre global du jeu. Dans le cas du contrôle de puissance efficace énergétiquement sur des canaux à accès multiples, l'état de chaque joueur est l'état de son canal ainsi que le niveau d'énergie dont il dispose et le champ moyen s'exprime à travers l'interférence perçue par le point d'accès. La modélisation de ce problème par un jeu à champ moyen permet d'en déduire des stratégies de contrôle de puissance innovantes.

Enfin le dernier problème général abordé dans cette thèse est celui de l'apprentissage des équilibre par des méthodes dynamiques. Dans un jeu, l'existence d'un équilibre dépend du niveau d'information dont disposent les joueurs. En d'autres termes, en fonction du niveau d'information disponible, différents équilibres peuvent être solutions du jeu. Mais dans un scénario pratique, le niveau d'information requis pour qu'un équilibre d'intérêt soit connu des joueurs n'est souvent pas réaliste. Pour pallier à cela, des méthodes d'apprentissage telles que les dynamiques de meilleures réponses [Fudenberg 1991], l'apprentissage par renforcement [Sutton 1998], *fictitious play* [Brown 1951] existent. Elles permettent aux agents du système de converger vers des états d'équilibres souhaités, sans pour autant disposer du niveau d'information nécessaire à la résolution formelle de ces équilibres. Au sein même de ces méthodes d'apprentissage, nous verrons que le niveau d'information sur le système influe sur les propriétés de convergence et sur les équilibres atteignables. De manière surprenante, nous verrons que dans certains scénarios, il est préférable d'avoir moins

d'information sur le système pour garantir la convergence vers un équilibre.

1.2 Structure et contributions

Ce manuscrit est organisé en trois chapitres qui reprennent les contributions principales de cette thèse.

Dans le chapitre 2, nous commençons par présenter le jeu de contrôle de puissance efficace énergétiquement tel qu'il a été introduit dans [Goodman 2000]. À partir de cette base, nous étudions un problème de report de gains de canaux dans une cellule où l'allocation de puissance est efficace énergétiquement. La caractérisation d'un équilibre dans le report stratégique des gains de canaux permet de comparer les performances d'efficacité énergétique de la cellule avec les reports stratégiques et les reports des gains de canaux réels. Puis la version répétée du jeu de contrôle de puissance efficace énergétiquement [Le Treust 2010] est présentée. À partir de ce modèle, nous proposons une extension stochastique du jeu répété. Dans ce nouveau jeu, nous caractérisons la région des efficacités énergétiques des stratégies d'équilibre et nous proposons une stratégie d'équilibre concrète qui offre de meilleures performances en termes d'efficacité énergétique que les stratégies de contrôle de puissance issues du jeu statique et du jeu répété.

Dans le chapitre 3, nous proposons un modèle de jeu différentiel stochastique pour le problème de contrôle de puissance efficace énergétiquement. Ce jeu permet de tenir compte à la fois de l'évolution des canaux dans le temps et de l'évolution de l'énergie disponible dans chaque terminal mobile. Bien que des conditions suffisantes d'existence d'un équilibre de Nash soient données, la caractérisation d'un équilibre de ce jeu est une tâche complexe, en particulier lorsque le nombre de joueurs est élevé. Pour contourner cette complexité, nous donnons des conditions de convergence du jeu différentiel stochastique vers un jeu à champ moyen [Lasry 2007]. Dans ce nouveau jeu, nous donnons des conditions suffisantes d'unicité d'un équilibre et nous illustrons les performances des stratégies de contrôle de puissance issues de ce modèle par des résultats numériques.

Dans le chapitre 4, nous présentons un jeu de positionnement de points d'accès mobiles dans lequel chaque point d'accès se positionne pour maximiser le trafic généré par les terminaux mobiles qui s'y connectent. L'étude de l'équilibre de Nash de ce jeu révélant que les positions d'équilibre ne peuvent pas être connues des joueurs avec des hypothèses raisonnables de connaissance du système, nous présentons deux méthodes permettant de converger vers l'équilibre. La première est la dynamique des meilleures réponses, dont la convergence est prouvée mais qui nécessite que chaque point d'accès connaisse son positionnement optimal étant donné les positions des autres points d'accès. La seconde est l'apprentissage par renforcement qui nécessite seulement que chaque point d'accès soit capable de connaître la valeur de son utilité pour un profil de positions données mais dont la convergence n'est pas assurée. Nous présentons également un problème de satisfaction de contraintes de qualité de service dans les réseaux distribués. Précisément, dans le cadre des jeux

de satisfaction [Perlaza 2012], nous présentons une règle de fonctionnement dont des conditions suffisantes de convergence vers un équilibre de satisfaction efficace sont données dans le cas général. Plus particulièrement, dans un jeu de contrôle de puissance dont les ensembles d'actions sont discrets, cette règle de fonctionnement présente l'avantage de converger vers un équilibre à partir de n'importe quel profil de puissances d'émission.

Le chapitre 5 conclue ce manuscrit et une sélection d'articles publiés pendant cette thèse est disponible en annexe, notamment pour consulter les preuves qui ne seraient pas données dans les chapitres 2, 3 et 4.

1.3 Publications

Cette thèse a donné lieu à la publication de plusieurs articles dans des communications avec actes et des revues avec comité de lecture. Nous les présentons ici selon leurs thèmes directeurs.

Contrôle de puissance efficace énergétiquement

La modélisation du problème de contrôle de puissance efficace énergétiquement du **jeu statique au jeu stochastique** a entraîné la publications de 3 articles de conférences internationales et un article de conférence nationale :

- F. Mériaux, Y. Hayel, S. Lasaulce et A. Garnaev. *Long-Term Energy Constraints on Power Control in Cognitive Radio Networks*. In IEEE International Conference on Digital Signal Processing (DSP), pages 842-848, 2011.
- F. Mériaux, M. Le Treust, S. Lasaulce et M. Kieffer. *A Stochastic Game Formulation of Energy-Efficient Power Control : Equilibrium Utilities and Practical Strategies*. In IEEE International Conference on Digital Signal Processing (DSP), pages 665-671, 2011.
- F. Mériaux, M. Le Treust, S. Lasaulce, and M. Kieffer. *Jeux Stochastiques et Contrôle de Puissance Distribuée*. In Colloque du Groupement de Recherche en Traitement du Signal et des Images (GRETSI), 2011.
- F. Mériaux, S. Valentin et S. Lasaulce. *An Energy-Efficient Power Allocation Game with Selfish Channel State Feedback in Cellular Networks*. In IEEE Conference on Performance Evaluation Methodologies and Tools (VALUE-TOOLS), pages 290-297, 2012.

L'extension du modèle vers les **jeux à champ moyen** a mené à la publication de 2 articles de conférences internationales et un article de revue :

- F. Mériaux et S. Lasaulce. *Mean-Field Games and Green Power Control*. In IEEE International Conference on Network Games, Control and Optimization (NetGCooP), pages 1-5, 2011.
- F. Mériaux, S. Lasaulce et H. Tembine. *Stochastic Differential Games and Energy-Efficient Power Control*. Dynamic Games and Applications, vol. 3, no. 1, pages 3-23, 2013.

- F. Mériaux, V. Varma et S. Lasaulce. *Mean Field Energy Games in Wireless Networks*. In IEEE Conference on Signals, Systems and Computers (ASILOMAR), pages 671-675, 2012.

Apprentissage et dynamiques de meilleures réponses

L'étude d'algorithmes convergeant vers des solutions d'équilibre de jeux a donné lieu à la publication de 2 articles de conférences internationales :

- F. Mériaux, S. Lasaulce et M. Kieffer. *More about Base Station Location Games*. In ICST Conference on Performance Evaluation Methodologies and Tools (VALUETOOLS), pages 372-380, 2011.
- F. Mériaux, S. M. Perlaza, S. Lasaulce, Z. Han et V. Poor. *Achievability of Efficient Satisfaction Equilibria in Self-Configuring Networks*. In Game Theory for Networks, vol. 105, pages 1-15. Springer Berlin Heidelberg, 2012.

Enfin, 2 articles de conférences internationales sortent du cadre abordé dans ce manuscrit :

- M. Mhiri, K. Cheikhrouhou, A. Samet, F. Mériaux et S. Lasaulce. *Energy-Efficient Spectrum Sharing in Relay-Assisted Cognitive Radio Systems*. In IEEE International Conference on Network Games, Control and Optimization (NetGCooP), pages 86-91, 2012.
- X. Chen, F. Mériaux et S. Valentin. *Predicting a User's Next Cell With Supervised Learning Based on Channel States*. In IEEE International Workshop on Signal Processing Advances for Wireless Communications (SPAWC), pages 1-5, 2013.

Contrôle de puissance efficace énergétiquement : du jeu statique au jeu stochastique

Sommaire

2.1	Jeu statique de contrôle de puissance	14
2.1.1	Modèle du réseau sans fil	14
2.1.2	Définition du jeu de contrôle de puissance	15
2.1.3	Équilibre de Nash	17
2.2	Report stratégique du gain du canal	17
2.2.1	Modélisation du système	18
2.2.2	Jeu statique de report des gains de canaux	20
2.2.3	Résultats numériques	23
2.3	Jeu répété de contrôle de puissance	26
2.3.1	Définition du jeu répété	27
2.3.2	Stratégie d'équilibre du jeu répété	29
2.4	Jeu stochastique de contrôle de puissance	30
2.4.1	Définition du jeu stochastique	31
2.4.2	Résultats analytiques	32
2.4.3	Résultats numériques	35
2.5	Conclusion	39

Le contrôle de puissance est depuis longtemps considéré comme un problème important pour les communications entre plusieurs utilisateurs qui partagent des ressources communes [Foschini 1993, Yates 1995]. Avec l'apparition de paradigmes nouveaux tels que les réseaux ad hoc [Gupta 1997], les bandes de fréquence non réglementées et la radio cognitive [Fette 2009, Mitola 1999], l'étude de protocoles de contrôle de puissance distribués devient particulièrement pertinente. En effet, dans de tels réseaux, les terminaux peuvent choisir librement leur stratégie de contrôle de puissance sans suivre des règles imposées par un nœud central. Les terminaux sont alors considérés comme des acteurs indépendants et il est raisonnable de considérer qu'ils sont également rationnels, c'est-à-dire que par le contrôle de leurs niveaux de puissance d'émission, les terminaux souhaitent maximiser leur qualité de communication. Dans ce contexte, il est naturel d'étudier le problème de contrôle de puissance de chaque terminal dans le cadre de la théorie des jeux, en considérant les

terminaux comme des joueurs cherchant chacun à maximiser leur propre fonction d'utilité en contrôlant leur puissance d'émission. La théorie des jeux permet notamment d'étudier l'existence et la multiplicité de stratégies de contrôle de puissance d'équilibre dont aucun terminal n'a intérêt à dévier unilatéralement (équilibre de Nash).

Dans un canal à accès multiple, le signal émis par un terminal est perçu par les autres terminaux comme une interférence pour leurs propres signaux. La qualité de transmission de chaque terminal dépend donc directement du niveau de puissance d'émission des autres terminaux. Un panorama des travaux précédents sur l'allocation de puissance dans les canaux à accès multiples est présenté dans [Belmega 2009]. Une caractéristique commune de ces travaux est que l'amélioration de la qualité de transmission des terminaux se partageant le canal est généralement obtenue en réduisant l'interférence entre ces terminaux. Cela signifie que, paradoxalement, pour améliorer la qualité globale de transmission des terminaux qui se partagent le canal, ces derniers ont intérêt à diminuer leur puissance d'émission. Deux approches visent à mettre en place des protocoles de contrôle de puissance qui remplissent cet objectif. La première consiste à minimiser la puissance d'émission sous certaines contraintes (typiquement des contraintes de débit minimal). La seconde consiste à maximiser une mesure d'efficacité énergétique qui correspond au rapport entre le nombre de bits d'information correctement transmis et la quantité d'énergie consommée par l'émission. Alors que la première approche vise à utiliser le minimum d'énergie pour remplir des objectifs donnés, la seconde se concentre sur l'efficacité énergétique de la transmission. En d'autres termes, pour un joule dépensé, quel est le maximum d'information qu'il est possible de transmettre ? Clairement, la minimisation d'énergie sous contrainte et la maximisation de l'efficacité énergétique sont deux approches différentes dont la pertinence dépend du contexte considéré (voir [Goodman 2000] et les travaux associés pour plus de justifications) et ne peuvent pas être comparées dans le cas général. Les résultats présentés dans ce chapitre concernent le second ensemble de travaux, dont l'objectif est de minimiser l'efficacité énergétique qui est mesurée comme le nombre moyen de bits d'information transmis avec succès par joule consommé par l'émetteur.

Dans la formulation d'origine proposée par [Goodman 2000] et reprise dans la plupart des travaux connexes [Meshkati 2006, Bonneau 2008, Lasaulce 2009, Buzzi 2011], le problème du contrôle de puissance efficace énergétiquement est modélisé par un jeu statique, dans lequel les gains de canaux des terminaux sont fixes. Bien que ce jeu permette d'obtenir simplement l'expression d'un équilibre unique, deux inconvénients apparaissent dans cette formulation. Premièrement, l'équilibre de ce jeu n'est pas optimal au sens de Pareto, c'est-à-dire que les puissances d'émission à l'équilibre de ce jeu ne permettent pas d'atteindre des valeurs d'efficacité énergétique qui dominent la région atteignable. Deuxièmement, cette formulation néglige l'évolution des canaux de transmission au cours du temps et ne permet pas la mise à jour des puissances d'émission sur chaque intervalle de temps.

Le premier de ces deux inconvénients peut, par exemple, être traité par des méthodes établissant un coût d'émission telles que [Saraydar 2002]. Cependant ces

méthodes imposent l'ajout d'un terme de coût à la fonction d'utilité du jeu, ce qui change en profondeur la structure et les propriétés du jeu. Concernant le second inconvénient, nous pourrions penser que pour les scénarios dans lesquels l'état des canaux provient de la réalisation iid d'une variable aléatoire donnée pour chaque intervalle de temps, les puissances d'émission ne sont pas corrélées d'un état au suivant. De ce fait, la modélisation du problème de contrôle de puissance au cours du temps reviendrait à considérer une séquence de jeux statiques joués de manière indépendante d'un intervalle au suivant. Mais même dans le cas où les états des canaux sont indépendants entre deux intervalles de temps, il existe des interactions stratégiques entre les terminaux qui apparaissent en raison de la répétition du jeu dans le temps et ces interactions ne sont pas présentes dans une séquence de jeux statiques. Pour tenir compte de ces interactions stratégiques, Le Treust et Lasaulce ont proposé une formulation du problème basée sur un jeu répété [Le Treust 2010]. La différence fondamentale entre leur formulation et une séquence de jeux statiques est que dans ce nouveau modèle, les terminaux gardent en mémoire les étapes passées du jeu. En conséquence, chaque terminal peut adapter son contrôle de puissance en fonction du comportement passé des autres terminaux. En particulier, il est possible d'augmenter le niveau de puissance pour punir des comportements non respectueux de l'intérêt général. En comparaison de l'approche classique d'établissement d'un coût d'émission [Saraydar 2002], une des caractéristiques fortes de leur formulation est que chaque terminal a uniquement besoin de la connaissance de son propre canal ainsi que d'un signal public pour atteindre un état d'équilibre Pareto-optimal. Cependant, ce jeu répété ne permet pas de modéliser correctement la variation des canaux au cours du temps puisque dans le modèle de jeu répété utilisé, les canaux sont considérés constants d'une étape à l'autre du jeu. De fait, la stratégie d'équilibre proposée par Le Treust et Lasaulce n'est pas optimale au sens de Pareto si les canaux varient d'une étape à la suivante.

Afin de tenir compte de la variation possible des canaux au cours du temps, nous avons alors fait évoluer le modèle du jeu pour passer d'un jeu répété à un jeu stochastique dans [Mériaux 2011c]. La formulation du jeu stochastique inclut notamment un état du jeu qui évolue de manière dynamique à chaque étape du jeu. Dans notre problème de contrôle de puissance, cet état peut être défini par l'ensemble des gains des canaux des terminaux mobiles, sur chaque intervalle de temps. Avec cette formulation stochastique, les travaux de [Fudenberg 2011, Hörner 2011] peuvent être utilisés pour obtenir un *Folk* théorème qui caractérise la région des utilités atteignables de ce jeu stochastique. Une stratégie de contrôle de puissance d'équilibre est également proposée pour obtenir de meilleures performances que la stratégie issue du jeu répété proposée dans [Le Treust 2010].

Ce premier chapitre est organisé de la manière suivante. Dans la partie 2.1, nous présentons la formulation de jeu statique initialement introduit dans [Goodman 2000] pour modéliser le problème de contrôle de puissance efficace énergétiquement. Dans la partie 2.2, nous proposons une variante de ce modèle lorsque l'allocation de puissance est réalisée par la station de base en fonction des gains de canaux reportés par les terminaux mobiles. Dans la partie 2.3, nous

repreons la formulation du jeu répété introduite par [Le Treust 2010], avant de proposer un modèle de jeu stochastique dans la partie 2.4.

2.1 Jeu statique de contrôle de puissance

Le but de cette partie est de passer en revue la formulation statique du problème de contrôle de puissance efficace énergétiquement proposée par Goodman et Mandayam [Goodman 2000]. Ce premier modèle occupe une place primordiale dans l'organisation de ce travail de thèse puisque l'ensemble des travaux présentés dans ce chapitre et au chapitre 3 en découlent.

2.1.1 Modèle du réseau sans fil

Dans un canal à accès multiple [Wyner 1974, Cover 2006], nous considérons $K \geq 1$ émetteurs mobiles et un récepteur. Ce réseau est décentralisé puisque le récepteur commun ne donne pas de consignes aux émetteurs concernant leur puissance d'émission. Les émetteurs sont donc libres de choisir leur propre puissance d'émission. Nous supposons que les émetteurs envoient leurs signaux par paquets au récepteur sur des canaux quasi-statiques, en même temps, sur la même bande de fréquence et sans mécanisme pour atténuer les interférences tels que la formation de faisceaux [Van Veen 1988]. Un paquet est défini comme une séquence de symboles de longueur finie, précédée d'une séquence d'apprentissage qui permet au récepteur d'estimer le canal de communication pour chaque paquet. Nous supposons donc que la transmission d'un paquet se fait sur une durée inférieure au temps de cohérence du canal. Le signal en bande de base reçu par la station de base s'écrit :

$$y(i) = \sum_{k=1}^K h_k(i)x_k(i) + z(i) \quad (2.1)$$

où $k \in \mathcal{K}$, $\mathcal{K} = \{1, \dots, K\}$, $x_k(i)$ représente le symbole transmis par l'émetteur k à l'instant i . La puissance du signal émis par chaque émetteur est notée $\mathbb{E}[|x_k|^2] = p_k \in \mathcal{A}_k$. Le bruit z est supposé suivre une distribution gaussienne de moyenne nulle et de variance σ^2 . Chaque canal h_k varie dans le temps mais est supposé constant pour la durée d'un paquet. Pour chaque émetteur $k \in \mathcal{K}$, le gain de canal est contenu dans un ensemble compact : $|h_k|^2 = \eta_k \in [\eta_k^{\min}, \eta_k^{\max}]$. Cette hypothèse traduit la sensibilité finie du récepteur et la contrainte d'une distance minimale entre émetteur et récepteur. Enfin, nous supposons que le récepteur décode le message de chaque émetteur séparément. La formulation du signal transmis correspond au modèle de canal à accès multiple issu de la théorie de l'information (voir [Belmega 2009] pour plus de précision au sujet du canal à accès multiple). Ce modèle peut facilement s'appliquer à des protocoles de communication spécifiques tels que le CDMA [Goodman 2000, Lasaulce 2009] ou le CDMA avec plusieurs porteurs [Meshkati 2006].

Pour chaque paquet de l'émetteur $k \in \mathcal{K}$, le rapport signal sur interférence plus bruit (RSIB) au récepteur est noté γ_k . Il s'écrit :

$$\gamma_k = \frac{p_k \eta_k}{\sum_{j \neq k} p_j \eta_j + \sigma^2}. \quad (2.2)$$

Comme exprimé dans (2.2), les signaux des autres émetteurs sont perçus comme de l'interférence pour le signal de l'émetteur k .

2.1.2 Définition du jeu de contrôle de puissance

Dans le réseau sans fil défini dans la partie précédente, nous considérons que les émetteurs cherchent à maximiser leur efficacité énergétique. Cette notion, introduite dans [Goodman 2000] est le rapport entre la quantité d'information correctement transmise au récepteur et l'énergie consommée pour l'émission. Cette métrique de performance peut être perçue comme un compromis entre le débit de la communication et la pollution électromagnétique causée par la transmission dans la région environnante. Pour chaque émetteur $k \in \mathcal{K}$, la fonction d'utilité efficace énergétiquement s'écrit :

$$u_k(p_1, \dots, p_K) = \frac{Rf(\gamma_k(p_1, \dots, p_K))}{p_k} \text{ bit/J}, \quad (2.3)$$

où R est une constante en bit/s qui représente le débit de communication en sortie de l'émetteur [Goodman 2000]. La fonction $f : \mathbb{R}^+ \rightarrow [0, 1]$ est une fonction sigmoïde qui représente le taux de succès de la transmission des paquets et qui dépend du RSIB γ_k . Nous rappelons qu'une fonction sigmoïde est convexe jusqu'à un certain point, puis concave sur le reste de son ensemble de définition. De plus, f est suffisamment régulière pour que u_k soit différentiable sur l'ensemble de définition $\mathcal{A}_k$. Pour plus de détails, voir [Rodriguez 2003, Meshkati 2005, Belmega 2011]. La puissance p_k étant exprimée en watt, l'unité de la fonction d'utilité est donc bit/joule.

Il existe plusieurs définitions de la fonction f qui assurent que ce soit bien une fonction sigmoïde dans l'intervalle $[0, 1]$. Par exemple, f peut être considérée comme le taux d'erreur de démodulation dépendant de la modulation utilisée sur le canal [Goodman 2000]. Dans les travaux de Belmega et Lasaulce [Belmega 2011] issus de la théorie de l'information, dans le cas d'un canal SISO, la fonction d'efficacité f associée au RSIB γ est définie comme :

$$\begin{aligned} f(\gamma) &= 1 - P_{\text{out}}(\gamma, a), \\ &= \exp\left(-\frac{a}{\gamma}\right), \end{aligned} \quad (2.4)$$

où P_{out} est la probabilité de coupure et $a = 2^{R/W} - 1$ est le seuil sous lequel le RSIB ne permet plus de transmettre avec un débit R . Cette expression est issue d'une approximation puisqu'en effet, P_{out} devrait dépendre de l'espérance du gain du canal, et non de la valeur instantanée du gain du canal. Nous justifions cette approximation par le fait que lorsqu'ils adaptent leurs puissances d'émission, les

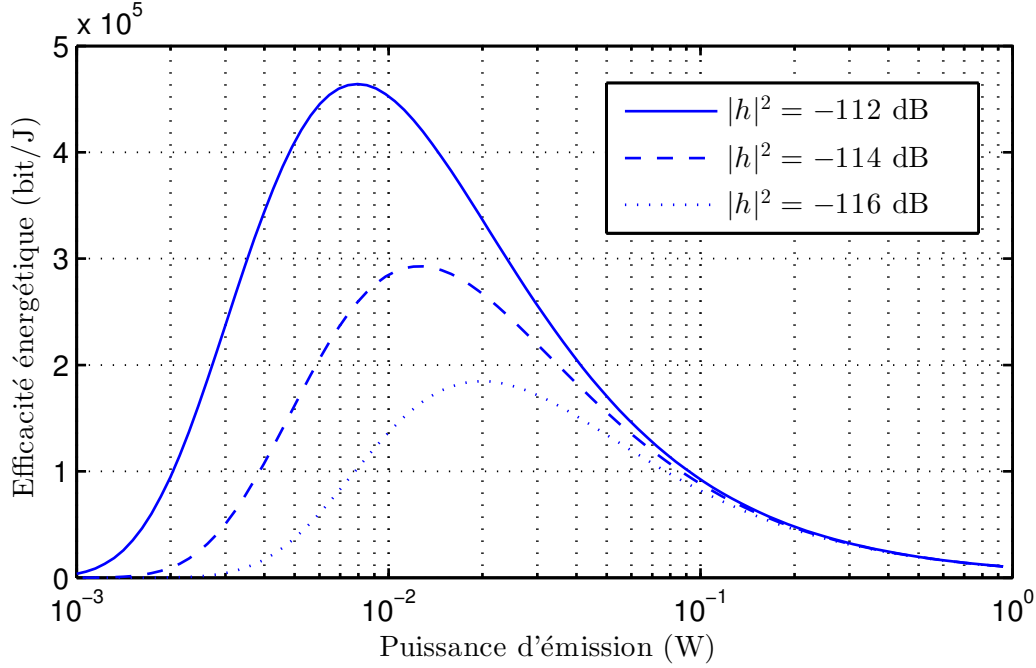


FIGURE 2.1 – Efficacité énergétique d’un émetteur en fonction de la puissance d’émission pour trois valeurs de gains de canaux. Les valeurs de paramètres sont fixées à $a = 1$, et $\sigma^2 = 5 \times 10^{-14}$ W.

émetteurs ne connaissent que leurs gains de canaux instantanés. Dans ce chapitre ainsi que le suivant, l’utilisation de cette fonction sera privilégiée.

Pour optimiser sa propre efficacité énergétique, chaque émetteur doit non seulement adapter sa puissance à la qualité du canal vers le récepteur, mais aussi aux puissances d’émissions des autres émetteurs. Ce problème peut s’exprimer sous forme du jeu présenté ci-dessous.

Définition 1 *Le jeu de contrôle de puissance efficace énergétiquement sous forme stratégique est défini par le triplet $\mathcal{G} = (\mathcal{K}, \{\mathcal{A}_k\}_{k \in \mathcal{K}}, \{u_k\}_{k \in \mathcal{K}})$ dans lequel :*

- $\mathcal{K} = \{1, \dots, K\}$ est l’ensemble des émetteurs ;
- $\mathcal{A}_k = [0, P_k^{\max}]$ est l’espace des actions de l’émetteur $k \in \mathcal{K}$;
- la fonction d’utilité de l’émetteur $k \in \mathcal{K}$ est donnée par

$$u_k(p_1, \dots, p_K) = \frac{Rf(\gamma_k(p_1, \dots, p_K))}{p_k}. \quad (2.5)$$

La figure. 2.1 montre l’allure typique de cette efficacité énergétique en fonction de la puissance d’émission, pour trois valeurs différentes de gain de canal. Il apparaît clairement sur cette figure que l’efficacité énergétique est quasiconcave vis-à-vis de la puissance d’émission et que maximiser l’efficacité énergétique n’est pas équivalent à maximiser le débit. En effet, contrairement à la maximisation de débit, il n’est pas optimal d’utiliser toute la puissance disponible pour maximiser l’efficacité én-

ergétique. Il est aussi intéressant de noter qu'à puissance d'émission fixée, l'efficacité énergétique par récepteur augmente avec la valeur du gain de canal.

2.1.3 Équilibre de Nash

Un concept important pour étudier le résultat d'un jeu est l'équilibre de Nash.

Définition 2 Dans le jeu $\mathcal{G}$, un équilibre de Nash est un profil d'actions $(p_1^*, \dots, p_K^*)$ tel que $\forall k \in \mathcal{K}$

$$\forall p_k \neq p_k^*, u_k(p_k, \underline{p}_{-k}^*) \leq u_k(p_k^*, \underline{p}_{-k}^*). \quad (2.6)$$

En d'autres termes, c'est un profil d'actions dont aucun joueur n'a intérêt à dévier unilatéralement. Précisément, le jeu statique de contrôle de puissance efficace énergétiquement est quasi-concave et possède donc un équilibre de Nash pur [Lasaulce 2011]. De plus, cet équilibre de Nash est unique et donné par l'expression suivante :

$$\forall i \in \{1, \dots, K\}, p_k^* = \frac{\sigma^2}{\eta_k} \frac{\beta^*}{1 - (K-1)\beta^*}, \quad (2.7)$$

où β^* est l'unique solution non-nulle de l'équation :

$$\gamma f'(\gamma) - f(\gamma) = 0, \quad (2.8)$$

qui est obtenue lorsque chaque émetteur maximise sa fonction d'utilité efficace énergétiquement en fonction de sa puissance d'émission.

Cet équilibre est qualifié de "non-saturé" si la puissance d'émission maximale de chaque émetteur, $P_k^{\max}$, est suffisamment élevée pour ne pas être atteinte à l'équilibre. En d'autres termes, chaque émetteur maximise son efficacité énergétique pour une puissance inférieure à $P_k^{\max}$. Une propriété importante de cet équilibre de Nash est que les émetteurs n'ont besoin de connaître que leur propre gain de canal η_k ainsi que le nombre d'émetteurs dans le système pour déterminer leur puissance d'émission d'équilibre. Cependant, il a été prouvé par [Goodman 2000, Meshkati 2006], que cet équilibre n'est pas optimal au sens de Pareto, c'est-à-dire que les efficacités énergétiques à l'équilibre ne dominent pas la région d'utilités atteignables.

2.2 Report stratégique du gain du canal

Dans cette partie, nous étudions un scénario de jeu statique construit à partir de la mesure d'efficacité énergétique présentée dans la partie 2.1. Une étude plus détaillée du problème d'allocation de puissance efficace énergétiquement que nous présentons ici se trouve dans [Mériaux 2012b]. En particulier, les preuves des différentes propriétés énoncées dans ce manuscrit sont présentes dans cet article.

Dans ce nouveau scénario, nous étudions les communications entre un émetteur (une station de base) et plusieurs récepteurs (des terminaux mobiles), sur des bandes de fréquence séparées. Disposant d'une puissance d'émission maximale fixée,

la station de base alloue cette puissance entre les différents terminaux mobiles de manière à maximiser l'efficacité énergétique de la cellule. Cette allocation est nécessairement réalisée en fonction des gains de canaux reportés par les terminaux à la station de base car les canaux ne sont pas symétriques (de la même manière que dans la norme LTE [3GPP 2009]). Les objectifs de la station de base et des terminaux sont différents puisque nous considérons que la station de base cherche à maximiser l'efficacité énergétique de la totalité de la cellule alors que les terminaux mobiles ne cherchent qu'à maximiser leur rapport signal à bruit (RSB). Ainsi, de manière rationnelle, les terminaux mobiles ont intérêt à reporter des gains de canaux qui favorisent leur intérêt individuel, même si cela diminue les performances globales du réseau.

Ce constat nous incite à étudier les effets de gains de canaux choisis de manière stratégique sur les performances de la cellule sans fil. Pour cela, nous modélisons le problème de report des gains de canaux sous forme d'un jeu statique dans lequel les joueurs sont les terminaux mobiles qui choisissent leur report de manière à maximiser la puissance qui leur est allouée par la station de base, les puissances allouées étant choisies par la station de base de manière à maximiser une fonction d'efficacité énergétique de la cellule. En limitant notre analyse à l'allocation de puissance, nous pouvons obtenir des résultats formels pour comparer une allocation de puissance avec les vrais gains de canaux et une allocation de puissance avec des gains de canaux reportés de manière stratégique. En conséquence, cette partie fournit des résultats importants pour les réseaux cellulaires efficace énergétiquement dans lesquels la fidélité des gains de canaux reportés n'est pas garantie.

En particulier, nos contributions sont les suivantes :

- Nous prouvons qu'il existe un unique équilibre de Nash pour le choix des gains de canaux reportés. Cela permet d'étudier les performances du réseau dans un état stable dont aucun terminal n'a unilatéralement intérêt à dévier.
- Nous comparons les performances obtenues avec les vrais gains de canaux et les reports stratégiques à l'équilibre. Pour les deux types de gains de canaux, nous donnons une formule explicite de l'efficacité énergétique de la cellule. Cela nous permet d'observer que les cellules de petite dimension sont plus robustes à des reports stratégiques de gains de canaux que des cellules avec des capacités d'émission plus importantes.

2.2.1 Modélisation du système

Modèle du réseau sans fil

Nous étudions une cellule dans laquelle une station de base alloue sa puissance d'émission entre K récepteurs mobiles, sur la bande descendante. Un utilisateur est identifié par son indice $k \in \mathcal{K} = \{1, \dots, K\}$ et la puissance de transmission allouée est notée $p_k \in [0, P]$ pour chaque récepteur. Le temps est divisé en intervalles et l'allocation de puissance est faite pour chaque intervalle. Pour étudier uniquement les conséquences de l'allocation de puissance, nous ignorons volontairement le problème

d'allocation de ressource sur plusieurs sous-bandes de fréquence et pour plusieurs intervalles de temps à la suite. En conséquence, chaque récepteur dispose de sa propre bande de largeur fixe W .

Sur chaque bande, nous supposons que le canal de la station de base vers le récepteur subit une atténuation quasi-statique avec un coefficient de canal h_k . Le processus d'atténuation est supposé iid Rayleigh, ce qui entraîne une distribution exponentielle pour les gains de canaux η_k . Pour chacun des récepteurs $k \in \mathcal{K}$, nous pouvons exprimer son rapport signal à bruit (RSB) instantané par :

$$\gamma_k(p_k, \eta_k) = \frac{p_k \eta_k}{\sigma^2}, \quad (2.9)$$

où σ^2 est la variance du bruit pour le récepteur k . Il est intéressant de noter que contrairement au modèle étudié en partie 2.1, les récepteurs ne souffrent pas d'interférence puisque les bandes de fréquences des récepteurs ne se superposent pas. Pour réaliser l'allocation de puissance, la station de base requiert un report de qualité de canal de la part de chaque récepteur, pour chaque intervalle de temps. Dans la présente étude, nous supposons que ce report est le gain du canal η_k , $\forall k \in \mathcal{K}$.

Efficacité énergétique de la cellule

La station de base réalise l'allocation de puissance, en prenant en compte l'efficacité énergétique associée à chacun des récepteurs mobiles. Contrairement à [Goodman 2000] où le contrôle de puissance est considéré sur le canal montant, nous étudions ici l'allocation de puissance sur le canal descendant. En conséquence, l'efficacité énergétique désigne ici le rapport entre la quantité d'information correctement transmise à un récepteur mobile et l'énergie dépensée par la station de base vers ce même récepteur. Nous notons $u_{ee}^{(k)}$ la fonction d'efficacité énergétique associée au $k^{\text{ième}}$ récepteur mobile. Elle s'écrit :

$$u_{ee}^{(k)}(p_k) = \frac{Rf(\gamma_k(p_k, \eta_k))}{p_k}, \quad (2.10)$$

où R est le débit fixe de transmission de la station de base en bit/s, et f est la fonction sigmoïde qui représente le taux de succès de la transmission de données. En ce qui concerne la puissance d'émission optimale par récepteur, nous pouvons vérifier que l'efficacité énergétique par récepteur peut être maximisée avec une puissance plus faible lorsque le gain du canal est meilleur. Pour l'allocation de puissance efficace énergétiquement à l'échelle de la cellule, cela signifie qu'il est plus intéressant d'allouer de la puissance aux récepteurs mobiles disposant d'un bon gain de canal, puisque ces récepteurs nécessitent moins de puissance pour offrir une bonne efficacité énergétique.

Nous pouvons à présent définir l'efficacité énergétique de la cellule qui est la somme de toutes les efficacités énergétiques par récepteur, c'est-à-dire :

$$u_{ee}(\underline{p}) = \sum_{k \in \mathcal{K}} u_{ee}^{(k)}(p_k), \quad (2.11)$$

où le vecteur $\underline{p}$ est composé de toutes les valeurs de puissance d'émission allouées par la station de base.

Les propriétés de l'efficacité énergétique de la cellule découlent des propriétés des efficacités énergétiques de chaque récepteur. En effet, pour tout $k \in \mathcal{K}$, la fonction $u_{ee}^{(k)}(p_k)$ est continue par rapport à p_k . Par conséquent, la somme (2.11) est continue par rapport à $\underline{p}$. Cependant, une telle analogie ne se retrouve pas pour une autre propriété de ces fonctions : bien que chaque fonction individuelle $u_{ee}^{(k)}(p_k)$ soit quasiconcave, leur somme n'est ni concave, ni quasiconcave. Le problème d'allocation optimal de puissance par la station de base n'est donc pas trivial. Nous invitons le lecteur à consulter [Mériaux 2012b] pour une étude plus détaillée de ce problème. Dans la suite de cette partie, nous considérons simplement que la station de base effectue une allocation optimale de puissance.

2.2.2 Jeu statique de report des gains de canaux

Définition du jeu

Dans cette partie, nous nous concentrons sur le comportement stratégique des récepteurs mobiles lorsqu'ils reportent l'état de leur canal. La différence principale avec les intérêts de la station de base est que les récepteurs mobiles ne sont pas concernés par l'efficacité énergétique sur le canal descendant puisque ce n'est pas leur propre puissance qui est utilisée. À la place de l'efficacité énergétique, ils ne cherchent qu'à maximiser leur propre RSB. Du point de vue d'un opérateur, l'énergie consommée par la station de base a bien plus d'importance que l'énergie consommée par les terminaux mobiles. Il y a donc une certaine logique à considérer que la station de base cherche à maximiser l'efficacité énergétique de la cellule alors que les récepteurs cherchent seulement à maximiser leur RSB.

Pour chaque récepteur mobile, le RSB est proportionnel à la puissance allouée par la station de base. Cette puissance allouée dépend du gain de canal que le récepteur en question reporte à la station de base, mais également des reports de tous les autres récepteurs mobiles dans la cellule. Puisque chaque récepteur préfère se voir allouer une forte puissance, il peut jouer avec la valeur du gain de canal qu'il reporte à la station de base afin de se voir allouer une puissance plus importante. Nous supposons donc que les récepteurs mobiles ont la liberté de reporter n'importe quel gain de canal à la station de base, et nous faisons appel à la théorie des jeux pour étudier le résultat d'un tel scénario. Pour chaque récepteur mobile $k \in \mathcal{K}$, deux valeurs de gain de canal sont importantes :

- la vraie valeur du gain de canal η_k , puisque que la capacité du canal du récepteur dépend de ce gain ;
- la valeur qu'il reporte à la station de base et que nous notons $g_k \in [0, G]$, avec G le gain maximal qu'un utilisateur peut reporter. La puissance d'émission est allouée au récepteur mobile en fonction de cette valeur.

Les deux termes apparaissent dans la fonction d'utilité de chaque récepteur, qui

est le RSB après l'allocation de puissance de la station de base. $\forall k \in \mathcal{K}$,

$$r_k(g_k, \underline{g}_{-k}) = \frac{p_k(g_k, \underline{g}_{-k})\eta_k}{\sigma^2}, \quad (2.12)$$

où $p_k(g_k, \underline{g}_{-k})$ est la puissance allouée par la station de base à l'émetteur k en fonction de tous les reports des gains de canaux. Nous pouvons maintenant définir le jeu étudié.

Définition 3 *Le jeu de report de canal est défini par le triplet $\mathcal{G}' = (\mathcal{K}, \{\mathcal{A}'_k\}_{k \in \mathcal{K}}, \{r_k\}_{k \in \mathcal{K}})$ dans lequel*

- $\mathcal{K}$ est l'ensemble des joueurs, qui représente les récepteurs mobiles.
- $\mathcal{A}'_k = [0, G]$ est l'ensemble des actions pour le joueur k . Dans ce jeu, une action $g_k \in \mathcal{A}'_k$ est le report du gain de canal à la station de base.
- La fonction d'utilité du joueur k est

$$r_k(g_k, \underline{g}_{-k}) = \frac{p_k(g_k, \underline{g}_{-k})\eta_k}{\sigma^2}.$$

c'est le RSB, qui dépend de la puissance allouée par la station de base en fonction de tous les reports des gains de canaux.

La figure 2.2 donne l'allure des fonctions d'utilité des joueurs de ce jeu en fonction du gain de canal reporté et de la puissance optimale associée pour maximiser l'efficacité énergétique. Il est intéressant de constater que lorsque qu'un récepteur reporte un gain de canal trop faible, aucune puissance ne lui est allouée. En effet, avec un canal trop faible, la puissance nécessaire pour maximiser l'efficacité énergétique du récepteur est forte et le gain d'efficacité énergétique est faible. La station de base préfère donc allouer sa puissance à d'autres récepteurs avec de meilleurs canaux. Au contraire, lorsqu'un récepteur reporte un gain de canal plus élevé, la station de base lui alloue de la puissance. Mais meilleur est le canal reporté, plus faible est la puissance allouée car nous rappelons qu'un récepteur avec un bon canal nécessite moins de puissance pour maximiser son efficacité énergétique (voir figure 2.1). Chaque récepteur est donc face à un compromis pour le choix du gain de canal qu'il reporte : il doit reporter un gain de canal suffisamment fort pour se voir allouer de la puissance, mais également suffisamment faible pour que cette puissance soit élevée.

Caractérisation de l'équilibre de Nash

Dans le jeu $\mathcal{G}'$, nous pouvons prouver qu'il existe un unique équilibre de Nash. La preuve de cette propriété peut se trouver dans [Mériaux 2012b].

Théorème 1 *Le jeu $\mathcal{G}'$ possède un unique équilibre de Nash. C'est le profil d'actions $\underline{g}^* = (g^*, \dots, g^*)$ où $g^* = \frac{Ka\sigma^2}{P}$ tel que*

$$\arg \max_p \frac{f(\gamma(p, g^*))}{p} = \frac{P}{K}. \quad (2.13)$$

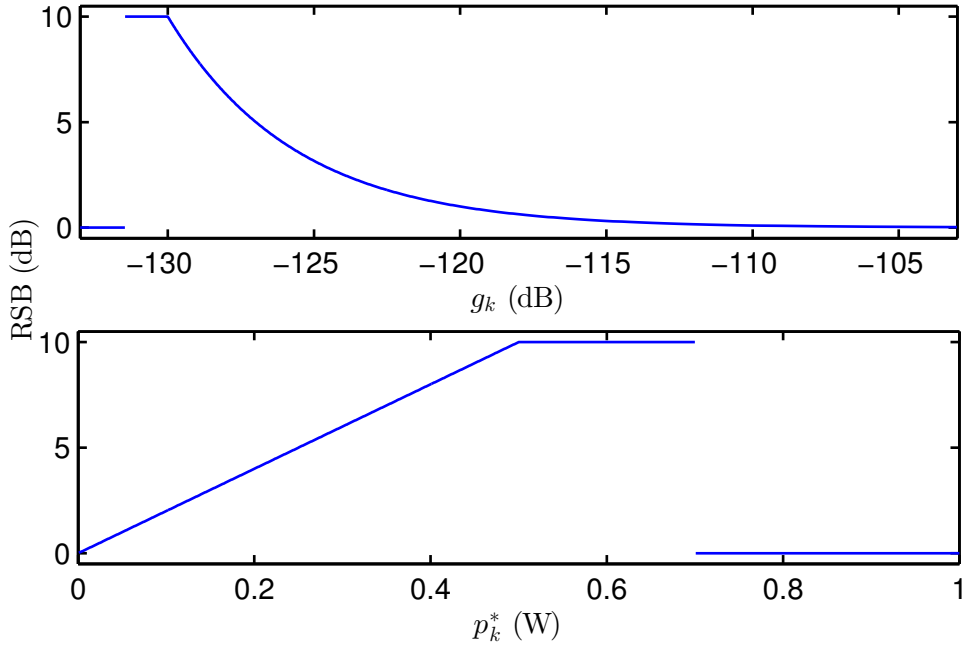


FIGURE 2.2 – RSB du récepteur k en fonction du gain de canal reporté et de la puissance optimale associée pour maximiser l'efficacité énergétique.

À cet équilibre, la puissance est uniformément allouée parmi les récepteurs mobiles et pour chaque récepteur $k \in \mathcal{K}$, le RSB s'écrit

$$r_k(\underline{g}^*) = \frac{P\eta_k}{K\sigma^2}. \quad (2.14)$$

Il est intéressant de noter qu'à cet équilibre, les récepteurs reportent tous le même état de canal et obtiennent la même puissance allouée par la station de base. De plus, cet équilibre dépend uniquement du nombre de récepteurs dans la cellule et de la contrainte de puissance de la station de base. Les gains des canaux réels n'apparaissent pas dans l'expression de cet équilibre.

Étant donné qu'à cet équilibre, les récepteurs mobiles ne reportent pas leurs gains de canal réels, il y a deux utilités efficaces énergétiquement d'intérêt pour la cellule. La première est l'utilité d'efficacité énergétique dans la cellule que la station de base pense correcte, étant donné les reports des récepteurs mobiles. Ce n'est pas la vraie valeur de l'utilité d'efficacité énergétique. Elle s'écrit :

$$\begin{aligned} u_{ee}\left(\frac{P}{K}, g^*\right) &= \sum_{k \in \mathcal{K}} \frac{K}{P} \exp\left(-\frac{aK\sigma^2}{Pg^*}\right), \\ &= \frac{K^2}{P} \exp(-1). \end{aligned} \quad (2.15)$$

Encore une fois, nous vérifions que cette utilité ne dépend que du nombre total de récepteurs dans la cellule et de la contrainte de puissance de la station de base. La deuxième utilité d'efficacité énergétique d'intérêt pour la cellule est évidemment

l'efficacité énergétique réelle de la cellule, qui dépend des gains de canaux réels. Cette utilité s'écrit :

$$u_{ee}\left(\frac{P}{K}, \eta_k\right) = \frac{K}{P} \sum_{k \in \mathcal{K}} \exp\left(-\frac{g^*}{\eta_k}\right). \quad (2.16)$$

Ainsi, nous pouvons exprimer le rapport entre l'efficacité énergétique réelle de la cellule et celle imaginée par la station de base :

$$\frac{u_{ee}\left(\frac{P}{K}, \eta_k\right)}{u_{ee}\left(\frac{P}{K}, g^*\right)} = \frac{1}{K} \sum_{k \in \mathcal{K}} \exp\left(1 - \frac{g^*}{\eta_k}\right). \quad (2.17)$$

Nous pouvons notamment vérifier que lorsque les gains reportés sont égaux aux gains réels, le rapport vaut 1 et si les gains reportés sont globalement plus élevés que les gains réels, le rapport devient très faible.

2.2.3 Résultats numériques

La figure 2.3 est issue d'un scénario avec deux récepteurs mobiles. Pour chacun d'entre eux, le gain de canal est fixé à -112 dB, et nous étudions comment l'efficacité énergétique de la cellule varie pour toutes les combinaisons possibles de puissances allouées par la station de base. Précisément, l'efficacité énergétique de la cellule est représentée en fonction du RSB de chaque récepteur mobile. La contrainte de puissance d'émission totale de la station de base est $P = 1$ W. D'une part, nous vérifions que l'efficacité énergétique de la cellule est maximisée lorsque les récepteurs mobiles reportent leurs gains de canaux réels. Mais dans ce cas, la contrainte de puissance de la station de base n'est pas saturée. En conséquence, ce n'est pas un point de Pareto pour les récepteurs mobiles. D'autre part, lorsque les récepteurs mobiles reportent les gains de canaux d'équilibre, un point de Pareto est atteint. Mais en contrepartie, l'efficacité énergétique de la cellule n'est pas maximale dans cette configuration.

Dans la figure 2.4, l'efficacité énergétique de la cellule est représentée en fonction du nombre de récepteurs mobiles dans la cellule. La variance du bruit, σ^2 est fixée à 5×10^{-14} W. Le paramètre a est fixé à 6. Les gains de canaux sont supposés suivre des distributions exponentielles. Pour chaque nombre de récepteurs mobiles dans la cellule, 10^4 réalisations sont calculées et les valeurs présentées sont moyennées sur ces réalisations. Comme pour la figure 2.3, nous comparons le cas pour lequel les récepteurs mobiles reportent leurs vrais gains de canal et le cas pour lequel ils modifient leurs reports de manière stratégique pour maximiser leur RSB. Nous étudions également l'influence de la contrainte de puissance de la station de base en considérant deux valeurs possibles : $P = 0,1$ W et $P = 1$ W. Sans surprise, pour chacune des contraintes de puissance, l'efficacité énergétique de la cellule est plus élevée quand les récepteurs reportent leurs vrais gains de canaux. Cette utilité augmente d'ailleurs avec le nombre de récepteurs dans la cellule simplement parce que c'est une somme sur l'ensemble des récepteurs. De façon intéressante, lorsque les récepteurs mobiles reportent leurs vrais gains de canaux, l'allocation de puissance

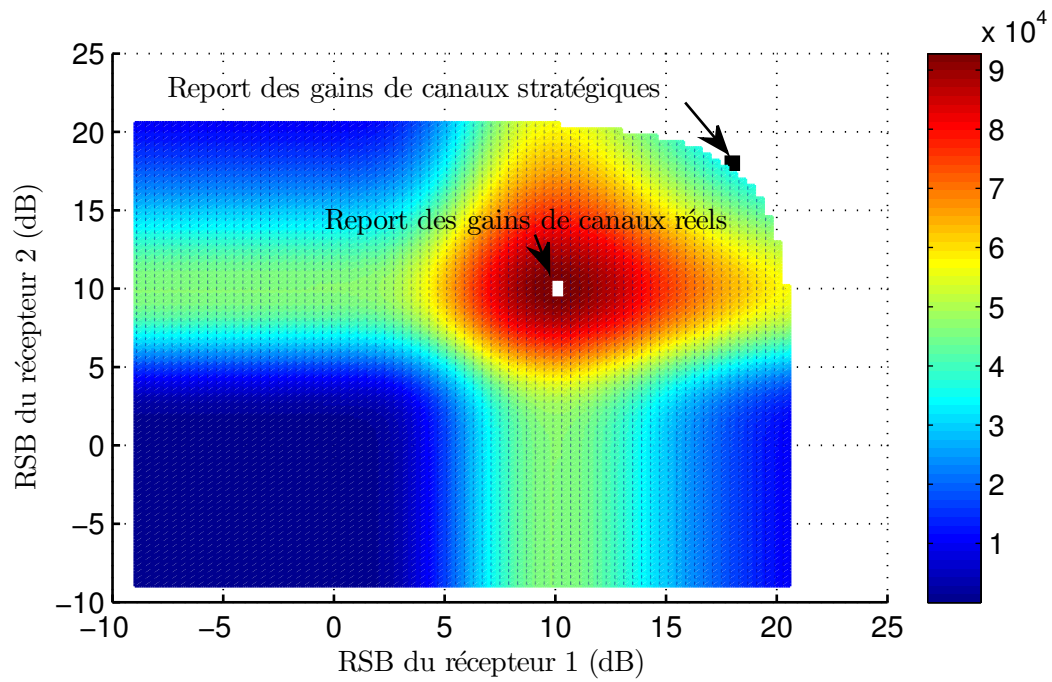


FIGURE 2.3 – Efficacité énergétique de la cellule en fonction des RSB des récepteurs 1 et 2. L'échelle de couleur représente l'utilité d'efficacité énergétique de la cellule en bit/joule. Les valeurs de paramètres sont fixées à $a = 10$, $\eta_1 = \eta_2 = -112$ dB, $P = 1$ W, et $\sigma^2 = 5 \times 10^{-14}$ W.

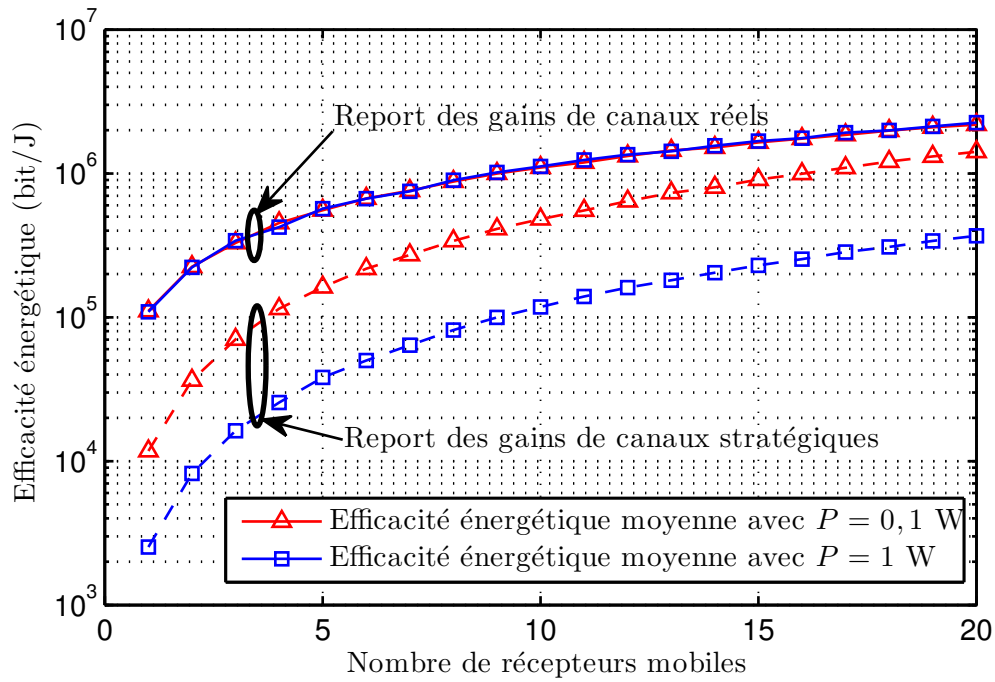


FIGURE 2.4 – Efficacité énergétique de la cellule en fonction du nombre de récepteurs mobiles.

par la station de base est la même quelle que soit la contrainte de puissance. Il y a deux explications à ce phénomène. Premièrement, quand la contrainte de puissance n'est pas saturée, la puissance allouée est la même quelle que soit la contrainte. Deuxièmement, nous rappelons que les récepteurs mobiles avec de mauvais gains de canaux ne sont pas intéressants en termes d'efficacité énergétique puisque la puissance qui maximise leur efficacité énergétique est élevée pour un gain faible d'efficacité énergétique. Allouer ou non de la puissance à ces récepteurs ne crée donc pas d'écart important en termes d'efficacité énergétique. Un autre phénomène qui mérite l'attention est que lorsque les récepteurs mobiles reportent les gains de canaux d'équilibre, l'efficacité énergétique de la cellule est moins bonne pour $P = 1 \text{ W}$ que pour $P = 0,1 \text{ W}$. Pour comprendre cela, nous rappelons qu'à l'équilibre, la contrainte de puissance de la station de base est saturée. Il y a donc plus de puissance allouée aux récepteurs mobiles dans le cas où $P = 1 \text{ W}$. Mais après un certain seuil, augmenter la puissance dans la cellule entraîne une diminution de l'efficacité énergétique de la cellule.

Dans la figure 2.5, le RSB moyen d'un récepteur mobile est représenté en fonction du nombre de récepteurs mobiles dans le même scénario que celui de la figure 2.4. Le RSB du récepteur mobile est donné en dB et pour chaque contrainte de puissance, le RSB est plus élevé quand les récepteurs mobiles reportent les gains de canaux d'équilibre. Comme le RSB des récepteurs mobiles à l'équilibre est proportionnel à la contrainte de puissance de la station de base, nous pouvons constater qu'il existe un gain de RSB entre les deux contraintes de puissance. Quand les récepteurs mobiles

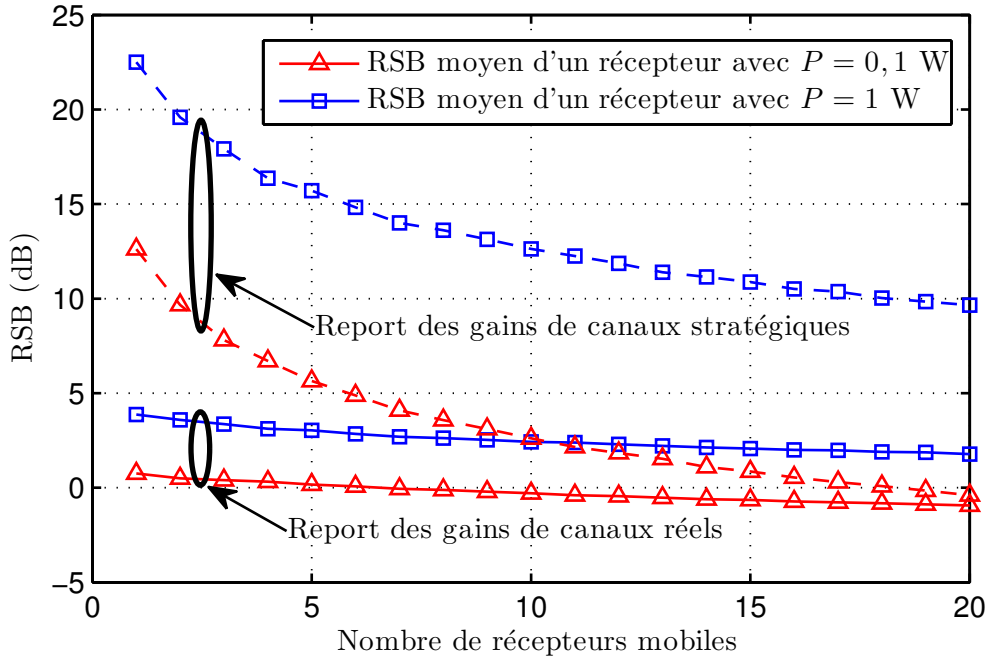


FIGURE 2.5 – RSB moyen d'un récepteur mobile en fonction du nombre de récepteurs mobiles dans la cellule.

reportent leurs gains de canaux réels, leur RSB est plus faible. Cependant, dans le cas où la contrainte de puissance de la station de base est plus élevée, cette dernière peut allouer de la puissance à plus de récepteurs. En conséquence, le RSB moyen des récepteurs mobiles en reportant leurs gains de canaux réels augmente avec la contrainte de puissance.

2.3 Jeu répété de contrôle de puissance

Dans cette partie, nous revenons au problème de contrôle de puissance efficace énergétiquement modélisé par un jeu statique en partie 2.1. Nous rappelons que cette formulation statique présente deux inconvénients majeurs :

- Bien que le jeu statique ait l'avantage de n'admettre qu'un unique équilibre de Nash dont il est facile d'obtenir une expression, cet équilibre ne donne pas des résultats satisfaisants en termes d'efficacité énergétique. En effet, les utilités d'équilibre ne sont pas optimales au sens de Pareto.
- Cette formulation ne permet pas de tenir compte de l'évolution des canaux au cours du temps et le fait que les terminaux adaptent leur puissance d'émission sur chaque paquet. En effet, considérer que cette évolution temporelle peut être modélisée par une séquence de jeux statiques indépendants les uns des autres n'est pas suffisant puisque la répétition du jeu fait intervenir des interactions stratégiques entre les joueurs qui ne sont pas présentes dans une séquence de jeux statiques.

Pour les deux raisons énoncées ci-dessus, la raison d'être de cette partie est de présenter une extension de la formulation initiale du problème de contrôle de puissance efficace énergétiquement : le jeu répété [Sorin 1992], qui offre l'avantage de tenir compte des interactions stratégiques des terminaux au cours du temps. En effet, l'avantage de la formulation répétée est que chaque émetteur conserve l'histoire du jeu. Il est donc possible que les émetteurs suivent des stratégies de contrôle de puissance globalement plus efficaces que l'équilibre de Nash du jeu statique. Dans cette partie, nous présentons le modèle de jeu répété utilisé dans [Le Treust 2010], ainsi qu'une stratégie de contrôle de puissance issue de ces travaux.

2.3.1 Définition du jeu répété

Dans le jeu répété que nous introduisons dans cette partie, les auteurs de [Le Treust 2010] proposent une stratégie de contrôle de puissance qui est globalement plus efficace énergétiquement au sens de Pareto que l'équilibre de Nash de la formulation statique. Ce gain d'efficacité énergétique est permis par le fait que dans un jeu répété, l'utilité que les joueurs cherchent à maximiser est l'espérance de leur efficacité énergétique tout au long des étapes successives du jeu. Ainsi, les émetteurs ne choisissent pas leur puissance d'émission pour maximiser un gain instantané mais pour assurer des performances sur le long terme, ce qui les incite à émettre à des valeurs de puissance qui limitent l'interférence dans le réseau. Afin de définir cette utilité, nous avons d'abord besoin de définir l'histoire du jeu et la stratégie d'un joueur.

Commençons par rappeler qu'un jeu répété se joue par étapes, qui correspondent aux instants auxquels les joueurs peuvent choisir leurs actions. En d'autres termes, les étapes du jeu correspondent à la mise à jour de la puissance d'émission des terminaux mobiles pour chaque paquet à transmettre. L'histoire du jeu représente la mémoire que chaque émetteur garde des étapes précédentes du jeu répété. Précisément, l'histoire d'un émetteur est composée des valeurs de ses puissances d'émission lors des étapes précédentes, ainsi que d'un signal public partagé par tous les émetteurs à chaque étape n du jeu répété :

$$\omega_n = \sigma^2 + \sum_{k=1}^K \eta_{k,n} p_{k,n}, \quad (2.18)$$

où $\eta_{k,n}$ et $p_{k,n}$ sont respectivement le gain du canal et la puissance d'émission de l'émetteur k à l'étape n . Ce signal public présente l'avantage de ne pas nécessiter d'émission supplémentaire dans le réseau puisque chaque émetteur peut l'obtenir par la connaissance de sa puissance d'émission, son gain de canal et son RSIB. Ce signal public est une fonction déterministe des puissances d'émission de tous les émetteurs dans la cellule. Nous notons $\Omega = \left[\sigma^2, \sigma^2 + \sum_{k=1}^K \eta_k^{\max} P_k^{\max} \right]$ l'intervalle dans lequel ce signal est défini.

Définition 4 *Le couple de vecteurs*

$$(\underline{\omega}_n, \underline{p}_{k,n}) = (\omega_1, \dots, \omega_{n-1}, p_{k,1}, \dots, p_{k,n-1}) \quad (2.19)$$

est appelé *histoire du jeu pour l'émetteur k à l'étape n* et appartient à l'ensemble $\mathcal{H}_n = \Omega^{n-1} \times \mathcal{A}_k^{n-1}$.

Lors du choix d'un niveau de puissance d'émission à l'étape n , nous supposons que chaque émetteur connaît sa propre histoire. Les émetteurs choisissent donc leur action à partir de leur histoire $(\underline{\omega}_n, \underline{p}_{k,n})$. À présent, nous pouvons définir les stratégies pures des émetteurs dans le jeu répété.

Définition 5 *Une stratégie pure τ_k pour le joueur $k \in \mathcal{K}$ est une séquence de fonctions causales $(\tau_{k,n})_{n \geq 1}$ avec*

$$\begin{aligned} \tau_{k,n} : \quad \mathcal{H}_n &\rightarrow \mathcal{A}_k \\ (\underline{\omega}_n, \underline{p}_{k,n}) &\mapsto p_{k,n}. \end{aligned} \quad (2.20)$$

Le vecteur des stratégies $\underline{\tau} = (\tau_1, \dots, \tau_K)$ est un profil de stratégies des émetteurs de la cellule. Un profil de stratégies $\underline{\tau}$ induit naturellement un unique plan d'action $(\underline{p}_n)_{n \geq 1}$ et une efficacité énergétique unique pour chaque émetteur. En effet, à chaque profil de puissances d'émission $\underline{p}_n = (p_{1,n}, \dots, p_{K,n})$ correspond une unique efficacité énergétique $u_k(\underline{p}_n)$ pour chaque émetteur. Nous donnons maintenant la définition de la fonction d'utilité pour le jeu répété qui est l'espérance de l'efficacité énergétique de chaque émetteur pour la durée du jeu. Nous présentons ici la fonction d'utilité du jeu escompté qui est un cas particulier de jeu répété¹.

Définition 6 *Soit $\underline{\tau} = (\tau_1, \dots, \tau_K)$ un profil de stratégie, l'utilité du joueur $k \in \mathcal{K}$ est définie par :*

$$v_k(\underline{\tau}) = \sum_{n=1}^{\infty} \lambda(1 - \lambda)^{n-1} u_k(\underline{p}_n), \quad (2.21)$$

où $\underline{p}_n$ est le profil des puissances d'émission induites par le profil de stratégie $\underline{\tau}$ à l'étape n . Le paramètre $0 < \lambda < 1$ est le facteur d'escompte du jeu. Il peut être interprété comme une probabilité d'arrêt du jeu ou comme une pondération entre l'efficacité énergétique instantanée et l'efficacité énergétique à long terme. Ce paramètre est supposé identique pour tous les émetteurs.

À présent, nous pouvons définir une stratégie d'équilibre du jeu répété et donner une stratégie d'équilibre explicite.

1. Les résultats exprimés dans cette partie existent également pour le jeu répété fini et le jeu infiniment répété, voir [Le Treust 2010].

2.3.2 Stratégie d'équilibre du jeu répété

Définition 7 *Un profil de stratégie $\underline{\tau}^*$ est un équilibre du jeu répété défini par $(\mathcal{K}, \{\mathcal{T}_k\}_{k \in \mathcal{K}}, \{v_k\}_{k \in \mathcal{K}})$ si $\forall k \in \mathcal{K}, \forall \tau_k \in \mathcal{T}_k$,*

$$v_k(\underline{\tau}^*) \geq v_k(\tau_k, \underline{\tau}_{-k}^*). \quad (2.22)$$

Contrairement au cas de la formulation statique du contrôle de puissance efficace énergétiquement, il existe plusieurs stratégies d'équilibre dans le jeu répété. La stratégie consistant à tout le temps émettre au niveau de puissance défini dans la partie 2.1 est d'ailleurs un des équilibres du jeu répété. Cependant l'efficacité énergétique de cette stratégie n'est pas satisfaisante. Dans [Le Treust 2010], une autre stratégie de contrôle de puissance d'équilibre est proposée. Cette stratégie est basée sur un point de fonctionnement qui offre une meilleure efficacité énergétique que le point d'équilibre de la formulation statique. Les auteurs partent du constat qu'à l'équilibre de Nash du jeu statique, le produit $p_k^* \eta_k$ est constant pour tous les émetteurs. Ils ajoutent alors cette égalité comme une contrainte à vérifier pour résoudre $\frac{\partial u_k}{\partial p_k}(\underline{p}) = 0$. L'unique solution du système associé est :

$$\forall k \in \mathcal{K}, \tilde{p}_k = \frac{\sigma^2}{\eta_k} \frac{\tilde{\gamma}}{1 - (K-1)\tilde{\gamma}} \quad (2.23)$$

où $\tilde{\gamma}$ est l'unique solution non-nulle de $\gamma[1 - (K-1)\gamma]f'(\gamma) - f(\gamma) = 0$.

Il est possible de vérifier que cette puissance d'émission offre une efficacité énergétique meilleure que la puissance de l'équilibre de Nash statique. À partir de ce point de fonctionnement, une stratégie d'équilibre peut être définie dans le jeu escompté.

Théorème 2 *Supposons que la condition suivante est respectée :*

$$\lambda \leq \frac{\eta_k^{\min} \delta(\beta^*, \tilde{\gamma})}{\eta_k^{\min} \delta(\beta^*, \tilde{\gamma}) + \eta_k^{\max} [(K-1)f(\beta^*) - \delta(\beta^*, \tilde{\gamma})]} \quad (2.24)$$

où $\delta(\beta^*, \tilde{\gamma}) = \frac{1-(K-1)\tilde{\gamma}}{\tilde{\gamma}} f(\tilde{\gamma}) - \frac{1-(K-1)\beta^*}{\beta^*} f(\beta^*)$. Alors, pour tout $k \in \mathcal{K}$, la stratégie suivante est un équilibre de Nash du jeu escompté pour n'importe quelle distribution des gains de canaux des émetteurs :

$$\forall i \geq 1, \tilde{\tau}_{k,n} = \begin{cases} \tilde{p}_k & \text{si } \forall n' < n, \omega_{n'} = \sigma^2 + K\sigma^2 \frac{\tilde{\gamma}}{1-(K-1)\tilde{\gamma}} \\ p_k^* & \text{autrement} \end{cases} \quad (2.25)$$

Plusieurs remarques sont utiles pour comprendre ce théorème :

- Selon les deux interprétations possibles de λ , l'inégalité qui doit être satisfaite signifie que la probabilité d'arrêt du jeu doit être suffisamment faible ou que l'efficacité énergétique instantanée doit être suffisamment moins importante que l'efficacité énergétique à long terme. Dans les deux cas, cela signifie que le futur du jeu répété doit être considéré comme important.

- Si un émetteur dévie de la stratégie de contrôle de puissance, tous les autres émetteurs utilisent la puissance d'émission correspondant à l'équilibre de Nash statique jusqu'à la fin du jeu. Pour que la stratégie proposée soit un équilibre du jeu répété, il est donc nécessaire que l'efficacité énergétique que peut gagner un émetteur en déviant de la stratégie à une étape donnée soit plus faible que ce qu'il perd si les autres émetteurs utilisent la puissance d'émission d'équilibre de la formulation statique jusqu'à la fin du jeu. La condition à vérifier par λ provient de cette nécessité.
- Enfin, chaque émetteur peut détecter si un autre émetteur dévie de la stratégie proposée grâce au signal public.

Bien que la formulation du problème de contrôle de puissance par un jeu répété permet d'obtenir une stratégie d'équilibre plus efficace énergétiquement que l'équilibre de la formulation statique, ce cadre de travail ne permet pas de tenir compte des variations des canaux à chaque étape du jeu de manière optimale. Pour cette raison, la stratégie proposée dans [Le Treust 2010] est optimale au sens de Pareto lorsque les canaux restent fixes tout au long du jeu. Mais si les canaux varient, cette stratégie n'est plus optimale. Pour pallier à cette sous-optimalité, nous proposons dans la partie 2.4 de modéliser le problème de contrôle de puissance par un jeu répété stochastique, qui permet d'inclure la variation des canaux dans le modèle du jeu.

2.4 Jeu stochastique de contrôle de puissance

Les auteurs de [Le Treust 2010] ont démontré qu'un modèle de jeu répété permet d'avoir une modélisation plus fine du problème, modélisation qui conduit à des solutions plus efficaces globalement. L'idée fondamentale et nouvelle en contrôle de puissance, et que nous exploitons dans cette partie, est qu'il ne faut pas supposer le contrôle de puissance indépendant d'un paquet à l'autre, et ceci même si les réalisations des gains des canaux sont indépendantes. Un modèle de jeu dynamique tel que le jeu répété permet de tenir compte du fait que les émetteurs interagissent plusieurs fois et ceci conduit à créer une corrélation entre les niveaux de puissances choisis par un émetteur au cours du temps. La contribution de cette partie est de généraliser les travaux de [Le Treust 2010] en relaxant une hypothèse de normalisation de l'utilité individuelle par le gain de canal. Pour faire cela, nous utilisons un modèle de jeux stochastiques [Shapley 1953], ce qui nous amène à supprimer la sous-optimalité en termes de performances induite par la normalisation nécessaire au modèle de jeu répété. Les travaux de [Fudenberg 2011, Hörner 2011] sont alors utilisés pour obtenir un *Folk* théorème qui caractérise la région des utilités atteignables de ce jeu stochastique. Nous présentons également une stratégie de contrôle de puissance explicite pour ce jeu. L'ensemble des résultats présentés dans cette partie sont issus de [Mériaux 2011c].

2.4.1 Définition du jeu stochastique

Dans la formulation stochastique du jeu de contrôle de puissance efficace énergétiquement, l'état des gains des canaux $\underline{\eta}_n = (\eta_{1,n}, \dots, \eta_{K,n}) \in \Gamma$ avec $\Gamma = \Gamma_1 \times \dots \times \Gamma_K$, peut varier dans un ensemble discret selon une loi d'évolution donnée (par exemple de manière iid, pour chaque paquet à transmettre). Les étapes du jeu stochastique correspondent aux instants auxquels les émetteurs peuvent mettre à jour leur puissance d'émission. D'une étape à l'autre, nous supposons que l'état des canaux $\underline{\eta}_n$ varie dans un ensemble discret en suivant une distribution de probabilités de transitions π . La chaîne de Markov engendrée par cette distribution de transition est supposée irréductible, c'est-à-dire que pour tout couple d'états de canaux $\underline{\eta}$ et $\underline{\eta}'$, nous avons $\pi(\underline{\eta}' | \underline{\eta}) > 0$. Enfin, un signal public est accessible à tous les émetteurs à chaque étape du jeu stochastique. Deux cas particuliers de signaux publics sont : (i) le niveau de puissance reçue par la station de base $\sum_{k=1}^K \eta_{k,n} p_{k,n} + \sigma^2$, (ii) la totalité de l'état des canaux ainsi que les puissances d'émission des autres émetteurs, c'est-à-dire $(\underline{\eta}_n, \underline{p}_n)$ (on parle alors de jeu avec observation parfaite). D'autres signaux publics sont bien sûr possibles et nous en introduirons d'ailleurs un en particulier dans la partie 2.4.2 pour mettre en place une stratégie d'équilibre avantageuse.

Nous pouvons à présent donner la définition du jeu stochastique.

Définition 8 *Un jeu stochastique est défini par :*

$$(\mathcal{K}, (\mathcal{T}_k)_{k \in \mathcal{K}}, (v_k)_{k \in \mathcal{K}}, (\Gamma_k)_{k \in \mathcal{K}}, \pi, \Theta, \phi), \quad (2.26)$$

avec $\mathcal{K}$ l'ensemble des joueurs, $\mathcal{T}_k$ l'ensemble des stratégies pour le joueur k , v_k la fonction d'utilité du joueur k sur le long terme, Γ_k l'intervalle des états de canaux accessibles au joueur k , π la probabilité de transition sur les états, Θ l'espace des observations et ϕ le signal public accessible à tous les joueurs.

Le jeu se déroule de la manière suivante : à partir d'une étape initiale associée à un état des canaux $\underline{\eta}_1$, l'état évolue d'une étape $n-1$ à une étape n selon $\pi(\cdot | \underline{\eta}_{n-1}) \in \Delta(\Gamma)$. À chaque étape, les émetteurs choisissent leur niveau de puissance d'émission en fonction du signal public :

$$\begin{aligned} \phi : \quad \Gamma \times \mathcal{A} &\rightarrow \Theta \\ (\underline{\eta}_n, \underline{p}_n) &\mapsto \theta_n \end{aligned} \quad (2.27)$$

où $\mathcal{A} = \mathcal{A}_1 \times \dots \times \mathcal{A}_K$.

La stratégie et l'utilité sur le long terme du joueur k sont définies comme suit.

Définition 9 *La stratégie du joueur $k \in \mathcal{K}$ est une suite de fonctions $(\tau_{k,n})_{n \geq 1}$ avec*

$$\begin{aligned} \tau_{k,n} : \quad \Theta^{n-1} \times \Gamma_k &\rightarrow \mathcal{A}_k \\ (\underline{\theta}_n, \eta_{k,n}) &\mapsto p_{k,n}, \end{aligned} \quad (2.28)$$

avec $\underline{\theta}_n = (\theta_1, \dots, \theta_{n-1})$ la suite des signaux publics émis jusqu'à l'étape n . Autrement dit, à l'histoire $(\underline{\theta}_n, \eta_{k,n}) \in \Theta^n$ (observations passées et état présent), on associe une action $p_{k,n} \in \mathcal{A}_k$.

La stratégie du joueur k est notée τ_k et le vecteur de stratégies $\underline{\tau} = (\tau_1, \dots, \tau_K)$ est un profil de stratégies. Un profil de stratégies $\underline{\tau}$ entraîne une unique séquence d'actions $(p_n)_{n \geq 1}$.

Définition 10 Soit $\underline{\tau}$ un profil de stratégies. L'utilité du joueur $k \in \mathcal{K}$ sachant que l'état initial du canal est $\underline{\eta}_1$ est définie par

$$v_k(\underline{\tau}, \underline{\eta}_1) = \sum_{n \geq 1} \lambda(1 - \lambda)^{n-1} \mathbb{E}_{\underline{\tau}, \pi} [u_k(p_n, \underline{\eta}_n) | \underline{\eta}_1] \quad (2.29)$$

où $(p_n)_{n \geq 1}$ est la séquence de puissances d'émission induite par le profil de stratégies $\underline{\tau}$. Nous reprenons ici une fonction d'utilité escomptée. L'interprétation du paramètre λ est la même que dans la partie 2.3.

À présent, définissons le concept d'équilibre de Nash d'un jeu stochastique pour un état initial est $\underline{\eta}_1$.

Définition 11 Un profil de stratégies $\underline{\tau}^*$ est un équilibre du jeu stochastique pour un état initial $\underline{\eta}_1$ si

$$\forall k \in \mathcal{K}, \forall \tau_k, v_k(\underline{\tau}^*, \underline{\eta}_1) \geq v_k(\tau_k, \underline{\tau}_{-k}^*, \underline{\eta}_1). \quad (2.30)$$

2.4.2 Résultats analytiques

Folk théorème

Nous donnons maintenant une caractérisation de l'ensemble des utilités d'équilibre pour le jeu stochastique. Pour cela, nous définissons les niveaux d'utilité minmax $\tilde{v}_k(\underline{\eta}_1)$ pour le joueur $k \in \mathcal{K}$ comme les niveaux d'utilité d'équilibre les moins favorables au joueur $k \in \mathcal{K}$ lorsque l'état initial du jeu est $\underline{\eta}_1$. En d'autres termes,

$$\tilde{v}_k(\underline{\eta}_1) = \inf_{\underline{\tau}_{-k}} \sup_{\tau_k} v_k(\tau_k, \underline{\tau}_{-k}, \underline{\eta}_1). \quad (2.31)$$

D'autre part, l'ensemble des utilités atteignables est notée $F(\underline{\eta}_1)$. Le résultat énoncé par Dutta dans [Dutta 1995] indique que si la chaîne de Markov induite par π est irréductible, alors les niveaux d'utilité minmax et l'ensemble des utilités atteignables sont indépendants de l'état initial $\underline{\eta}_1$.

$$\lim_{\lambda \rightarrow 0} \tilde{v}_k(\underline{\eta}_1) = \tilde{v}_k, \forall \underline{\eta}_1 \quad (2.32)$$

$$\lim_{\lambda \rightarrow 0} F(\underline{\eta}_1) = F, \forall \underline{\eta}_1 \quad (2.33)$$

À partir de ces deux ensembles, nous pouvons définir l'ensemble des utilités atteignable et individuellement rationnelles.

Définition 12 *L'ensemble des utilités atteignables et individuellement rationnelles est :*

$$F^* = \{\underline{x} \in F \mid x_k \geq \tilde{v}_k, \forall k \in \mathcal{K}\} \quad (2.34)$$

Le *Folk* théorème nous donne la correspondance entre cet ensemble et l'ensemble des utilités d'équilibre.

Théorème 3 *Supposons que tous les joueurs voient le même signal public. Alors pour tout profil d'utilité $\underline{v} \in F^*$, il existe λ_0 tel que pour tout $\lambda < \lambda_0$, il existe une stratégie d'équilibre du jeu stochastique dont l'utilité à long terme vaut $\underline{v} \in F^*$.*

La preuve est basée sur le théorème 2 dans [Hörner 2011]. Il faut noter qu'une telle caractérisation de la région d'utilités d'équilibre est très puissante. En effet, la technique classique pour obtenir la région d'utilités d'équilibre consisterait à déterminer toutes les stratégies possibles pour les joueurs puis de calculer les utilités correspondantes. Dans un jeu très simple où chaque joueur n'aurait le choix qu'entre deux niveaux de puissance à chaque étape, il faudrait considérer 2^N stratégies possibles, avec N le nombre d'étapes du jeu. D'après [Dutta 1995], le *Folk* théorème nous autorise à considérer uniquement les stratégies dites de Markov sans perte d'optimalité pour obtenir la région d'équilibre. Nous rappelons qu'une stratégie de Markov est une stratégie pour laquelle les actions choisies à un étape donnée ne dépendent que de l'état du jeu à cette étape. Le nombre de stratégies à étudier se réduit donc à $2^{|\Gamma|}$ avec $|\Gamma|$ le nombre d'états de canaux.

Plan de sélections des meilleurs utilisateurs

Obtenir une région d'utilités atteignables est une chose, mais il nous reste à déterminer formellement des stratégies efficaces dans cette région. C'est ce que nous proposons de faire avec l'introduction d'un plan de *Sélection des Meilleurs Utilisateurs* (SMU).

La stratégie proposée est basée sur le point de fonctionnement présenté dans la partie 2.3 :

$$\forall k \in \mathcal{K}, \tilde{p}_{k,n} = \frac{\sigma^2}{\eta_{k,n}} \frac{\tilde{\gamma}}{1 - (K-1)\tilde{\gamma}}. \quad (2.35)$$

Contrairement au cas du jeu répété où les gains des canaux sont constants, quand ces derniers varient à chaque étape, la stratégie consistant à ce que chaque joueur émette au point de fonctionnement (2.35) n'est plus optimale. Il se trouve qu'on obtient de meilleurs résultats en termes de bien-être social (c'est-à-dire lorsqu'on cherche à maximiser la somme des efficacités énergétiques des émetteurs) si l'ensemble des joueurs émettant au point de fonctionnement est réduit. Cette approche est intitulée stratégie SMU, elle est caractérisée de la manière suivante.

À chaque étape n du jeu, le récepteur fixe $\mathcal{K}^n \subset \mathcal{K}$, l'ensemble des joueurs émettant au point de fonctionnement (2.35) pour maximiser la somme des utilités instantanées des émetteurs. Pour chaque joueur $k \in \mathcal{K}$:

- Si $k \in \mathcal{K}^n$, il lui est recommandé d'émettre au point de fonctionnement (2.35) à l'étape n .

- Si $k \notin \mathcal{K}^n$, il lui est demandé de ne pas émettre à cette étape.

Il faut bien noter que le comportement des émetteurs n'est pas imposé, le récepteur envoie seulement des recommandations aux émetteurs. Un mécanisme de punition est établi si un émetteur dévie de la stratégie : les autres émetteurs jouent l'équilibre de Nash statique pour le restant du jeu.

Propriétés du plan SMU

Bien qu'on puisse penser que la stratégie SMU nécessite des opérations complexes au niveau du récepteur, il se trouve que ce dernier n'a pas à comparer toutes les combinaisons possibles d'émetteurs. En effet, l'ensemble optimal des émetteurs de cette stratégie comprend toujours l'émetteur avec le meilleur gain de canal et un certain nombre d'autres émetteurs parmi ceux avec les meilleurs gains de canaux. Ainsi, le récepteur a uniquement besoin de classer les gains de canaux des émetteurs par ordre décroissant et fixer le gain de canal minimum pour séparer les émetteurs qui ont intérêt à émettre des autres. Cette opération est beaucoup moins coûteuse qu'une comparaison de toutes les combinaisons possibles et plus particulièrement lorsque le nombre d'émetteurs est grand. En termes de complexité, pour K émetteurs, au lieu de comparer 2^K combinaisons, le récepteur ne doit en comparer que K .

Proposition 1 *À débit d'émission égal, l'ensemble de k émetteurs qui offre les meilleures performances d'efficacité énergétique en émettant au point de fonctionnement (2.35) est l'ensemble des k émetteurs avec les meilleurs gains de canaux.*

La preuve de cette proposition est donnée dans [Mériaux 2011c].

Pour que la stratégie SMU soit un équilibre du jeu stochastique, il est nécessaire qu'aucun émetteur n'ait intérêt à dévier du plan d'action proposé. En d'autres termes, l'équilibre de la stratégie est assuré si le maximum (en termes d'utilité) que peut gagner un joueur en déviant à une étape du jeu est inférieur à ce qu'il va perdre en étant puni par les autres joueurs jusqu'à la fin du jeu. Nous obtenons alors la condition d'équilibre suivante :

Théorème 4 *La stratégie SMU est un équilibre du jeu stochastique si $\forall k \in \mathcal{K}$*

$$\lambda \leq \frac{\mathbb{E}[u_k(\underline{p}^{\text{smu}}, \underline{\eta})] - \mathbb{E}[u_k(\underline{p}^*, \underline{\eta})]}{\frac{R\eta_{\max}}{\sigma^2} \frac{f(\beta^*)}{\beta^*} + \mathbb{E}[u_k(\underline{p}^{\text{smu}}, \underline{\eta})] - \mathbb{E}[u_k(\underline{p}^*, \underline{\eta})]}, \quad (2.36)$$

avec $\underline{p}^{\text{smu}}$ le profil de puissance résultant de l'application de la stratégie SMU et $\underline{p}^*$ et β^* respectivement le profil de puissance et le RSIB correspondant à l'équilibre de Nash de la formulation statique.

L'efficacité énergétique issue de l'application de la stratégie SMU peut être comparée aux efficacités énergétiques issues de la stratégie d'équilibre de la formulation statique, de la stratégie de partage de temps pour laquelle seul l'émetteur avec le meilleur gain de canal émet (et dont le profil de puissance d'émission est noté $\underline{p}^{\text{ts}}$) et de la stratégie où tous les émetteurs émettent au point de fonctionnement.

Proposition 2 *À débit d'émission égal et pour des gains de canaux iid, nous avons :*

$$\forall k \in \mathcal{K}, \mathbb{E}[u_k(\underline{p}^{\text{smu}}, \underline{\eta})] \geq \mathbb{E}[u_k(\underline{p}^*, \underline{\eta})] \quad (2.37)$$

$$\forall k \in \mathcal{K}, \mathbb{E}[u_k(\underline{p}^{\text{smu}}, \underline{\eta})] \geq \mathbb{E}[u_k(\tilde{\underline{p}}, \underline{\eta})] \quad (2.38)$$

$$\forall k \in \mathcal{K}, \mathbb{E}[u_k(\underline{p}^{\text{smu}}, \underline{\eta})] \geq \mathbb{E}[u_k(\underline{p}^{\text{ts}}, \underline{\eta})] \quad (2.39)$$

Information nécessaire à l'établissement du plan SMU

Dans le plan SMU, les émetteurs adaptent leurs puissances d'émission au niveau du point de fonctionnement (2.35) si le récepteur leur recommande d'émettre. Cela signifie qu'ils ont besoin de connaître la recommandation du récepteur. De plus, pour émettre le bon niveau de puissance (2.35), ils doivent aussi connaître le nombre d'émetteurs qui vont être actifs en même temps qu'eux, ainsi que leur gain de canal. Nous considérons donc dans le plan SMU que le récepteur envoie les recommandations à tous les émetteurs de la cellule, ainsi que le nombre d'émetteurs auxquels il recommande d'émettre. En ce qui concerne les gains de canaux, nous supposons que chaque émetteur connaît son propre canal. Le dernier signal nécessaire à l'établissement du plan SMU permet à chaque émetteur de détecter si un autre émetteur dévie du plan d'action. Comme dans la partie 2.3, ce signal peut simplement être le RSIB de chaque émetteur.

2.4.3 Résultats numériques

Pour l'obtention de résultats numériques, nous utilisons la fonction d'efficacité $f(\gamma) = e^{-\frac{\alpha}{\gamma}}$, définie dans la partie 2.1.

La figure 2.6 est obtenue pour un scénario comprenant 10 émetteurs dont les gains de canaux peuvent seulement atteindre deux états $\eta_{\min}$ et $\eta_{\max}$ avec des probabilités $\{\frac{1}{2}, \frac{1}{2}\}$. Elle donne les gains respectifs de la stratégie issue du jeu répété et du plan SMU par rapport à l'équilibre de Nash issu du jeu statique, en fonction du rapport entre $\eta_{\max}$ et $\eta_{\min}$. Cette figure montre que la stratégie SMU est particulièrement intéressante lorsque l'écart entre les gains de canaux est important. En effet, nous pouvons vérifier que pour un rapport $\frac{\eta_{\max}}{\eta_{\min}} = 1$, l'efficacité énergétique du plan SMU est la même que pour le point de fonctionnement (2.35) et pour $\frac{\eta_{\max}}{\eta_{\min}} > 1$, le plan SMU devient plus efficace. Dans le cas plus général où les gains de canaux peuvent prendre de nombreuses valeurs différentes, les réalisations des canaux comprennent des valeurs hétérogènes, ce qui est donc favorable au plan SMU.

La figure 2.7 illustre la région atteignable pour un jeu à 2 joueurs et 2 états de canaux (avec $\frac{\eta_{\max}}{\eta_{\min}} = 4$). La frontière *minmax* délimite la région d'équilibre. Les utilités moyennes de SMU, du point de fonctionnement et de l'équilibre de Nash de la formulation statique sont également représentées à l'intérieur de cette région. Notons que la stratégie SMU Pareto-domine les autres stratégies considérées.

La simulation présentée en figure 2.8 compare les utilités instantanées moyennes de quatre mécanismes de contrôle de puissance en fonction du nombre d'émetteurs.

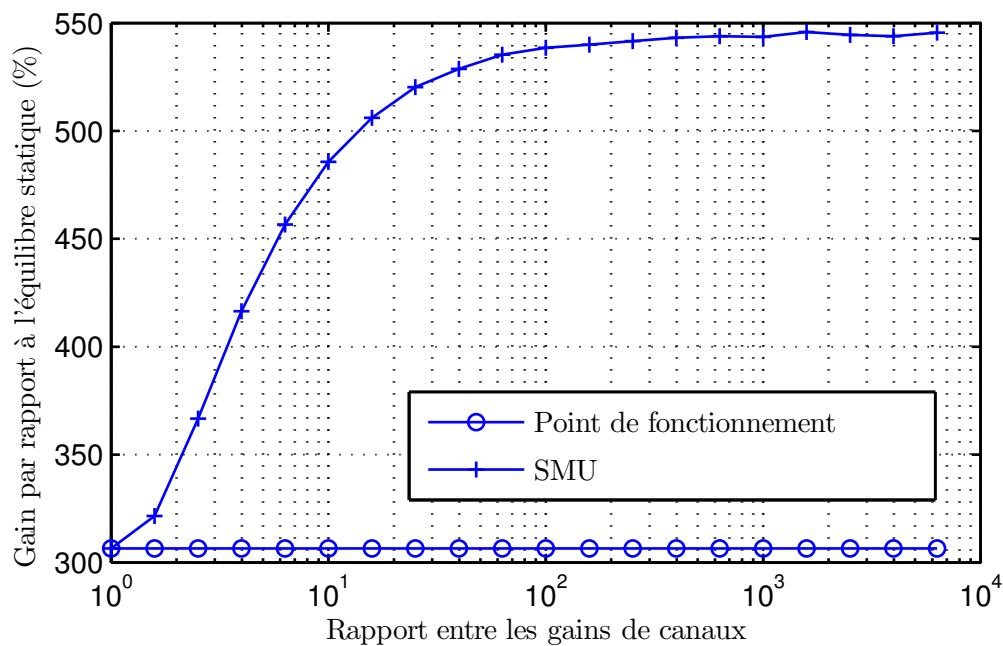


FIGURE 2.6 – Comparaison de l'efficacité énergétique du plan SMU et du point de fonctionnement par rapport à l'équilibre de Nash du jeu statique en fonction du rapport entre les gains de canaux atteignables dans un système à 10 émetteurs.

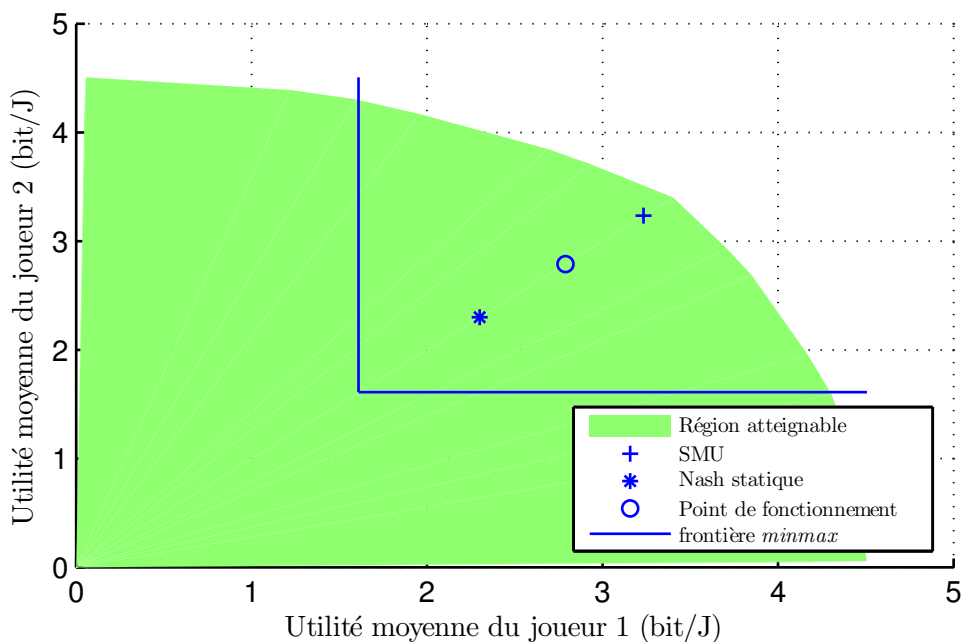


FIGURE 2.7 – Région atteignable et utilités moyennes de diverses stratégies pour un jeu à 2 joueurs.

Pour cette simulation, on considère un nombre fini de gains de canal. La loi d'évolution des gains des canaux suit la propriété de Markov, c'est-à-dire qu'il existe une matrice de probabilité de transition entre l'état des canaux à l'instant n et l'état des canaux à l'instant $i+1$. Cette matrice ainsi que les états de gains de canal accessibles sont les mêmes pour tous les joueurs. À travers l'étude de ces quatre mécanismes, nous étudions les performances atteignables en fonction du caractère centralisé ou décentralisé du mécanisme ainsi que de la quantité d'information disponible sur le système. Ces mécanismes sont les suivants :

- Une version centralisée de SMU, dans laquelle le récepteur choisit qui émet à chaque tour et impose la puissance d'émission en connaissant les gains des canaux à l'instant n . Dans le modèle considéré, les émetteurs appliquent à l'instant $n+1$ la puissance d'émission décidée à l'instant n . Ce retard se justifie par un temps de transmission entre le récepteur et les émetteurs.
- SMU, pour lequel le récepteur décide uniquement l'ensemble des émetteurs conseillé à chaque tour du jeu. Chaque émetteur connaissant le gain de son canal et le nombre d'émetteurs qui vont transmettre avec lui, il fixe lui-même sa puissance d'émission. De la même manière que précédemment, on prend en compte le retard de transmission entre le récepteur et les émetteurs. L'ensemble des joueurs qui émettent à l'instant $n+1$ est donc décidé par le récepteur à l'instant n .
- La stratégie reposant sur le point de fonctionnement développée dans [Le Treust 2010]. L'approche est encore plus décentralisée puisque tous les émetteurs fixent leur puissance à chaque tour en connaissant le gain de leur canal et le nombre de joueurs sans recommandation de la part du récepteur.
- Un équilibre de Nash "myope". Dans ce cas, les émetteurs n'ont aucune information sur le système mis à part l'espérance du gain de leur canal et le nombre de joueurs. Ils se contentent donc de jouer l'équilibre de Nash statique.

Il est intéressant de noter que SMU offre de meilleures performances que les trois autres mécanismes. En ce qui concerne l'approche centralisée, le fait que la puissance d'émission soit connue des émetteurs avec un temps de retard par rapport à l'état des gains des canaux est un véritable handicap qui n'est compensé que pour un nombre suffisant d'émetteurs.

La figure 2.9 offre une représentation graphique de la distribution des différentes configurations de recommandations qu'un émetteur peut rencontrer au cours des différentes étapes d'un jeu stochastique (dans un cas où les distributions des canaux des émetteurs sont identiques). Le signal 1 correspond à la recommandation d'émettre alors que 2 correspond à la recommandation de ne pas émettre. Pour chacune des recommandations, nous indiquons également le nombre total d'émetteurs qui ont la recommandation d'émettre. La simulation est faite pour 5 émetteurs.

La figure 2.10 fait référence au théorème 4. Elle représente la valeur maximale que le facteur d'escompte λ peut prendre pour que le plan SMU soit un équilibre du jeu stochastique. Comme l'écart entre l'efficacité énergétique du plan SMU et l'efficacité énergétique de l'équilibre de Nash de la formulation statique croît avec le nombre d'émetteurs dans le système, la valeur maximale du facteur d'escompte

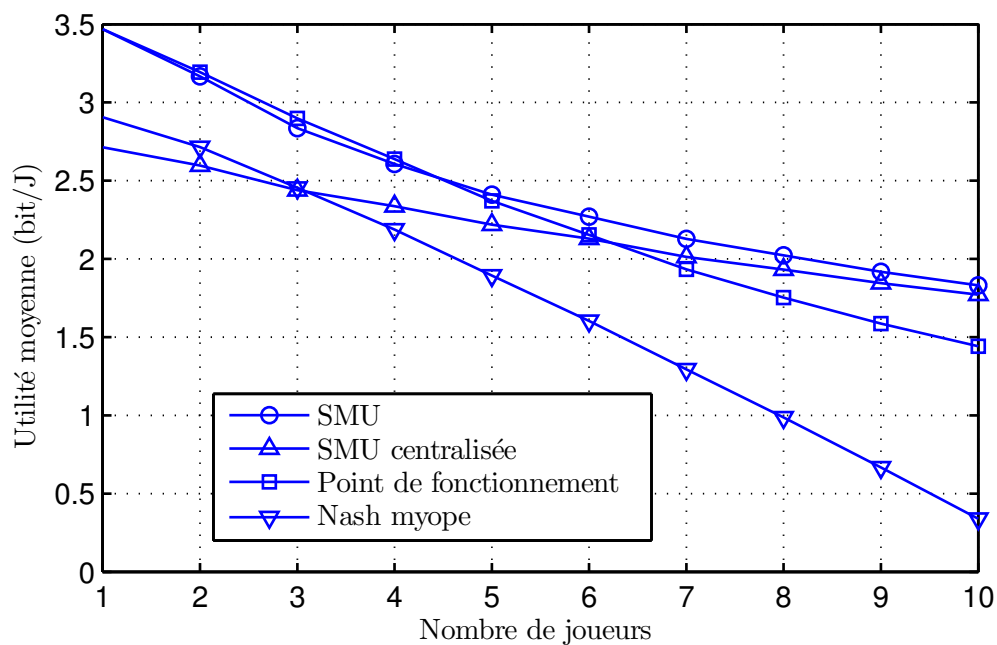


FIGURE 2.8 – Utilités moyennes de quatre mécanismes de contrôle de puissance en fonction du nombre d'émetteurs.

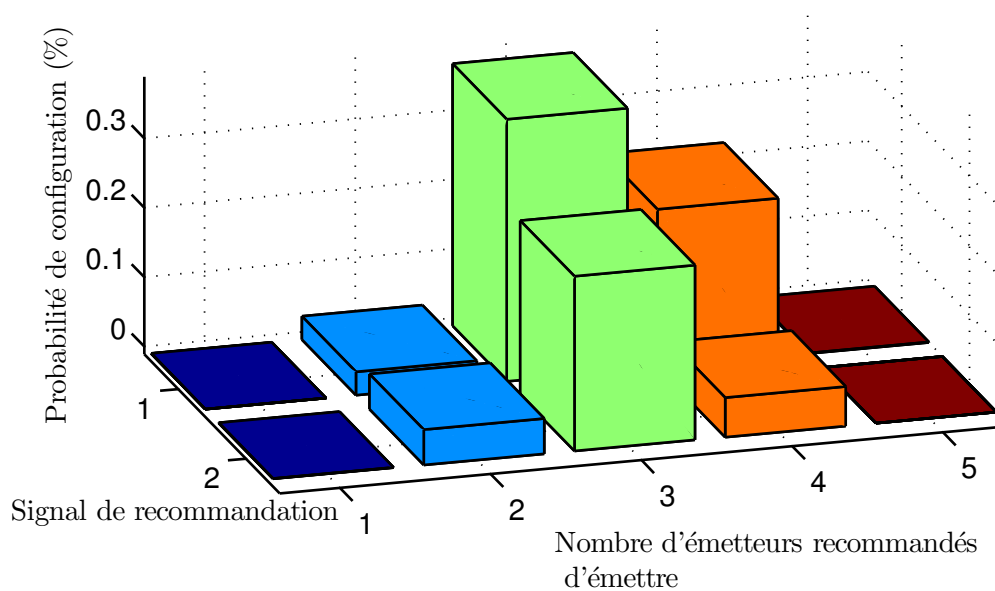


FIGURE 2.9 – Distribution des différentes configurations de recommandation avec le plan SMU dans un système à 5 émetteurs.

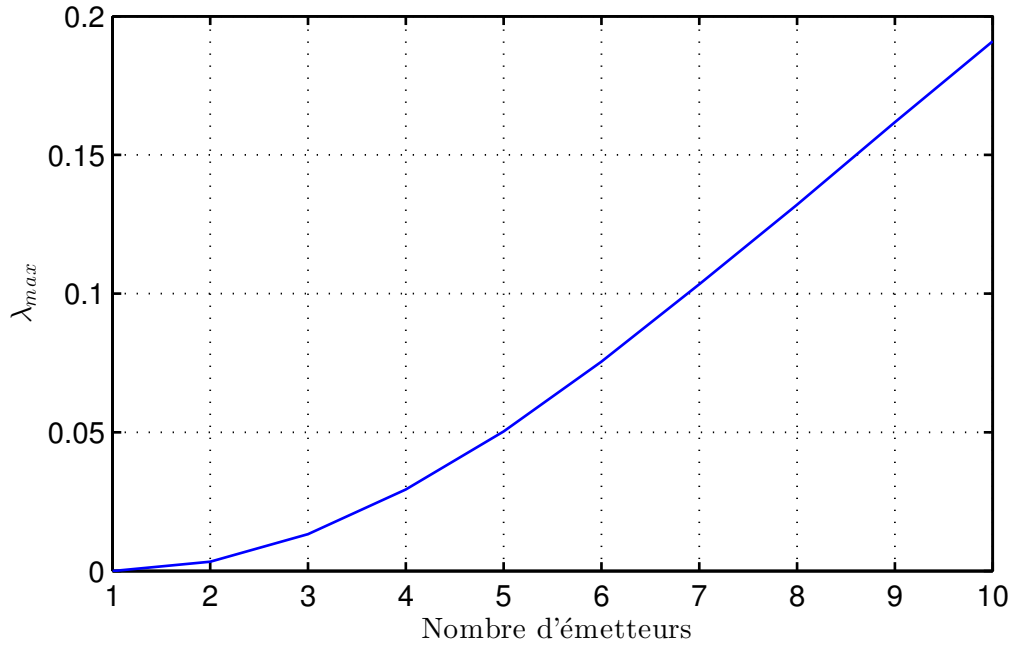


FIGURE 2.10 – Valeur maximale du coefficient d'escompte λ pour que le plan SMU soit un équilibre du jeu stochastique.

augmente également.

2.5 Conclusion

Dans cette partie, nous avons étendu le problème de contrôle de puissance efficace énergétiquement initialement introduit dans [Goodman 2000]. Notre première contribution est de proposer un jeu statique pour étudier un problème d'allocation de puissance efficace énergétiquement (à l'échelle d'une cellule) lorsque les terminaux mobiles sont libres de reporter n'importe quelle valeur de gain de canal. Plus précisément, nous avons étudié le conflit d'intérêt entre des terminaux mobiles choisissant la valeur de leur report de gain de canal pour maximiser leur RSB et une station de base qui alloue sa puissance d'émission de façon à maximiser l'efficacité énergétique de la cellule. La formulation du problème par un jeu statique nous permet de prouver l'existence d'un unique équilibre pour le choix des reports des gains de canaux. En comparaison avec le cas pour lequel les terminaux mobiles reportent leurs vrais gains de canaux, les résultats numériques issus de notre modèle indiquent que ces reports stratégiques permettent d'augmenter le RSB des terminaux, au détriment de l'efficacité énergétique de la cellule. Cependant, cette dégradation de l'efficacité énergétique est moins importante pour des cellules avec une puissance d'émission faible. En conséquence, les cellules de petite dimension sont plus robustes à des reports stratégiques de gains de canaux que des cellules avec une forte puissance d'émission.

Notre seconde contribution est de proposer un modèle de jeu stochastique pour modéliser le problème de contrôle de puissance initial. Ce jeu stochastique présente l'avantage de tenir compte des interactions stratégiques entre les terminaux mobiles au cours du temps (de même que le jeu répété introduit dans [Le Treust 2010]), ainsi que de la variation de l'état des canaux dans le temps. Cette approche nous permet notamment de caractériser la région des utilités des stratégies d'équilibre. Il apparaît également qu'en raison des interactions à long terme entre les émetteurs, ces derniers peuvent avoir intérêt à ne pas émettre certains paquets si leur condition de canal est trop dégradée. Cela nous mène à établir une stratégie de contrôle de puissance d'équilibre fondée sur le partage temporel qui se montre performante en termes d'efficacité énergétique. Les perspectives de ce travail sont d'intégrer dans le contrôle de puissance plusieurs aspects visant à mieux prendre en compte les caractéristiques des flux d'information dans des réseaux réels : la possibilité de tolérer un retard sur l'émission d'un paquet ; la possibilité que le flux de paquets soit sporadique ; le fait que la taille de stockage des paquets à l'émetteur soit finie.

Jeux à champ moyen et contrôle de puissance efficace énergétiquement

Sommaire

3.1	Formulation du jeu différentiel stochastique	43
3.1.1	Définition du jeu différentiel stochastique	43
3.1.2	Équilibre de Nash du jeu différentiel stochastique	45
3.1.3	Cas du jeu à un joueur	47
3.2	Analyse du jeu à champ moyen	47
3.2.1	Convergence vers le jeu à champ moyen	47
3.2.2	Solution du problème de réponse au champ moyen	49
3.2.3	Unicité de la solution	51
3.2.4	Résultats numériques pour le jeu à champ moyen	51
3.3	Conclusion	54

Dans ce chapitre, nous modifions le jeu stochastique introduit en partie 2.4 du chapitre précédent pour obtenir une modélisation plus fine du problème de contrôle de puissance efficace énergétiquement. Il résulte de ce nouveau modèle un jeu différentiel stochastique [Bressan 2011] dont l'étude est complexe. Face à la complexité de ce nouveau modèle, nous faisons appel à la théorie des jeux à champ moyen [Lasry 2007] pour étudier l'existence et l'unicité d'un équilibre de Nash.

Commençons par rappeler que le jeu stochastique introduit dans la partie 2.4 présente l'avantage de modéliser le problème de contrôle de puissance efficace énergétiquement en tenant compte de la variation des canaux au cours du temps, ainsi que des interactions entre les joueurs sur le long terme. Bien qu'il soit possible de déterminer heuristiquement des stratégies d'équilibres de ce jeu stochastique, (comme nous l'avons fait en partie 2.4 du chapitre précédent) ces stratégies ne sont pas forcément optimales. Motivés par le besoin d'obtenir un modèle plus facilement manipulable et la volonté d'introduire dans le modèle des considérations énergétiques à long terme, nous apportons les modifications suivantes.

- La première modification est d'introduire une contrainte d'énergie à long terme pour les terminaux mobiles. Cette contrainte d'énergie permet de tenir compte des relations entre l'état de charge des batteries des terminaux mobiles et les stratégies de contrôle de puissance : un terminal dont la batterie est vide

ne peut plus émettre et un terminal qui émet vide progressivement sa batterie. Cette contrainte d'énergie est complémentaire avec les considérations énergétiques de court terme induites par la maximisation de l'efficacité énergétique. À notre connaissance, cette contrainte à long terme est une nouveauté dans la littérature sur le contrôle de puissance efficace énergétiquement. En effet, dans les travaux de [Goodman 2000] et les travaux qui suivirent [Meshkati 2006, Bonneau 2008, Lasaulce 2009, Buzzi 2011], les terminaux mobiles émettent tout le temps, ce qui revient à ne considérer aucune contrainte sur l'énergie disponible.

- La seconde modification apportée au modèle du jeu stochastique concerne l'évolution des canaux au cours du temps. Alors que les canaux du jeu stochastique varient dans un espace d'état discret selon une chaîne de Markov, nous considérons à présent que les canaux évoluent de manière continue selon une équation différentielle stochastique. Comme justifié dans [Agarwal 2012], ce modèle est un cas limite d'un modèle de canal à évanouissement par bloc corrélé dans lequel le canal évolue en temps discret selon un processus de Gauss-Markov. Il présente l'avantage de simplifier l'étude analytique de l'évolution du système. D'autres travaux utilisant cette loi continue d'évolution de canal sont [Charalambous 2005, Feng 2007].

Ces deux modifications introduisent donc des lois d'évolution continues pour les quantités d'énergie restantes dans les batteries des terminaux mobiles, ainsi que pour les états des canaux. En considérant que l'état de la batterie et l'état du canal d'un terminal constituent son état individuel, le nouveau modèle correspond donc à un jeu différentiel stochastique. En comparaison des travaux de [Huang 2003], ce jeu différentiel stochastique de contrôle de puissance se différencie notamment par le fait qu'il n'est pas linéaire quadratique. Pour ce jeu, nous pouvons donner une condition d'existence d'un équilibre de Nash. Mais la complexité de la caractérisation des performances du réseau distribué augmente drastiquement avec le nombre de terminaux dans le réseau et le même constat est valable pour déterminer les stratégies de contrôle de puissance d'équilibre. Pour résoudre ce problème de complexité, nous faisons appel aux jeux à champ moyen.

Les jeux à champ moyen [Lasry 2007] permettent d'étudier la limite d'un jeu différentiel stochastique lorsque le nombre de joueurs tend vers l'infini. Le principe est que dans un jeu avec un très grand nombre de joueurs, l'influence individuelle d'un joueur sur le cours du jeu tend à être nulle. Cela permet de considérer qu'un joueur donné ne joue plus en fonction de l'ensemble discret d'autres joueurs, mais en fonction de l'agrégation continue des autres joueurs dans un champ moyen. L'étude de la stratégie optimale d'un joueur donné ne nécessite donc plus de connaître tous les états des autres joueurs, mais uniquement l'état du champ moyen. Cette formulation simplifie grandement l'étude du jeu différentiel stochastique. Il semble que le travail le plus pertinent au sujet des jeux à champ moyen pour le contrôle de puissance est [Tembine 2010]. Par rapport à cette référence, le travail présenté dans ce chapitre se distingue par une fonction d'utilité différente (aucune hypothèse de contrôle linéaire quadratique n'est faite ici), une loi d'évolution des canaux plus

réaliste et le fait que le niveau de batterie de chaque terminal fait partie de l'état du terminal.

Ce chapitre est organisé de la façon suivante. Dans la partie 3.1, nous introduisons le jeu différentiel stochastique qui modélise le problème de contrôle de puissance efficace énergétiquement sous des contraintes d'énergie à long terme et nous donnons une condition pour l'existence d'un équilibre de Nash. Dans la partie 3.2, l'étude de ce jeu sous l'hypothèse que le nombre de joueurs tend vers l'infini est permise par la formulation d'un jeu à champ moyen. Nous donnons une condition pour l'unicité de l'équilibre de Nash et nous illustrons, par des résultats numériques, des stratégies de contrôle de puissance d'équilibre pour ce jeu.

3.1 Formulation du jeu différentiel stochastique

3.1.1 Définition du jeu différentiel stochastique

Dans cette partie, nous présentons un jeu différentiel stochastique [Bressan 2011] qui est construit à partir du jeu statique $\mathcal{G}$ défini en partie 2.1. Nous supposons que le temps est continu, c'est-à-dire $t \in \mathbb{R}$. Cette hypothèse a été discutée dans d'autres travaux sur le contrôle de puissance tels que [Foschini 1993, Olama 2006, Tembine 2010]. Cette hypothèse est à la fois pertinente pour les scénarios dans lesquels les canaux sont soumis à des atténuations rapides et en tant que cas limite des scénarios où les canaux sont soumis à une atténuation lente. L'horizon de temps du jeu est fini, nous le définissons comme l'intervalle allant de T à T' . Cependant, la méthodologie développée ici peut être étendue au cas de l'horizon infini (avec des utilités moyennes au cours du temps) car les processus sous-jacents sont ergodiques (le modèle de canal proposé est ergodique et la dynamique de consommation d'énergie l'est également).

Définition 13 (Modèle de jeu différentiel stochastique) *Le jeu différentiel stochastique de contrôle de puissance est défini par le quadruplet*

$\widehat{\mathcal{G}} = (\mathcal{K}, \{\mathcal{X}_k\}_{k \in \mathcal{K}}, \{\mathcal{S}_k\}_{k \in \mathcal{K}}, \{U_k\}_{k \in \mathcal{K}})$ *pour lequel :*

- $\mathcal{K}$ est l'ensemble des joueurs. À nouveau, c'est l'ensemble des émetteurs du système ;
- $\mathcal{X}_k \subset \mathbb{R}^3$ est l'espace d'état du joueur k . L'état du joueur k à l'instant t est défini par $\underline{X}_k(t) = [E_k(t), \underline{h}_k(t)]^T$ et suit la loi d'évolution suivante :

$$d\underline{X}_k(t) = \begin{bmatrix} -p_k(t) \\ \frac{1}{2}(\underline{\mu} - \underline{h}_k(t)) \end{bmatrix} dt + \begin{bmatrix} 0 \\ \nu \end{bmatrix} d\underline{\mathbb{W}}_k(t); \quad (3.1)$$

- $E_k(t)$ et $\underline{h}_k(t)$ sont respectivement l'énergie disponible et le canal du joueur k à l'instant t ;
- $\forall k \in \mathcal{K}$, $\underline{\mathbb{W}}_k(t)$ sont des processus de Wiener mutuellement indépendants de dimension 2 ;
- $\underline{\mu} \geq 0$ et $0 \leq \nu < +\infty$ sont des constantes dont les interprétations physiques sont données dans la proposition 3 ;

- $\mathcal{S}_k$ est l'ensemble des stratégies de contrôle pour le joueur k [Basar 1999]. Une stratégie de contrôle est notée $\hat{\tau}_k$ avec

$$\begin{aligned} \hat{\tau}_k : [T, T'] \times \underline{\mathcal{X}} &\rightarrow \mathcal{A}_k \\ (t, \underline{X}(t)) &\mapsto p_k, \end{aligned} \quad (3.2)$$

- où $\underline{\mathcal{X}} = \prod_{k \in \mathcal{K}} \mathcal{X}_k$ et $\underline{X}(t) = (\underline{X}_1(t), \dots, \underline{X}_K(t))^T$ est le profil d'états ;
- La fonction d'utilité moyenne U_k est définie par :

$$U_k(\hat{\tau}, \underline{X}(T)) = \mathbb{E} \left[\int_T^{T'} u_k(\hat{\tau}(t, \underline{X}(t)), \underline{X}(t)) dt + q(\underline{X}(T')) \right], \quad (3.3)$$

où $\hat{\tau} = (\hat{\tau}_1, \dots, \hat{\tau}_K)$ est le profil de stratégies de contrôle et $q(\underline{X}(T'))$ est l'utilité de l'état final (le modèle permet donc que certains états finaux soient préférables à d'autres). Nous rappelons enfin que u_k est l'efficacité énergétique instantanée (voir partie 2.1), donnée par :

$$u_k(\underline{p}(t), \underline{X}(t)) = \frac{Rf(\gamma_k(\underline{p}(t)))}{p_k(t)} \text{ bit/J}. \quad (3.4)$$

Les raisons pour choisir les dynamiques proposées pour l'état $\mathcal{X}_k$ sont les suivantes. Le terme $dE_k(t) = -p_k(t)dt$ signifie que la variation d'énergie pendant dt est proportionnelle à la puissance consommée par l'émission. Le modèle proposé permet donc de prendre en compte qu'une émission a un coût énergétique. Cela s'applique parfaitement pour les terminaux mobiles disposant uniquement d'une réserve d'énergie finie, tels que téléphones portables, les ordinateurs portables, les petites stations de bases qui doivent être autonomes énergétiquement, etc. De tels terminaux ont des batteries qui contiennent une quantité d'énergie finie et nécessitent donc d'être rechargées lorsqu'elles sont vides. Notre modèle est valide pour une période comprise entre deux recharges et sur un tel horizon, l'énergie disponible est une fonction décroissante avec le temps. Il faut bien comprendre que la prise en compte des énergies disponibles dans les batteries des terminaux change radicalement les stratégies de contrôle de puissance par rapport aux modèles étudiés dans le chapitre 2. Par exemple, si la batterie d'un terminal est vide, la puissance d'émission optimale doit être nulle et ne peut donc pas suivre les niveaux de puissances recommandés précédemment.

En ce qui concerne l'horizon de temps de notre modèle, T et T' peuvent être choisis arbitrairement tant que $T' \geq T$. Bien sûr, les paramètres de la loi d'évolution doivent être stationnaires sur cet intervalle, c'est-à-dire que $\underline{\mu}$ et $\underline{\nu}$ doivent être fixés. La taille de l'horizon, mesuré en seconde, dépend de l'application visée. Dans le cas où l'évolution des canaux modélise les effets des pertes dues à la propagation, cette évolution est lente et l'horizon de temps est de l'ordre de la minute. Dans le cas où la loi d'évolution des canaux modélise des atténuations rapides, l'horizon est de l'ordre de la seconde. Typiquement, dans un réseau cellulaire, le contrôle de puissance est effectué par blocs dont une durée typique est 1 ms.

Tout en étant simple, la dynamique du canal $\underline{h}_k(t)$ fait apparaître plusieurs phénomènes caractéristiques des communications sans fil. Avant de commenter ces phénomènes, nous établissons les propriétés du processus aléatoire $\underline{h}_k(t)$.

Proposition 3 (Propriétés de la dynamique du canal) *Soit $\underline{h}_k(t) = (a_k(t), b_k(t))$ un processus aléatoire suivant la dynamique définie dans le jeu différentiel stochastique $\widehat{\mathcal{G}}$, alors nous avons :*

$$\begin{aligned} \lim_{t \rightarrow +\infty} \mathbb{E}[\underline{h}_k(t)] &= \underline{\mu}, \\ \lim_{t \rightarrow +\infty} \mathbb{E}[|\underline{h}_k(t)|^2] - \mathbb{E}[|\underline{h}_k(t)|]^2 &= 2\nu^2. \end{aligned} \quad (3.5)$$

Les fonctions de densité de probabilité stationnaires

$m_a : \mathbb{R} \rightarrow \mathbb{P}(\mathbb{R})$, $m_b : \mathbb{R} \rightarrow \mathbb{P}(\mathbb{R})$ des deux composantes a_k , b_k de $\underline{h}_k$ sont :

$$\begin{cases} m_a(a_k) = \frac{1}{\nu\sqrt{2\pi}} e^{-\frac{(a_k - \mu_a)^2}{2\nu^2}}, \\ m_b(b_k) = \frac{1}{\nu\sqrt{2\pi}} e^{-\frac{(b_k - \mu_b)^2}{2\nu^2}}. \end{cases} \quad (3.6)$$

La preuve de ce résultat est donnée dans [Mériaux 2013]. Cela montre que la dynamique proposée permet notamment de modéliser des canaux de Rice, c'est-à-dire des canaux dont le gain moyen est non-nul. Cela est possible en fixant $|\underline{\mu}| \geq 0$, qui représente le paramètre de Rice. Les effets des évanouissements lents ou rapides peuvent également être pris en compte en choisissant une valeur de variance $2\nu^2$ respectivement faible ou élevée. Il est aussi intéressant de noter que cette dynamique de canal peut aussi être perçue comme un cas limite du modèle de Gauss-Markov en temps discret, modèle qui permet de tenir compte de la corrélation temporelle des gains de canaux [Agarwal 2012]. Pour finir notre discussion sur le choix de cette dynamique, nous verrons en partie 3.2 qu'elle possède également une propriété importante pour la dynamique du jeu à champ moyen qui sera étudié.

3.1.2 Équilibre de Nash du jeu différentiel stochastique

À présent, nous pouvons définir un équilibre de Nash du jeu différentiel stochastique $\widehat{\mathcal{G}}$ et énoncer un résultat d'existence.

Définition 14 (Équilibre de Nash de $\widehat{\mathcal{G}}$) *Un profil de stratégie de contrôle $\widehat{\tau}^* = (\widehat{\tau}_1^*, \dots, \widehat{\tau}_K^*)$ est un équilibre de Nash en boucle fermée du jeu différentiel stochastique si et seulement si $\forall k \in \mathcal{K}$, $\widehat{\tau}_k^*$ est une solution du problème de contrôle :*

$$\sup_{\widehat{\tau}_k} \mathbb{E} \left[\int_T^{T'} u_k \left(\widehat{\tau}_k(t, \underline{X}(t)), \widehat{\tau}_{-k}^*(t, \underline{X}(t)), \underline{X}(t) \right) dt + q(\underline{X}(T')) \right], \quad (3.7)$$

sous contrainte que

$$d\underline{X}(t) = \begin{bmatrix} -[\widehat{\tau}_k(t, \underline{X}(t)), \widehat{\tau}_{-k}^*(t, \underline{X}(t))]^T \\ \frac{1}{2}(\underline{\mathbb{1}} \otimes \underline{\mu} - \underline{h}(t)) \end{bmatrix} dt + \begin{bmatrix} 0 \\ \nu \end{bmatrix} d\underline{W}(t), \quad (3.8)$$

où $\underline{1} = (1, \dots, 1) \in \mathbb{R}^K$, $\underline{h}(t) = (\underline{h}_1(t), \dots, \underline{h}_K(t))^T$, $\otimes$ désigne le produit de Kronecker, $\underline{\mathbb{W}}(t) = (\underline{\mathbb{W}}_1(t), \dots, \underline{\mathbb{W}}_K(t))^T$.

En raison de la prise en compte des niveaux d'énergies des batteries des différents émetteurs du système, l'étude d'un équilibre de Nash est différente des jeux étudiés au chapitre 2. Cependant, nous pouvons énoncer le théorème suivant concernant l'existence d'un équilibre de Nash.

Théorème 5 (Existence d'un équilibre de Nash dans $\widehat{\mathcal{G}}$) *Une condition suffisante pour l'existence d'un équilibre de Nash dans $\widehat{\mathcal{G}}$ est que tout couple $(\varphi_0, \gamma_0) \in \mathbb{R}^2$ tel que $f'(\gamma_0)\gamma_0 - f(\gamma_0) = \varphi_0\gamma_0^2$, vérifie $2\varphi_0 - f''(\gamma_0) \neq 0$.*

La preuve détaillée de cette condition d'existence est donnée dans [Mériaux 2013]. Elle vient du fait qu'une condition suffisante pour l'existence d'un équilibre de Nash pour le jeu différentiel stochastique est l'existence d'une solution à l'équation de Hamilton-Jacobi-Bellman-Fleming [Fleming 2006] pour chaque terminal :

$$\begin{aligned} 0 = \sup_{\widehat{\tau}_k} & \left[u_k(\widehat{\tau}(t, \underline{X}(t)), \underline{X}(t)) - \widehat{\tau}_k(t, \underline{X}(t)) \frac{\partial V_k(t, \underline{X}(t))}{\partial E_k} \right] \\ & + \frac{1}{2} \langle \underline{\mu} - \underline{h}_k(t), \nabla_{\underline{h}_k} V_k(t, \underline{X}(t)) \rangle_{\mathbb{R}^2} + \frac{\partial V_k(t, \underline{X}(t))}{\partial t} + \frac{\nu^2}{2} \Delta_{\underline{h}_k} V_k(t, \underline{X}(t)), \end{aligned} \quad (3.9)$$

où V_k est la fonction de Bellman définie par :

$$V_k(T, \underline{X}(T)) = \sup_{\widehat{\tau}_k} \mathbb{E} \left[\int_T^{T'} u_k(\widehat{\tau}(t, \underline{X}(t)), \underline{X}(t)) dt + q(\underline{X}(T')) \right]. \quad (3.10)$$

Le théorème 5 peut s'appliquer à toutes les fonctions d'efficacité utilisées dans la littérature. En particulier, pour la fonction associée à la probabilité de coupure [Belmega 2011] $f(x) = e^{-\frac{a}{x}}$, avec $a \geq 0$, nous avons :

$$\begin{cases} e^{-\frac{a}{\gamma_0}} \left(\frac{a}{\gamma_0} - 1 \right) = \varphi_0 \gamma_0^2, \\ 2\varphi_0 - \left(\frac{a^2}{\gamma_0^4} - \frac{2a}{\gamma_0^3} \right) e^{-\frac{a}{\gamma_0}} \neq 0, \end{cases} \quad (3.11)$$

ce qui donne

$$2\varphi_0 - \left(\frac{a}{\gamma_0} - 2 \right) \frac{a\varphi_0}{a - \gamma_0} \neq 0. \quad (3.12)$$

Bien que nous n'ayons pas de résultat formel permettant de justifier que cette condition soit tout le temps vérifiée. De nombreuses simulations semblent indiquer que cela est le cas. Cette condition n'est donc pas restrictive.

Alors que l'unicité de l'équilibre de Nash est une propriété attractive du jeu statique (voir partie 2.1), cette propriété n'est pas facile à vérifier pour le jeu différentiel stochastique $\widehat{\mathcal{G}}$. Au contraire, ce type de jeux possède généralement un grand nombre d'équilibres. La détermination explicite d'un équilibre de Nash n'est également pas une tâche aisée puisque qu'en conséquence de (3.9), la détermination d'un équilibre de Nash nécessite de résoudre un système de K équations d'Hamilton-Jacobi-Bellman-Fleming [Fleming 2006], couplées par les états $\underline{X}(t)$. Il existe cependant

deux cas-limites de $\widehat{\mathcal{G}}$ pour lesquels les problèmes d'existence et d'unicité deviennent plus simples. Le premier cas particulier est lorsqu'il y a uniquement un joueur. Le second cas particulier est quand le nombre de joueur est grand, ce qui conduit à l'étude d'un jeu à champ moyen. Le premier cas est détaillé dans la partie suivante et le second est traité dans la partie 3.2.

3.1.3 Cas du jeu à un joueur

Une des raisons pour lesquelles nous analysons le cas du jeu à un joueur est de séparer les effets des contraintes d'énergie à long terme et les effets des interactions entre les joueurs puisque ces deux effets peuvent indépendamment inciter les terminaux à ne pas émettre. En effet, dans un système avec plusieurs terminaux, ces derniers peuvent avoir intérêt à ne pas émettre en permanence pour limiter le niveau d'interférence (ce qui fait des stratégies de partage de temps d'émission des équilibres naturels, comme observé dans [Mériaux 2011a]). Par contre, dans le jeu à un joueur, il n'y a pas d'interactions possibles et nous montrons, à travers des équations simples, que la contrainte d'énergie à long terme incite le joueur à ne pas tout le temps émettre, en fonction de l'état de son canal.

Dans la partie précédente, nous avons montré que déterminer une stratégie optimale de contrôle de puissance revient à résoudre l'équation suivante en $\gamma_1(t)$:

$$f'(\gamma_1(t))\gamma_1(t) - f(\gamma_1(t)) = \frac{\gamma_1(t)^2}{R} \frac{\partial V_1(t, \underline{X}_1(t))}{\partial E_1} \frac{\sigma^4}{|\underline{h}_1(t)|^4}. \quad (3.13)$$

En notant $\gamma_1^*(t)$ la plus grande solution de l'équation précédente, une stratégie optimale de contrôle de puissance s'écrit :

$$\widehat{\tau}_1^*(t, \underline{X}_1(t)) = \frac{\sigma^2}{|\underline{h}_1(t)|^2} \gamma_1^*(t), \quad (3.14)$$

avec $\gamma_1^*(t) \geq 0$. Il est intéressant de remarquer que $\gamma_1^*(t) = 0$ peut être vrai pour certains intervalles de temps. En particulier, cela dépend du gain de canal $|\underline{h}_1(t)|^2$. Si ce dernier est trop faible, émettre ne serait pas efficace énergétiquement, ce qui conduit à une puissance d'émission nulle. En d'autres termes, si le canal d'un terminal n'est pas de qualité suffisamment bonne, le terminal n'a pas intérêt à émettre. Cette dernière constatation peut avoir un intérêt pratique dans le contexte de la radio cognitive pour que d'autres émetteurs puissent capter les ressources radio alors laissées libres. Cette idée est plus largement développée dans [Mériaux 2011a].

3.2 Analyse du jeu à champ moyen

3.2.1 Convergence vers le jeu à champ moyen

Dans la partie précédente, nous avons mentionné que le jeu différentiel stochastique devient de plus en plus difficile à analyser quand le nombre de joueurs augmente. Néanmoins, notre système présente une structure particulière qui peut être

exploitée pour simplifier le problème lorsque K est grand. En effet, du point de vue d'un joueur donné, ce qui va influencer la qualité de sa transmission est une somme pondérée des actions des autres joueurs et non les actions des autres joueurs individuellement. La quantité numérique qui affecte l'utilité du joueur k est la suivante :

$$I_k(t) = \frac{1}{K} \sum_{j \in \mathcal{K}, j \neq k} p_j(t) |\underline{h}_j(t)|^2. \quad (3.15)$$

Dans un réseau sans fil, cette quantité est simplement l'interférence qui affecte un récepteur lorsqu'il cherche à décoder le signal de l'émetteur k sans procédé d'annulation d'interférence. Ici, cette quantité est normalisée. Plusieurs applications pratiques permettent de justifier cette normalisation [Tulino 2004]. Par exemple, elle est pleinement justifiée dans les systèmes CDMA à étalement aléatoire [Meshkati 2006]. Prouver que le jeu différentiel stochastique converge vers un jeu à champ moyen revient à prouver que cette interférence converge lorsque le nombre de joueurs tend vers l'infini. Pour prouver cette convergence, la traditionnelle loi des grands nombres ne peut pas être appliquée puisque les processus aléatoires $p_j(t) |\underline{h}_j(t)|^2$ sont dépendant les uns des autres. Cependant, il existe des conditions sous lesquelles la convergence est assurée. Comme détaillé dans [Tembine 2011], l'une d'entre elles est la propriété d'échangeabilité qui est définie ci-dessous.

Définition 15 (Échangeabilité) *Les états $X_1, \dots, X_K$ sont dits échangeables en loi avec la stratégie en boucle fermée α s'ils génèrent une distribution de probabilité conjointe qui est invariante par permutation des indices des joueurs, c'est-à-dire*

$$\forall K, \mathcal{L}(\underline{X}_1, \dots, \underline{X}_K \mid \alpha) = \mathcal{L}(\underline{X}_{\Pi(1)}, \dots, \underline{X}_{\Pi(K)} \mid \alpha), \quad (3.16)$$

pour toute permutation Π définie sur $\{1, \dots, K\}$.

Pour garantir cette propriété dans le jeu étudié, nous faisons les hypothèses suivantes (H) :

- (i) chaque joueur connaît seulement son état individuel instantané,
- (ii) chaque joueur utilise un contrôle homogène et admissible :

$$\begin{aligned} \hat{\tau}_k(t, \underline{X}_k(t)) &= \alpha_t(\underline{X}_k(t)), \\ \text{et } \mathbb{E} \left[\int_T^{T'} \alpha_t(\underline{X}_k(t))^2 dt \right] &< +\infty, \end{aligned} \quad (3.17)$$

- (iii) les états initiaux $\underline{X}_k(T)$ sont identiques pour tous les joueurs.

En conséquence de la propriété d'échangeabilité dans $\widehat{\mathcal{G}}$, le jeu comprend maintenant des joueurs avec des conditions initiales et des lois d'évolution symétriques. À partir de ce constat, nous notons donc $\underline{s} = (E, \underline{h})^T$ la variable d'état d'un joueur (qui ne nécessite plus d'être indexé) et $\underline{s}(t) = (E(t), \underline{h}(t))^T$ sa réalisation à l'instant t .

Proposition 4 (Convergence vers le jeu à champ moyen) *Si les hypothèses (H) sont vérifiées, alors le jeu différentiel stochastique $\widehat{G}$ converge vers un jeu à champ moyen $\widehat{G} = (\mathcal{X}, \mathcal{S}, \widehat{U})$ où :*

- $\mathcal{X} \subset \mathbb{R}^3$ est l'espace d'état de chaque joueur. Les dynamiques d'état d'un joueur sont données par le système d'équations différentielles stochastiques suivantes :

$$\begin{cases} dE(t) = -\alpha_t(\underline{s}(t))dt, \\ d\underline{h}(t) = \frac{1}{2}(\underline{\mu} - \underline{h}(t))dt + \nu d\underline{W}(t). \end{cases} \quad (3.18)$$

où $\underline{\mu}$ et ν ne dépendent pas du temps.

- $\mathcal{S}$ est l'espace des stratégies α telles qu'à chaque instant $t \in [T, T']$:

$$\begin{aligned} \alpha_t : \mathcal{X} &\rightarrow \mathcal{A} \\ \underline{s} &\mapsto p. \end{aligned} \quad (3.19)$$

- La fonction d'utilité moyenne de chaque joueur est donnée par :

$$\widehat{U}(\alpha, \underline{s}(T)) = \mathbb{E} \left[\int_T^{T'} \hat{u}(\underline{s}(t), m_t, \alpha_t) dt + q(\underline{s}(T')) \right], \quad (3.20)$$

où m_t est le champ moyen, c'est à dire la distribution des états à l'instant t et :

$$\hat{u}(\underline{s}, m_t, \alpha_t) = \frac{Rf(\hat{\gamma}(\underline{s}, m_t, \alpha_t))}{\alpha_t(\underline{s})}, \quad (3.21)$$

$$\hat{\gamma}(\underline{s}, m_t, \alpha_t) = \frac{\alpha_t(\underline{s})|\underline{h}|^2}{\sigma^2 + \widehat{I}(m_t, \alpha_t)}, \quad (3.22)$$

$$\widehat{I}(m_t, \alpha_t) = \int_{\underline{s}} |\underline{h}|^2 \alpha_t(\underline{s}) m_t(d\underline{s}). \quad (3.23)$$

La preuve détaillée de cette proposition est donnée dans [Mériaux 2013].

3.2.2 Solution du problème de réponse au champ moyen

Nous pouvons maintenant définir le concept de solution du système d'équations stochastiques différentielles (3.18).

Définition 16 Soit $T, T' > 0$ tel que $[T, T']$ est l'horizon du jeu. La distribution d'état $m_t(\underline{s})$ est une solution faible de la dynamique d'état (3.18) si m_t est intégrable sur $[T, T']$ et pour toute fonction ψ à support compact infiniment continue et différentiable sur $[T, T'] \times \mathbb{R}^3$, on a

$$\mathbb{E}_{m_T}[\psi_T(\cdot)] + \int_T^{T'} \mathbb{E}_{m_t} \left[\frac{\partial \psi_t(\cdot)}{\partial t} - \langle \nabla_{\underline{s}} \psi_t, D^*(t, \underline{s}) \rangle_{\mathbb{R}^2} + \frac{\nu^2}{2} \Delta_{\underline{s}} \psi_t \right] dt = 0, \quad (3.24)$$

où $\mathbb{E}_{m_t}$ est l'espérance par rapport à m_t et $D^*(t, \underline{s})$ est le vecteur de déviation $(-\alpha_t(\underline{s}), \frac{1}{2}(\underline{\mu} - \underline{h}))^T$.

En appliquant la formule d'Itô [Karatzas 1991], nous pouvons déduire de la définition 16 qu'une distribution d'état m_t , solution faible de l'équation aux dérivées partielles (3.18), vérifie l'équation de *Fokker-Planck-Kolmogorov forward* :

$$\frac{\partial m_t}{\partial t} - \frac{\partial}{\partial E}(m_t \alpha_t) + \text{div}_{\underline{h}}(m_t \frac{1}{2}(\underline{\mu} - \underline{h})) - \frac{\nu^2}{2} \Delta_{\underline{h}} m_t = 0. \quad (3.25)$$

D'autre part, pour n'importe quelle trajectoire m_t , solution faible de (3.18), nous pouvons formuler le problème de contrôle optimal pour lequel chaque joueur sélectionne la meilleure stratégie de réponse à une trajectoire m_t . Le problème de réponse au champ moyen est de déterminer une stratégie de contrôle α qui génère une trajectoire du champ moyen m_t^* et qui est elle-même la meilleure réponse à m_t^* :

$$\widehat{V}_T(\underline{s}(T)) = \sup_{\alpha} \mathbb{E} \left[\int_T^{T'} \widehat{u}(\underline{s}(t), m_t^*, \alpha_t) dt + q(\underline{s}(T')) \right], \quad (3.26)$$

où m_T et $q(\cdot)$ sont supposés connus.

De même que dans la partie 3.1, l'équation (3.26) peut se développer en équation d'Hamilton-Jacobi-Bellman-Fleming. En couplant cette équation à (3.25), nous obtenons donc le système d'équations dont une solution donne un équilibre du jeu à champ moyen :

$$\begin{cases} \frac{\partial \widehat{V}_t}{\partial t} + \tilde{H}(\underline{s}, \frac{\partial \widehat{V}_t}{\partial E}, m_t) + \frac{1}{2} \langle \underline{\mu} - \underline{h}, \nabla_{\underline{h}} \widehat{V}_t \rangle_{\mathbb{R}^2} + \frac{\nu^2}{2} \Delta_{\underline{h}} \widehat{V}_t = 0, \\ \frac{\partial m_t}{\partial t} + \frac{\partial}{\partial E} (m_t \frac{\partial}{\partial u'} \tilde{H}(\underline{s}, \frac{\partial \widehat{V}_t}{\partial E}, m_t)) + \text{div}_{\underline{h}} (m_t \frac{1}{2} (\underline{\mu} - \underline{h})) = \frac{\nu^2}{2} \Delta_{\underline{h}} m_t, \end{cases} \quad (3.27)$$

avec

$$\tilde{H}(\underline{s}, u', m) = \sup_{\alpha_t} \{ \widehat{u}(\underline{s}, m, \alpha_t) - \alpha_t u' \}. \quad (3.28)$$

Comme pour le cas du jeu différentiel stochastique, la première équation est une équation de Hamilton-Jacobi-Bellman-Fleming. Elle est maintenant couplée avec une équation de Fokker-Planck-Kolmogorov. La première est une équation *backward* alors que la seconde est une équation *forward*. La différence fondamentale entre le jeu différentiel stochastique et le jeu à champ moyen est que dans le premier, l'obtention d'un équilibre nécessite la résolution d'un système d'équations (3.9) dont la dimension dépend du nombre de terminaux dans le réseau alors que dans le second, l'équilibre s'obtient par la résolution d'un système de deux équations. Ainsi cette formulation est bien moins complexe et nous verrons dans la partie 3.2.4 qu'elle permet la caractérisation de stratégies de contrôle de puissance d'équilibre. Les seules informations requises au préalable pour résoudre ce système d'équations sont la distribution initiale des états du réseau ainsi que la fonction d'utilité de l'état final. Sous les hypothèses (H), la distribution initiale est implicitement connue puisqu'elle est uniforme mais dans le cas général, elle peut être apprise par le récepteur commun (via un report des émetteurs du réseaux) et diffusée à tous les terminaux dans un unique signal public. Concernant la fonction d'utilité finale, nous pouvons raisonnablement considérer qu'elle est commune et connue par tous les émetteurs. Une fois le système résolu, chaque terminal peut connaître sa puissance optimale d'émission en fonction de l'état de son canal et l'énergie de sa batterie. En d'autres termes, la solution du jeu fournit à chaque terminal une fonction dépendant de l'état de son canal et de l'état de sa batterie dont le résultat est le contrôle de puissance d'équilibre.

3.2.3 Unicité de la solution

Une condition suffisante pour l'unicité de la solution du problème de réponse au champ moyen peut être donnée. Tout d'abord, nous rappelons la définition de la positivité d'un opérateur.

Définition 17 (Positivité d'un opérateur) *L'opérateur $\mathcal{O} : \mathcal{E} \rightarrow \mathcal{E}$ est dit positif (noté par $\mathcal{O} \succ 0$) si*

$$\forall x \neq 0_{\mathcal{E}}, \langle x, \mathcal{O}x \rangle_{\mathcal{E}} > 0,$$

où $0_{\mathcal{E}}$ est l'élément neutre de $\mathcal{E}$.

Théorème 6 (Unicité de la solution du problème de réponse au champ moyen)

Une condition suffisante pour l'unicité de la solution du problème de réponse au champ moyen est que pour tout triplet $(\underline{s}, u', m) \in \mathbb{R}^3 \times \mathbb{R} \times \mathbb{P}(\mathbb{R}^3)$,

$$\begin{aligned} -\frac{1}{m} \frac{\partial}{\partial m} \tilde{H}(\underline{s}, u', m) &\succ 0, \\ \frac{\partial^2}{\partial u'^2} \tilde{H}(\underline{s}, u', m) &> 0, \\ \frac{\partial^2}{\partial m \partial u'} \tilde{H}(\underline{s}, u', m) - \frac{1}{2} \frac{(\frac{\partial^2}{\partial m \partial u'} \tilde{H}(\underline{s}, u', m))^2}{\frac{\partial^2}{\partial u'^2} \tilde{H}(\underline{s}, u', m)} &\succ 0. \end{aligned} \tag{3.29}$$

La preuve de cette proposition est donnée dans [Mériaux 2013].

Dans la partie suivante, nous présentons des résultats numériques illustrant des stratégies de contrôle de puissance issues de l'analyse du jeu à champ moyen.

3.2.4 Résultats numériques pour le jeu à champ moyen

La contribution de cette partie est d'étudier un cas particulier du modèle développé en partie 3.2 pour lequel l'obtention d'une stratégie de contrôle de puissance d'équilibre est simple. Par le biais de résultats numériques, nous étudions l'efficacité énergétique de cette solution en comparaison avec celles des stratégies issues du jeu statique et du jeu répété respectivement définis dans les parties 2.1 et 2.3.

Cas limite des dynamiques de canal lentes

Le scénario que nous étudions dans cette partie est celui des dynamiques de canal lentes par rapport aux dynamiques des niveaux d'énergie. En d'autres termes, pour des canaux à atténuation lente et pour une durée de jeu suffisamment courte, les coefficients des canaux peuvent être considérés constants sur l'intervalle de temps $[T, T']$. La seule composante de l'état des terminaux qui varie au cours du jeu est donc le niveau d'énergie de chaque terminal. Cela revient à considérer que l'énergie

de la batterie définie seule l'état d'un terminal et dans ce cas, le problème de réponse au champ moyen devient :

$$\begin{cases} \frac{\partial \widehat{V}_t}{\partial t} + \sup_{\alpha_t} \{ \widehat{u}(\underline{s}, m_t, \alpha_t) - \alpha_t \frac{\partial \widehat{V}_t}{\partial E} \} = 0, \\ \frac{\partial m_t}{\partial t} - \frac{\partial}{\partial E} (m_t \alpha_t^*) = 0, \end{cases} \quad (3.30)$$

avec $\alpha_t^* = \arg \max_{\alpha_t} \{ \widehat{u}(\underline{s}, m_t, \alpha_t) - \alpha_t \frac{\partial \widehat{V}_t}{\partial E} \}$. La première équation permet d'obtenir $\widehat{V}_t$ et α_t^* en fonction de m_t . La seconde équation donne m_t en fonction de α_t^* . À partir de la distribution initiale m_T et de la fonction d'utilité de l'état final $V_{T'}$, chaque joueur peut obtenir une stratégie de contrôle de puissance α^* qui, pour chaque instant t , donne la puissance d'émission en fonction de l'état du canal et du niveau d'énergie. La stratégie de contrôle de puissance résultant de ce cas est illustrée par des résultats numériques issus de [Mériaux 2012c] dans la partie suivante.

Résultats numériques

Les paramètres choisis pour obtenir des résultats numériques sont les suivants. Lorsque les émetteurs émettent vers le récepteur commun, ils le font avec un débit $R = 1$ Mbps. La puissance du bruit au récepteur est $\sigma^2 = 0,1$ (puissance normalisée). Les gains des canaux sont considérés uniformes et $|h(t)|^2 = 1$ (on notera que n'importe quelle autre distribution statique peut être considérée). La fonction d'efficacité est $f(\gamma) = \exp(-\frac{0,9}{\gamma})$. Les courbes 3.1 à 3.4 sont obtenues pour une durée du jeu de 10 s alors que les courbes 3.5 et 3.6 sont obtenues pour une durée du jeu de 120 s. La distribution initiale d'énergie des terminaux m_T est spécifiée pour chaque figure. La condition finale $\widehat{V}_{T'} = q(s(T'))$ est fixée à 0, pour que la comparaison avec les stratégies issues du jeu statique et du jeu répété escompté soit juste. En effet, il n'y a pas de récompense d'utilité à l'état final dans ces deux types de jeux. Enfin, nous choisissons un niveau maximum d'énergie disponible de $E_{\max} = 20$ J.

Dans la figure 3.1, est représentée la puissance d'émission d'équilibre d'un terminal en fonction de l'énergie disponible dans sa batterie ainsi que la durée du jeu. La figure montre que les terminaux commençant le jeu avec un niveau d'énergie élevé (20 J) commencent par émettre avec un niveau de puissance élevé. Puis ce niveau de puissance diminue avec le temps. Par contre, les terminaux commençant le jeu avec peu d'énergie n'émettent pas avec des valeurs de puissance élevées au début du jeu. Il est intéressant de remarquer que ces terminaux attendent que d'autres terminaux déchargent leurs batterie avant de commencer à émettre. Ceci leur permet de subir moins d'interférence. Ce phénomène est également mis en lumière dans la figure 3.2 qui représente l'évolution de la puissance d'émission pour trois niveaux d'énergie initiale : 1 J, 5 J et 20 J.

La figure 3.3 représente la distribution d'énergie dans les terminaux en fonction du temps lorsque les terminaux suivent la stratégie d'émission d'équilibre. Dans ce cas précis, les terminaux débutent le jeu avec une distribution d'énergie uniforme et nous remarquons que la proportion de terminaux possédant beaucoup d'énergie diminue avec le temps, alors qu'une proportion croissante de terminaux voient leurs

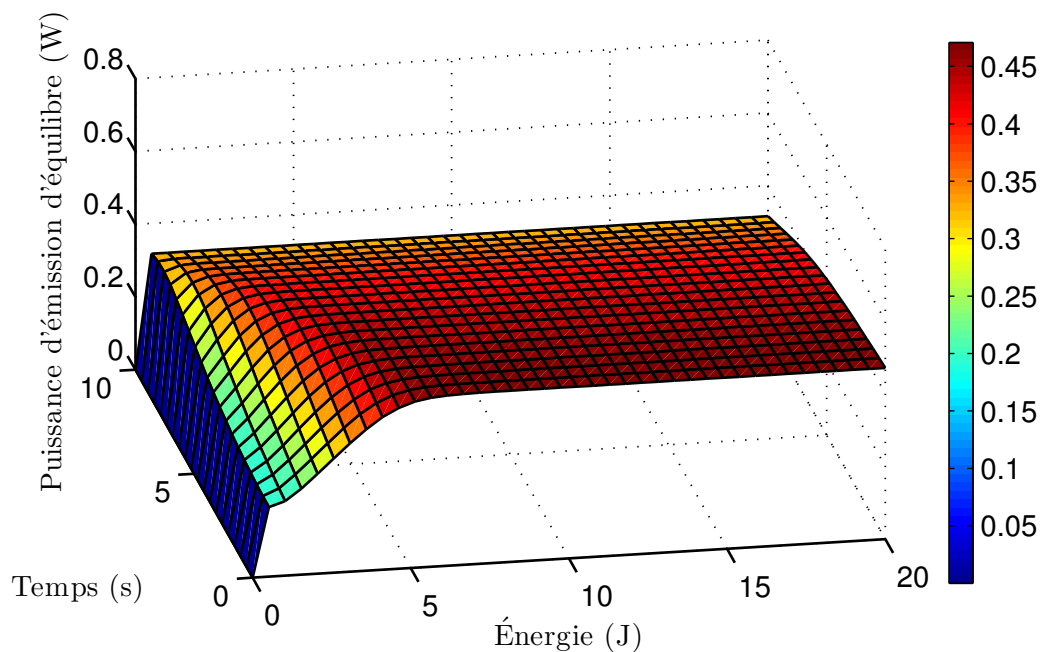


FIGURE 3.1 – Stratégie d'émission d'équilibre du jeu à champ moyen en fonction du temps et de l'énergie disponible.

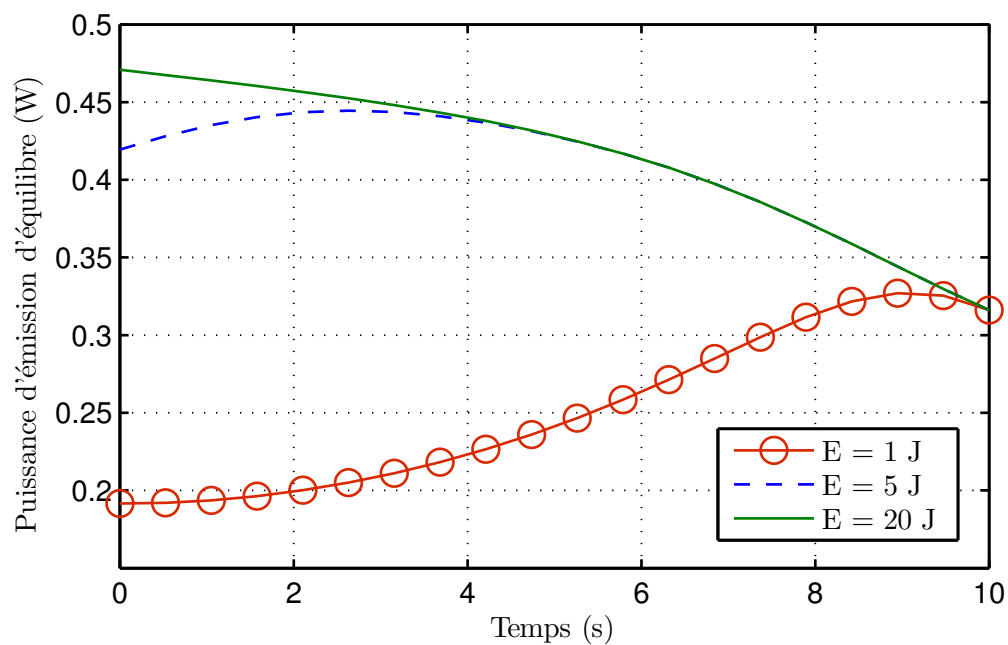


FIGURE 3.2 – Stratégie d'émission d'équilibre du jeu à champ moyen en fonction du temps pour 3 quantités d'énergie initiale.

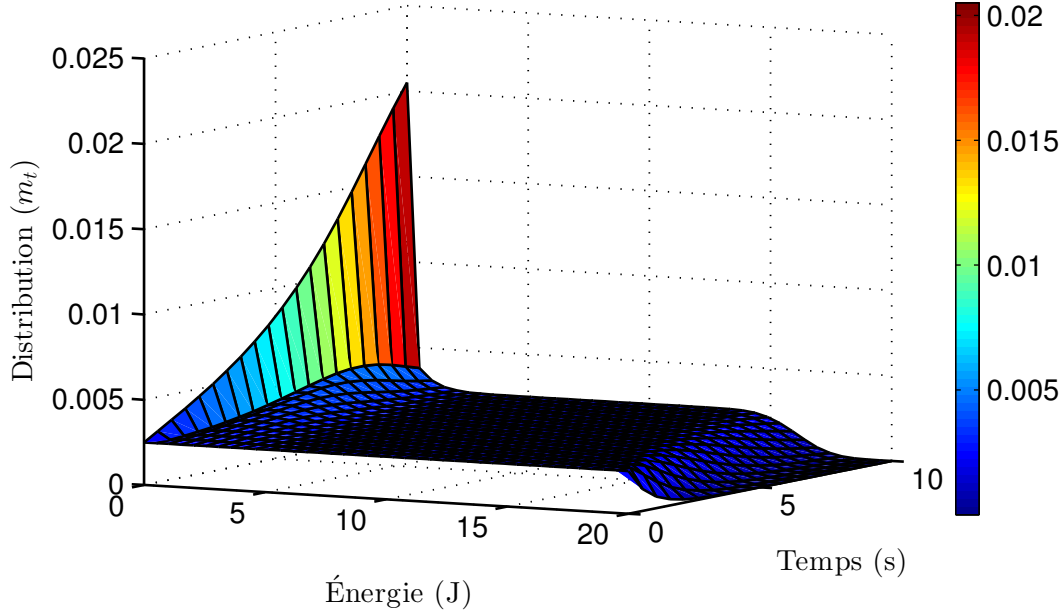


FIGURE 3.3 – Évolution de la distribution d'énergie associée à la stratégie d'émission d'équilibre du jeu à champ moyen pour une distribution d'énergie initiale uniforme.

batterie se vider. Un comportement similaire est observé dans la figure 3.4, à la différence qu'il n'y a que des terminaux avec une quantité élevée d'énergie au début de ce dernier cas. La distribution de l'énergie dans les terminaux du système évolue donc de manière regroupée.

Puisque la solution du problème de réponse au champ moyen est déterminée pour une durée du jeu donnée, la stratégie de puissance d'émission d'équilibre dépend également de cette durée. Pour cette raison, il est intéressant de faire croître la durée du jeu à 120 s, comme cela est fait dans la figure 3.5, pour en observer l'impact sur la stratégie d'émission. D'une manière générale, une durée du jeu plus longue incite les terminaux à consommer moins de puissance pour que leurs batteries restent chargées plus longtemps. Cette faible consommation offre aux terminaux une meilleure efficacité énergétique que dans le cas des stratégies d'émission issues du jeu statique ou du jeu répété standard, comme cela est illustré dans la figure 3.6.

3.3 Conclusion

Dans ce chapitre, nous avons proposé une formulation du problème de contrôle de puissance efficace énergétiquement initialement introduite dans [Goodman 2000] sous forme d'un jeu différentiel stochastique. Cette formulation permet d'optimiser l'efficacité énergétique globale du réseau, tout en tenant compte des contraintes d'énergies à long terme et des effets de propagation tels que la corrélation temporelle

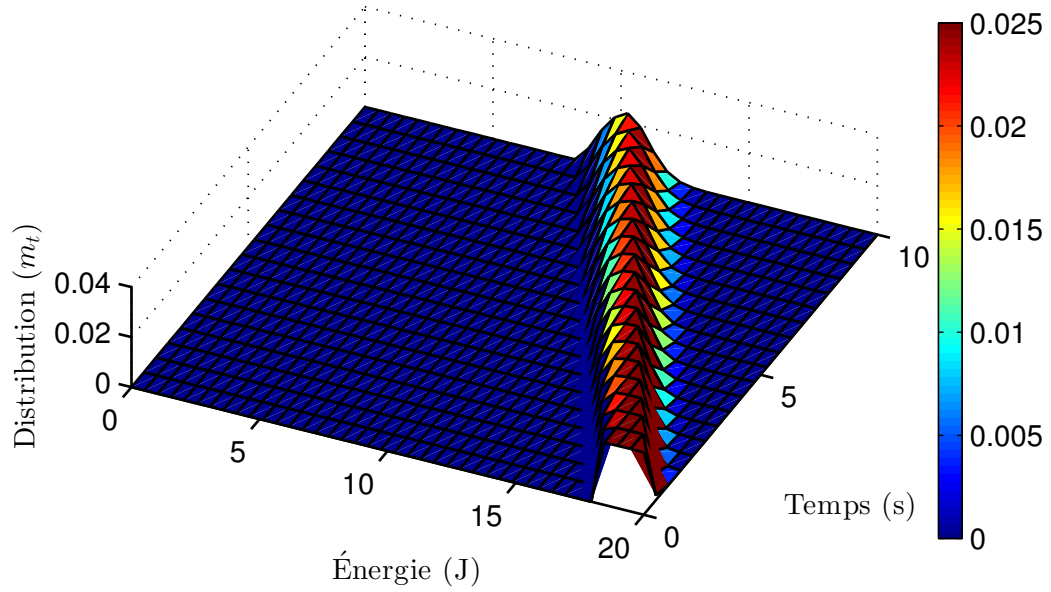


FIGURE 3.4 – Évolution de la distribution d'énergie associée à la stratégie d'émission d'équilibre du jeu à champ moyen pour une distribution d'énergie initiale non-uniforme.

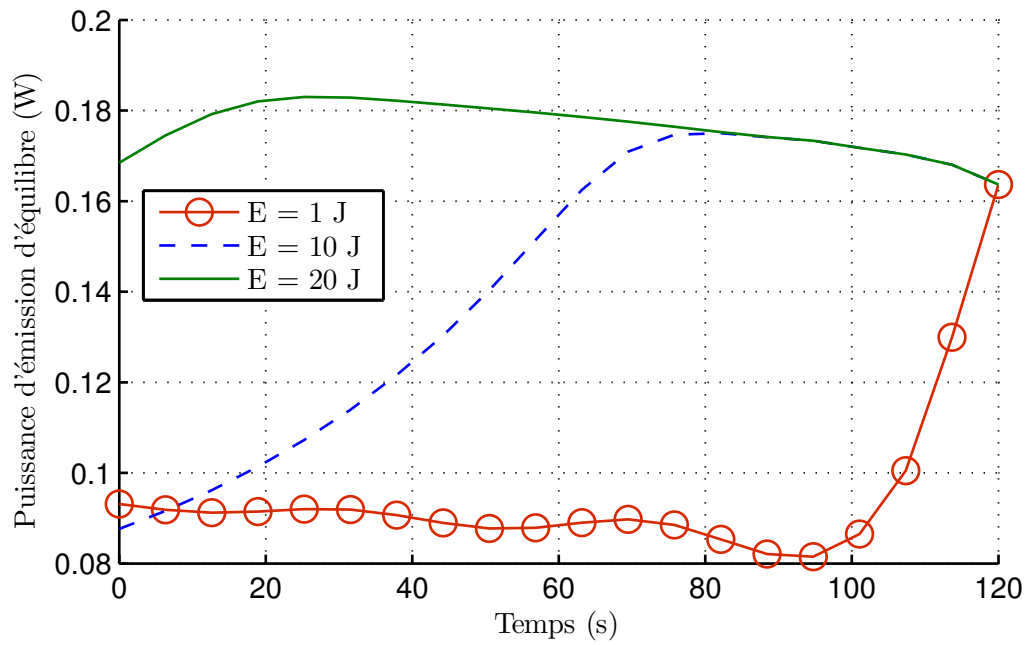


FIGURE 3.5 – Stratégies d'émission d'équilibre du jeu à champ moyen en fonction du temps, pour 3 niveaux d'énergie initiale, pour une durée du jeu longue.

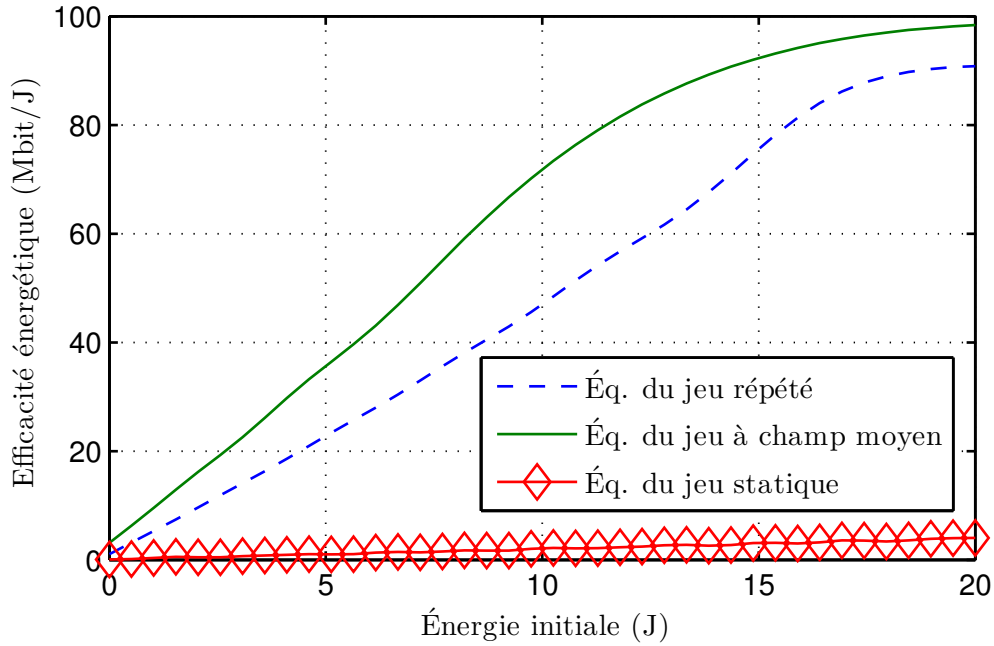


FIGURE 3.6 – Comparaison de l'utilité moyenne d'efficacité énergétique pour la stratégie d'émission d'équilibre du jeu à champ moyen avec les stratégies d'équilibre issues du jeu statique et du jeu répété standard.

des gains de canaux. Le problème est que ce modèle devient trop complexe à manipuler lorsque le nombre de terminaux dans la cellule devient grand. Au lieu de considérer les réseaux de grande dimension comme un problème, ils peuvent être perçus comme un atout puisque sous l'hypothèse d'un grand nombre de joueurs, le jeu différentiel stochastique peut être approché par un jeu à champ moyen. Sous l'hypothèse de la connaissance de l'état individuel, l'idée est précisément d'exploiter le grand nombre de joueur pour simplifier l'analyse du problème. Ce nouveau jeu simplifie l'analyse d'un équilibre puisque l'expression de l'équilibre nécessite seulement la résolution d'un système de deux équations, à partir de la connaissance de la distribution initiale d'état du système. Cette connaissance peut être obtenue par la diffusion d'un signal public dans le réseau et remarquablement, la complexité de ce signal ne dépend pas de la taille du réseau puisque la quantité d'information à transmettre ne dépend pas du nombre de terminaux dans la cellule. À l'issue de la résolution du système, chaque terminal obtient une stratégie de contrôle de puissance dépendant uniquement de l'état de son canal et de sa batterie et aucun signal public supplémentaire n'est nécessaire. Ce cadre nous permet également d'obtenir des conditions suffisantes pour l'existence et l'unicité d'un équilibre de stratégies de contrôle de puissance.

Alors que dans le cas général, les équations aux dérivées partielles résultant du jeu à champ moyen restent compliquées à résoudre, nous avons étudié un cas particulier pour lequel des résultats numériques peuvent être obtenus. Ces résultats sont

encourageant, bien qu'ils aient seulement été générés pour le cas des canaux quasi-statiques. Nous rappelons certaines conclusions issues de ces résultats qui peuvent être intéressantes pour des ingénieurs et d'autres chercheurs :

1. Les terminaux disposant d'une faible énergie initiale émettent d'abord avec une faible puissance, puis cette puissance augmente au cours du temps, pour le cas où la distribution d'énergie initiale est uniforme et la durée $T' - T$ est courte.
2. Dans les mêmes conditions, les terminaux disposant d'une forte énergie initiale émettent d'abord avec une forte puissance, puis cette puissance décroît progressivement au cours du temps.
3. Les terminaux avec une énergie initiale faible émettent d'abord avec une puissance faible qui décroît avec le temps, avant d'augmenter cette puissance quand suffisamment d'autres terminaux ont vidé leurs batterie, quand la distribution d'énergie initiale est uniforme et que $T' - T$ est grand.
4. Quand $T' - T$ est suffisamment grand, l'équilibre du jeu à champ moyen est plus performant que d'autres équilibres classiques (de [Goodman 2000] et [Le Treust 2010]) en termes d'efficacité énergétique.

Une solution numérique du cas général qui comprend des canaux variant avec le temps en plus du niveau d'énergie est une perspective de travail futur. D'autres axes de travail consistent à étudier un réseau multi-cellulaire, des stations de base distribuées ainsi que le cas des multiples porteuses.

Dynamiques de meilleures réponses et apprentissage pour les réseaux distribués

Sommaire

4.1	Jeu de positionnement de points d'accès	61
4.1.1	Modélisation du système	62
4.1.2	Étude de l'équilibre de Nash	66
4.1.3	Convergence vers l'équilibre de Nash	67
4.2	Jeu de satisfaction de contraintes de qualité de service . . .	71
4.2.1	Formulation du jeu de satisfaction	73
4.2.2	Convergence de la DMR vers un ESE	74
4.2.3	Convergence de la dynamique de réponse robuste aveugle (DRRA) vers un ESE	78
4.2.4	Résultats numériques	79
4.3	Conclusion	80

Dans les chapitres précédents, l'étude de jeux de contrôle de puissance efficace énergétiquement nous a conduit à étudier l'existence et l'unicité d'équilibres de Nash. Nous rappelons brièvement qu'un équilibre de Nash est un profil d'actions (pour les jeux statiques) ou de stratégies (pour les jeux répétés) dont aucun joueur n'a unilatéralement intérêt à dévier. Dans un système distribué, un tel point représente donc un état stable qu'aucun joueur n'a intérêt à quitter et c'est cette propriété de stabilité qui justifie l'intérêt de ces points. Pour les équilibres étudiés précédemment, les joueurs avaient une connaissance du jeu telle qu'ils pouvaient en déduire leurs actions ou leurs stratégies d'équilibre. Malheureusement le niveau de connaissance du jeu nécessaire à la détermination d'un équilibre n'est pas forcément disponible dans tous les cas d'études. Dans ce chapitre, nous étudions deux problèmes pour lesquels nous prouvons qu'il existe des états d'équilibre mais que les joueurs ne peuvent pas connaître à cause du manque d'information disponible. Pour contourner ce manque d'information, pour chacun des deux problèmes, nous étudions des dynamiques de meilleures réponses (DMR) [Fudenberg 1991] et des techniques d'apprentissage [Sutton 1998] qui permettent aux joueurs de converger en plusieurs étapes vers ces états d'équilibre.

Les travaux que nous présentons dans ce chapitre sortent du cadre du contrôle de puissance efficace énergétiquement et sont donc totalement indépendants des chapitres 2 et 3. Le premier problème considéré a été présenté pour la première fois dans [Mériaux 2011b]. C'est un jeu de positionnement de points d'accès sans fil pour maximiser le débit offert aux terminaux mobiles qui y sont connectés. Ce problème rentre dans le cadre des jeux de positionnement, qui ont été étudiés de longue date pour étudier notamment comment positionner des points de ventes qui cherchent à maximiser leurs résultats respectifs [Hotelling 1929]. Ce n'est que plus récemment que des travaux tels que [Altman 2009] et [Silva 2010] ont transposé ces problèmes dans le domaine des télécommunications. Dans le travail que nous présentons, le problème est : étant donné une distribution spatiale de terminaux mobiles et des règles d'association entre les terminaux mobiles et les points d'accès, où des points d'accès ont-ils intérêt à se positionner pour maximiser le débit qui transite par eux-mêmes sur les canaux montants ? Nous verrons que sous certaines hypothèses de rapport signal à interférence plus bruit (RSIB) faible, nous pouvons prouver qu'il existe un équilibre de Nash pour le problème en question. Plus précisément, dans le cas où les points d'accès ont des caractéristiques symétriques, nous montrons que l'ensemble des positions d'équilibre est unique. En d'autres termes, si une condition d'ordre est imposée aux points d'accès, l'équilibre de Nash est également unique. Cependant, si la condition d'ordre n'est pas connue des points d'accès (ce qui est l'hypothèse la plus réaliste), il existe plusieurs configurations de positionnement stables. Il n'est donc pas immédiat pour les joueurs de savoir quelle position d'équilibre viser. Pour résoudre ce problème, une solution est de mettre en place des techniques de meilleures réponses ou d'apprentissage pour qu'en plusieurs itérations, les joueurs convergent vers des positions d'équilibre. Nous verrons que ces différentes techniques ne requièrent pas le même niveau de connaissance du système et en conséquence, n'ont pas les mêmes garanties et vitesses de convergence. Précisément, nous prouvons la convergence des DMR séquentielles ou simultanées et nous appliquons également un apprentissage par renforcement dont l'utilisation nécessite moins d'information sur le système.

Le second problème considéré est tiré de [Mériaux 2012a]. Il concerne la garantie du respect de contraintes de qualité de service (QDS) dans les réseaux décentralisés auto-configurables (RDAC). Dans de tels réseaux, il n'existe pas de contrôleur central pouvant imposer les configurations d'émission/réception des différents terminaux dans le réseau afin d'assurer un état globalement satisfaisant. Il revient donc directement aux terminaux mobiles de choisir leur configuration de communication. Ainsi, de l'ensemble des choix des terminaux va découler si les contraintes de QDS sont satisfaites par le système. Pour tenir compte à la fois de la liberté d'action des terminaux mobiles et des contraintes de QDS à respecter, nous modélisons le problème dans le cadre des jeux de satisfaction, introduits dans [Perlaza 2012]. Dans ce type de jeux, un équilibre a la particularité d'être un état pour lequel tous les joueurs satisfont leurs contraintes de QDS. Cependant, dans le cas général, la caractérisation d'un équilibre n'est pas faisable avec des hypothèses réalistes sur la connaissance du système par chaque terminal. Pour que le réseau opère dans un état satisfaisant, il

est donc nécessaire de mettre en place des règles de fonctionnement qui convergent vers un équilibre. Pour un problème de contrôle de puissance sur le canal montant pour respecter des contraintes de RSIB, de telles règles ont déjà été proposées par [Yates 1995] et [Wu 2001]. Sous l'hypothèse qu'il existe un état d'équilibre et que les ensembles d'actions soient continus, la méthode introduite dans [Yates 1995] a l'avantage de converger à partir de n'importe quel état de départ. Cependant, dans le cas où les ensembles d'actions sont discrets, la règle de fonctionnement issue de [Wu 2001] ne converge qu'à partir de certains états particuliers du système. Dans cette seconde partie, notre contribution principale est donc de proposer une règle de fonctionnement inspirée de la DMR qui présente les avantages de ne demander à chaque joueur qu'une connaissance de son propre état et qui, pour le problème de contrôle de puissance avec des ensembles d'actions discrets, offre la garantie de converger vers un équilibre de satisfaction à partir de n'importe quel point de départ.

4.1 Jeu de positionnement de points d'accès

Les terminaux mobiles sont actuellement en train de gagner en autonomie décisionnelle pour permettre une meilleure utilisation des ressources dans les réseaux sans fil. Ainsi, les terminaux mobiles peuvent choisir avec quelle technologie accéder au réseau et à quel point d'accès ou station de base ils se connectent. Dans un futur proche, nous pourrions imaginer que cela soit fait indépendamment de l'opérateur mobile gérant le point d'accès (partage de point d'accès, itinérance). Lorsqu'un opérateur mobile déploie ses points d'accès pour son réseau mobile, il doit donc faire face à des considérations nouvelles. Si son objectif est de maximiser le trafic perçu par ses propres points d'accès (d'une part pour satisfaire ses utilisateurs et d'autre part pour maximiser les frais d'itinérance), il doit prendre en compte la présence de point d'accès concurrents pour choisir où positionner ses propres points d'accès. Si chaque opérateur mobile raisonne de la même manière, ce problème de positionnement des points d'accès peut être étudié dans le cadre de la théorie des jeux et plus précisément dans le contexte des jeux de positionnement.

L'histoire des jeux de positionnement débute avec les travaux de Hotelling [Hotelling 1929] qui introduisent la notion de concurrence spatiale dans une situation de duopole. Plus précisément, deux entreprises s'affrontent pour positionner leurs points de vente respectifs sur un segment (qui modélise une route) peuplé de clients potentiels. Le positionnement des deux points de vente a pour conséquence de diviser le segment en régions d'influence convexes pour chacune des entreprises. Un vaste inventaire des jeux de positionnement est présenté dans [Gabszewicz 1992]. Dans [Plastria 2001], Plastria donne un aperçu général des méthodes de positionnement d'un nouveau point de vente dans un environnement où des concurrents sont déjà positionnés. Les jeux de positionnement sont étendus au contexte des réseaux sans fil dans des travaux tels que [Altman 2009, Silva 2010]. La spécificité de ce contexte applicatif est que les interactions entre des points d'ac-

cès sont provoquées par les interférences mutuelles. Ce constat rend compliqué le problème d'association entre les terminaux mobiles et les points d'accès. En effet, puisque les associations entre les points d'accès et les terminaux mobiles dépendent du rapport signal à bruit plus interférence (RSIB) perçu par les terminaux, les associations dépendent non seulement du positionnement des points d'accès mais aussi de la répartition des terminaux mobiles déjà connectés aux points d'accès. Alors que [Silva 2010] se concentre sur un problème de positionnement de points d'accès en prenant en compte les canaux descendants, dans [Altman 2009], le positionnement des points d'accès et l'association des terminaux mobiles se font en fonction des canaux montants et sont traités comme un jeu de Stackelberg [Stackelberg 1934]. Le cadre de notre travail est proche de [Altman 2009] mais se différencie par l'obtention de plusieurs résultats nouveaux :

- Comme dans [Altman 2009], nous considérons que les terminaux mobiles fonctionnent en émission et sont répartis le long d'un axe unidimensionnel. Cependant, nous supposons que chaque terminal choisit le point d'accès le plus proche pour se connecter (par exemple, en se basant sur une bande de fréquence ne subissant pas d'interférences). Cela permet d'exprimer simplement la fonction d'utilité des points d'accès. En conséquence, l'existence d'un équilibre de Nash peut être démontrée.
- En raison de la symétrie du problème, de multiples équilibres de Nash existent. Précisément, tous les équilibres comprennent les mêmes positions et ce n'est que la position occupée par chaque point d'accès qui varie d'un équilibre à l'autre. En fixant l'ordre de positionnement de chaque point d'accès, nous pouvons donc montrer qu'il n'existe qu'un unique équilibre de Nash.
- Sous l'hypothèse que la distance au sol des points d'accès est négligeable en comparaison de la distance typique entre deux points d'accès, le jeu peut être simplifié sous une forme de duopole de Cournot [Cournot 1963].
- En supposant que la position des points d'accès peut varier de manière dynamique (ce qui serait par exemple pertinent dans les cas des cellules de petite dimension pour lequel uniquement une portion des cellules sont actives simultanément), nous montrons que les DMR et l'apprentissage par renforcement [Bush 1955, Sastry 1994, Narendra 1989] permettent aux points d'accès de converger vers des positions d'équilibre.

Cette partie est organisée de la manière suivante. Dans la partie 4.1.1, nous présentons la modélisation du réseau ainsi que le jeu de positionnement des points d'accès. La partie 4.1.2 donne l'étude de l'équilibre de Nash. Enfin, la partie 4.1.3 présente comment les points d'accès peuvent converger vers l'équilibre de Nash en utilisant une DMR et l'apprentissage par renforcement.

4.1.1 Modélisation du système

Considérons un plan dans $\mathbb{R}^2$. Dans ce plan, se trouvent un nombre indéterminé de terminaux mobiles ainsi que K points d'accès. Nous notons $\mathcal{K} = \{1, \dots, K\}$ l'ensemble des points d'accès. Lorsqu'un terminal mobile en position $x \in \mathbb{R}^2$ est

connecté au point d'accès k en position $x_k \in \mathbb{R}^2$, nous pouvons exprimer le RSIB de ce terminal :

$$\gamma_k(x) = \frac{P\eta_k(x)}{\sigma^2 + I_k}, \quad (4.1)$$

où P est la puissance d'émission du terminal mobile (que nous supposons constante pour tous les terminaux), σ^2 est la puissance du bruit au point d'accès k (supposée identique pour tous les points d'accès) et $\eta_k(x)$ est l'atténuation du canal montant entre le terminal mobile en position x et le point d'accès k . L'interférence perçue par le point d'accès k en position x est I_k et sera définie ultérieurement. Dans le réseau considéré, le positionnement d'un point d'accès est une décision dont les bénéfices seront essentiellement évalués sur le long terme. Nous pouvons donc raisonnablement supposer qu'à cette échelle de temps, seul l'affaiblissement de propagation est vraiment significatif pour modéliser les canaux entre les terminaux mobiles et le point d'accès. En conséquence, nous posons :

$$\eta_k(x) = \left(|x - x_k|^2 + \varepsilon^2\right)^{-\frac{\xi}{2}}, \quad (4.2)$$

où $|x|$ est la norme ℓ_2 , ε est la distance au sol du point d'accès et $\xi \geq 2$ est l'exposant d'affaiblissement. Le cas $\xi = 2$ correspond à la situation où le terminal mobile et le point d'accès sont en ligne de vue. Une valeur plus élevée de ξ est utilisée lorsqu'il y a des obstacles entre le terminal et le point d'accès.

En décodant les signaux des différents terminaux mobiles sans annulation d'interférence, il n'y a pas de hiérarchie dans le décodage des différents signaux reçus à chaque point d'accès. En conséquence, le signal en provenance du terminal en position x est décodé avec toute l'interférence et la capacité de Shannon entre le terminal et le point d'accès k s'écrit :

$$C_k(x) = \log(1 + \gamma_k(x)). \quad (4.3)$$

Sans perte de généralité, puisque la puissance d'émission des terminaux mobiles n'est pas un facteur différenciant, nous fixons $P = 1$. De plus, nous supposons que les conditions d'affaiblissement de canal sont les mêmes pour tous les terminaux, c'est-à-dire qu'ils partagent le même exposant d'affaiblissement ξ . Avec ces hypothèses, (4.3) devient :

$$C_k(x) = \log \left(1 + \frac{\left(|x - x_k|^2 + \varepsilon^2\right)^{-\frac{\xi}{2}}}{\sigma^2 + I_k} \right). \quad (4.4)$$

Quand plusieurs points d'accès sont positionnés dans le plan considéré, les terminaux mobiles peuvent choisir le point d'accès avec lequel ils souhaitent se connecter. Dans ce travail, nous émettons l'hypothèse forte que cette association se fait en choisissant le point d'accès le plus proche. Cela est notamment réalisable pratiquement si les terminaux mobiles choisissent leur association en se basant sur des signaux émis par les points d'accès dans des bandes de fréquences différentes des bandes utilisées sur les canaux montants et qui ne se recouvrent pas entre elles.

Ainsi, nous supposons que les terminaux mobiles choisissent toujours de s'associer au point d'accès le plus proche, ce qui crée des cellules convexes pour chacun des points d'accès. Par exemple, considérons un système avec deux points d'accès 1 et 2 aux positions x_1 et x_2 et un terminal mobile en position x . Nous avons :

$$C_X = \begin{cases} \log \left(1 + \frac{(|x-x_1|^2 + \varepsilon^2)^{-\frac{\xi}{2}}}{\sigma^2 + I_1} \right) & \text{si } |x - x_1|^2 \leq |x - x_2|^2 \\ \log \left(1 + \frac{(|x-x_2|^2 + \varepsilon^2)^{-\frac{\xi}{2}}}{\sigma^2 + I_2} \right) & \text{si } |x - x_2|^2 \leq |x - x_1|^2. \end{cases} \quad (4.5)$$

Fonction d'utilité d'un point d'accès

Nous considérons que la fonction d'utilité d'un point d'accès est la somme des capacités de Shannon qu'il offre aux terminaux mobiles qui lui sont associés. Sous l'hypothèse que les terminaux mobiles sont suffisamment nombreux pour être représentés par une distribution continue $d(x)$, la fonction d'utilité du point d'accès $k \in \mathcal{K}$ peut s'exprimer comme la somme continue des capacités de Shannon des terminaux connectés à k :

$$Q_k(\underline{x}) = \int_{S_k(\underline{x})} d(z) \log \left(1 + \frac{(|z - x_k|^2 + \varepsilon^2)^{-\frac{\xi}{2}}}{\sigma^2 + I_k(\underline{x})} \right) dz, \quad (4.6)$$

où $S_k(\underline{x})$ est le sous-ensemble du plan dans lequel les terminaux mobiles sont connectés avec le point d'accès k et $\underline{x} = (x_1, \dots, x_K)$ est le vecteur des positions de tous les points d'accès. Dans ce travail, nous considérons uniquement une distribution uniforme des terminaux mobiles avec $d(x) = 1$.

En ce qui concerne les interférences, nous rappelons que la pire situation est considérée, c'est-à-dire qu'il n'y a pas de mécanismes tels que la formation de voies [Litva 1996] pour en diminuer les effets. Nous considérons également qu'il n'y a pas d'interférence entre des terminaux mobiles connectés à des points d'accès différents en raison de la réutilisation de fréquences. Seuls les terminaux connectés à un même point d'accès génèrent donc de l'interférence entre eux. Ce contexte est proche des travaux présentés dans [Altman 2009] pour lesquels deux points d'accès concurrents utilisent différentes bandes de fréquence. Dans le scénario que nous étudions, chacun des K points d'accès utilise sa propre bande de fréquence.

Sans mécanisme d'élimination de l'interférence au point d'accès k , nous obtenons :

$$I_k(\underline{x}) = \int_{S_k(\underline{x})} (|z - x_k|^2 + \varepsilon^2)^{-\frac{\xi}{2}} dz. \quad (4.7)$$

L'utilité du point d'accès k (4.6) devient donc :

$$Q_k(\underline{x}) = \int_{S_k(\underline{x})} \log \left(1 + \frac{(|z - x_k|^2 + \varepsilon^2)^{-\frac{\xi}{2}}}{\sigma^2 + \int_{S_k(\underline{x})} (|z' - x_k|^2 + \varepsilon^2)^{-\frac{\xi}{2}} dz'} \right) dz. \quad (4.8)$$

Approximation de l'utilité d'un point d'accès

Pour se concentrer sur le cas critique où le nombre de terminaux mobiles connectés à chaque point d'accès est élevé, nous nous plaçons dans des régimes de faibles RSIB. Dans ce cas, la puissance du signal reçu de chaque terminal mobile est faible devant le niveau d'interférence dans (4.8). Avec cette hypothèse, une expression simplifiée de (4.8) peut être obtenue :

$$\begin{aligned} Q_k(\underline{x}) &\approx \frac{\int_{\mathcal{S}_k(\underline{x})} (|z - x_k|^2 + \varepsilon^2)^{-\frac{\xi}{2}} dz}{\sigma^2 + \int_{\mathcal{S}_k(\underline{x})} (|z' - x_k|^2 + \varepsilon^2)^{-\frac{\xi}{2}} dz'} \\ &= \frac{I_k(\underline{x})}{\sigma^2 + I_k(\underline{x})}. \end{aligned} \quad (4.9)$$

Nous noterons que cette approximation rend l'utilité basée sur la capacité de Shannon équivalente à l'utilité basée sur le RSIB telle qu'elle est étudiée dans [Altman 2009]. En d'autres termes, dans des régimes de faible RSIB, il est équivalent de travailler avec une utilité basée sur la capacité de Shannon ou le RSIB.

De plus, puisque la fonction $g(t) = t/(\sigma^2 + t)$ est strictement croissante sur l'intervalle $[0, +\infty]$, maximiser l'approximation de $Q_k(\underline{x})$ ou $I_k(\underline{x})$ est équivalent. L'utilité définitive que nous utiliserons dans le jeu défini dans la prochaine partie est donc $I_k(\underline{x})$, $\forall k \in \mathcal{K}$.

Il peut sembler surprenant que chaque point d'accès cherche à maximiser son interférence. Cependant, sous l'hypothèse qu'il n'y a aucun mécanisme d'élimination d'interférence, maximiser l'interférence I_k revient à maximiser la surface $\mathcal{S}_k$ associée aux terminaux mobiles connectés au point d'accès k . En d'autres termes, dans le scénario considéré, maximiser l'interférence d'un point d'accès revient à maximiser le nombre de terminaux mobiles associés à ce même point d'accès. C'est donc une fonction d'utilité raisonnable.

Définition du jeu de positionnement

Dans cette partie, nous étudions le cas de K points d'accès en concurrence sur un segment de longueur $L \in \mathbb{R}$. Nous rappelons que chaque point d'accès utilise une bande de fréquence différente des autres points d'accès, ce qui permet d'éviter les interférences entre cellules. La distribution des terminaux mobiles est uniforme sur le segment et l'ensemble des positions accessibles pour les points d'accès est $[0, L]$.

Définition 18 *La forme stratégique du jeu de positionnement est donnée par le triplet $\mathcal{G}^{(2)} = (\mathcal{K}, \{\mathcal{L}_k\}_{k \in \mathcal{K}}, \{I_k\}_{k \in \mathcal{K}})$, où*

- $\mathcal{K} = \{1, \dots, K\}$ est l'ensemble des joueurs, c'est-à-dire les points d'accès.
- $\{\mathcal{L}_k\}_{k \in \mathcal{K}}$ est l'ensemble des actions que les joueurs peuvent jouer. Ici

$$\mathcal{L}_k = [0, L]. \quad (4.10)$$

- $\{I_k\}_{k \in \mathcal{K}}$ est l'ensemble des utilités des joueurs.

4.1.2 Étude de l'équilibre de Nash

L'objectif de cette partie est d'étudier l'existence d'un équilibre de Nash dans le jeu défini dans la partie 4.1.1 ainsi que de donner une caractérisation de cet équilibre.

Existence

Théorème 7 *Dans le jeu introduit en définition 18, il existe au moins un équilibre de Nash.*

La preuve de cet théorème découle des conditions de Rosen [Rosen 1965] pour l'existence d'un équilibre (voir [Mériaux 2011b] pour la preuve complète).

Multiplicité des équilibres de Nash

En ce qui concerne la multiplicité des équilibres de Nash, puisque les caractéristiques des points d'accès (hauteur et puissance du bruit à la réception) sont supposées identiques, il est intéressant de noter qu'une permutation des positions des points d'accès mène à un système identique. Ainsi, sans condition d'ordre sur les points d'accès, il y a $K!$ équilibres de Nash symétriques pour lesquels l'ensemble des positions d'équilibre est toujours le même. Pour simplifier l'étude de la multiplicité de l'équilibre, nous introduisons donc la condition suivante sur les ensemble d'actions du jeu $\mathcal{G}^{(2)}$:

$$\mathcal{L}_k = \{x_k \in [0, L] \mid 0 < x_1 < \dots < x_K < L\}. \quad (4.11)$$

Avec cette condition d'ordre, il est possible de prouver qu'il existe un unique équilibre.

Théorème 8 *Dans le jeu introduit en définition 18, soumis à la condition (4.11), il existe un unique équilibre de Nash.*

La preuve de ce théorème est donnée dans [Mériaux 2011b]. Elle repose sur la condition de concavité strictement diagonale [Rosen 1965].

L'unicité de l'équilibre de Nash sous condition d'ordre peut sembler être un résultat faible en comparaison d'un résultat général d'unicité. Cependant, dans un contexte de positionnements dynamiques, les positions initiales des points d'accès suffisent à déterminer l'équilibre de Nash observé (après convergence).

Caractérisation de l'équilibre de Nash

Une caractérisation de l'équilibre est donnée dans cette partie, ainsi que des exemples pour des petites valeurs de K . Pour $\xi > 2$, lorsque chaque point d'accès

maximise sa fonction d'utilité, une solution dans le domaine réel doit satisfaire :

$$\begin{cases} x_1 = \frac{2\sqrt{(5.2^{\frac{2}{\xi}} - 2^{\frac{4}{\xi}} - 4)\varepsilon^2 + 2^{\frac{2}{\xi}}x_2^2 - 2^{\frac{2}{\xi}}x_2}}{4 - 2^{\frac{2}{\xi}}}, \\ x_k = \frac{x_{k+1} + x_{k-1}}{2}, \forall k \in \{2, \dots, K-1\}, \\ x_K = \frac{-2\sqrt{(5.2^{\frac{2}{\xi}} - 2^{\frac{4}{\xi}} - 4)\varepsilon^2 + 2^{\frac{2}{\xi}}(L - x_{K-1})^2 - 2^{\frac{2}{\xi}}x_{K-1} + 4L}}{4 - 2^{\frac{2}{\xi}}}. \end{cases} \quad (4.12)$$

Voir [Mériaux 2011b] pour le détail des calculs. Pour le cas $\xi = 2$, nous obtenons :

$$\begin{cases} x_1 = \sqrt{2\varepsilon^2 + 2x_2^2} - x_2, \\ x_k = \frac{x_{k+1} + x_{k-1}}{2}, \forall k \in \{2, \dots, K-1\}, \\ x_K = -\sqrt{10\varepsilon^2 + 2(L - x_{K-1})^2} - x_{K-1} + 2L. \end{cases} \quad (4.13)$$

En particulier, pour le jeu avec deux point d'accès et $\xi = 2$, nous obtenons :

$$\begin{cases} x_1 = L - \frac{1}{2}\sqrt{2L^2 - 4\varepsilon^2}, \\ x_2 = \frac{1}{2}\sqrt{2L^2 - 4\varepsilon^2}. \end{cases} \quad (4.14)$$

4.1.3 Convergence vers l'équilibre de Nash

Du point de vue de chaque point d'accès, connaître l'équilibre caractérisé dans la partie 4.1.2 n'est pas une tâche aisée. En effet, cela nécessiterait que chaque point d'accès connaisse la fonction d'utilité des autres points d'accès du système ainsi que la condition d'ordre (4.11). Dans un scénario pratique, cela est difficilement envisageable. C'est pourquoi nous présentons maintenant des techniques permettant d'atteindre les positions de l'équilibre de Nash d'une manière totalement distribuée : les DMR et l'apprentissage par renforcement. Le temps est également découpé en étapes pendant lesquelles les points d'accès peuvent mettre à jour leurs positions.

Dynamiques de meilleures réponses

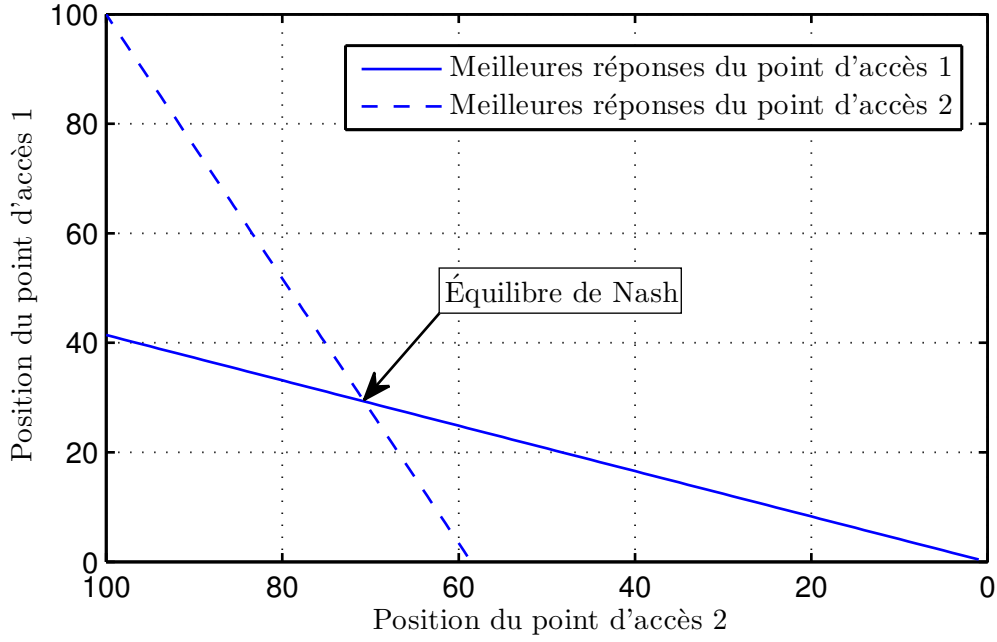
Le principe de la meilleure réponse est que chaque joueur est capable de choisir l'action qui maximise sa fonction d'utilité, pour un profil d'actions des autres joueurs donné. Précisément cette correspondance s'écrit :

$$MR_k(\underline{x}_{-k}) = \arg \max_{x_k \in \mathcal{L}_k} I_k(x_k, \underline{x}_{-k}). \quad (4.15)$$

La DMR est donc définie comme la mise à jour répétée des actions des joueurs en suivant leurs meilleures réponses. Deux dynamiques sont considérées dans cette partie :

- la DMR simultanée, pour laquelle tous les joueurs mettent à jour leurs positions en même temps en fonction des actions de l'étape précédente, c'est-à-dire que pour tout joueur $k \in \mathcal{K}$, son action à l'instant $n + 1$ est donnée par :

$$x_k^{(n+1)} = MR_k(\underline{x}_{-k}^{(n)}), \quad (4.16)$$

FIGURE 4.1 – Meilleures réponses des points d'accès 1 et 2 pour $\xi = 2$.

- la DMR séquentielle, pour laquelle les joueurs mettent à jour leurs positions chacun leur tour, c'est-à-dire que pour tout joueur $k \in \mathcal{K}$, son action à l'instant $n + 1$ est donnée par :

$$x_k^{(n+1)} = MR_k(x_1^{(n+1)}, \dots, x_{k-1}^{(n+1)}, x_{k+1}^{(n)}, \dots, x_K^{(n)}). \quad (4.17)$$

En ce qui concerne la convergence de ces dynamiques, le théorème suivant est démontré dans [Mériaux 2011b].

Théorème 9 *Si la hauteur ε des points d'accès est négligeable devant la distance caractéristique entre deux points d'accès et que les points d'accès ne sont pas superposables, les DMR simultanée et séquentielle convergent vers un équilibre de Nash à partir de n'importe quel profil de positions initiales.*

Pour le cas d'un système avec deux point d'accès, dans la figure 4.1, nous comparons les meilleures réponses séquentielles des points d'accès 1 et 2 pour $\xi = 2$. Une conséquence du fait de négliger ε est que les deux meilleures réponses sont des fonctions linéaires de l'action de l'autre joueur. Leur point d'intersection correspond à l'équilibre de Nash du jeu.

Apprentissage par renforcement

Dans cette partie, nous modifions le modèle du système étudié pour pouvoir utiliser l'apprentissage par renforcement. La modification consiste à considérer que les espaces d'actions des joueurs sont maintenant discrets. Ainsi nous notons $\tilde{\mathcal{L}} =$

$\{y_1, \dots, y_M\}$ l'ensemble des positions possibles pour tous les points d'accès (M étant la cardinalité de $\tilde{\mathcal{L}}$). Cet ensemble correspond à la discrétisation de l'ensemble $[0, L]$. Il faut bien noter que cette discrétisation des espaces d'actions donne naissance à un nouveau jeu dont l'étude d'un équilibre n'est pas faite dans ce travail. L'idée sous-jacente est qu'avec un maillage de positions suffisamment dense, les positions vers lesquelles converge l'algorithme d'apprentissage par renforcement sont proches des positions d'équilibre du jeu $\mathcal{G}^{(2)}$.

Nous implémentons une version discrète d'un algorithme d'apprentissage stochastique, ce qui signifie que l'espace d'action est discret et qu'à chaque étape, les actions sont choisies selon des distributions de probabilité qui sont mises à jour à chaque itération de l'algorithme, en fonction des valeurs des utilités observées. La seule information disponible pour chaque point d'accès est la valeur de son utilité après chaque itération. Il n'est donc pas nécessaire que le point d'accès connaisse l'expression de sa fonction d'utilité.

Définissons $\underline{o}_k^{(n)} = (o_{k1}^{(n)}, \dots, o_{kM}^{(n)})$ le vecteur des distributions de probabilité des positions du point d'accès k à l'étape n :

$$P \left[x_k^{(n)} = y_m \right] = o_{km}^{(n)}, \quad m \in \{1, \dots, M\}. \quad (4.18)$$

La routine appliquée par chaque point d'accès est alors la suivante :

1. Initialisation de la distribution de probabilité :

$$\forall k \in \mathcal{K}, \forall m \in \{1, \dots, N\}, \quad o_{km}^{(0)} = \frac{1}{M}. \quad (4.19)$$

2. A chaque étape n , chaque point d'accès k choisit une position $x_k^{(n)}$ selon la distribution de probabilité $\underline{o}_k^{(n)}$.
3. Chaque point d'accès observe $I_k^{(n)}$.
4. Chaque point d'accès met à jour sa distribution de probabilité $\underline{o}_k^{(n)}$:

$$\begin{aligned} o_{km}^{(n+1)} &= o_{km}^{(n)} - \zeta I_k^{(n)} o_{km}^{(n)}, \text{ si } x_k^{(n)} \neq y_m \\ o_{km}^{(n+1)} &= o_{km}^{(n)} + \zeta I_k^{(n)} \sum_{s \neq m} o_{ks}^{(n)}, \text{ si } x_k^{(n)} = y_m. \end{aligned} \quad (4.20)$$

5. Si $\underline{o}_k^{(n+1)} = \underline{o}_k^{(n)}$, la dynamique s'arrête. Autrement, elle reprend à l'étape 2.

Les figures 4.2 et 4.3 illustrent l'évolution des distributions de probabilité de deux points d'accès. Les paramètres de simulation sont les suivants : chaque point d'accès a le même ensemble de positions accessibles $\{10, 30, 50, 70, 90\}$ et le pas de l'algorithme d'apprentissage est $\zeta = 0.01$.

En fonction du choix de ζ , deux phénomènes peuvent se produire.

- Plus la valeur de ζ est grande, plus rapide est le temps de convergence de l'algorithme,
- Cependant, si ζ est choisi trop grand, l'algorithme peut converger vers des positions qui ne correspondent pas à l'équilibre du jeu défini dans la partie 4.1.1.

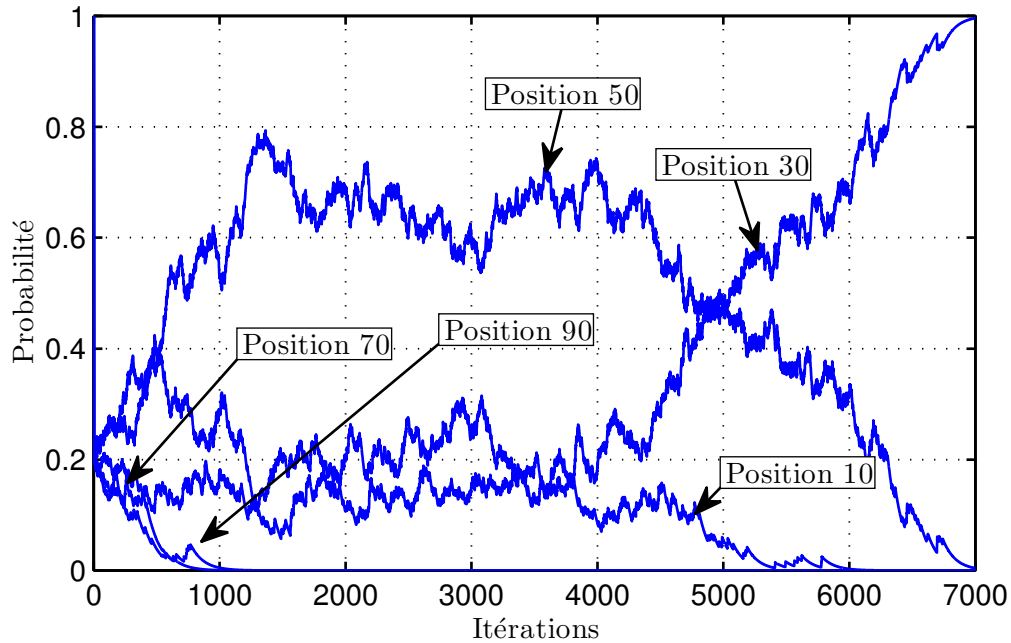


FIGURE 4.2 – Évolution des probabilités de positionnement du point d'accès 1 pour un apprentissage discret stochastique.

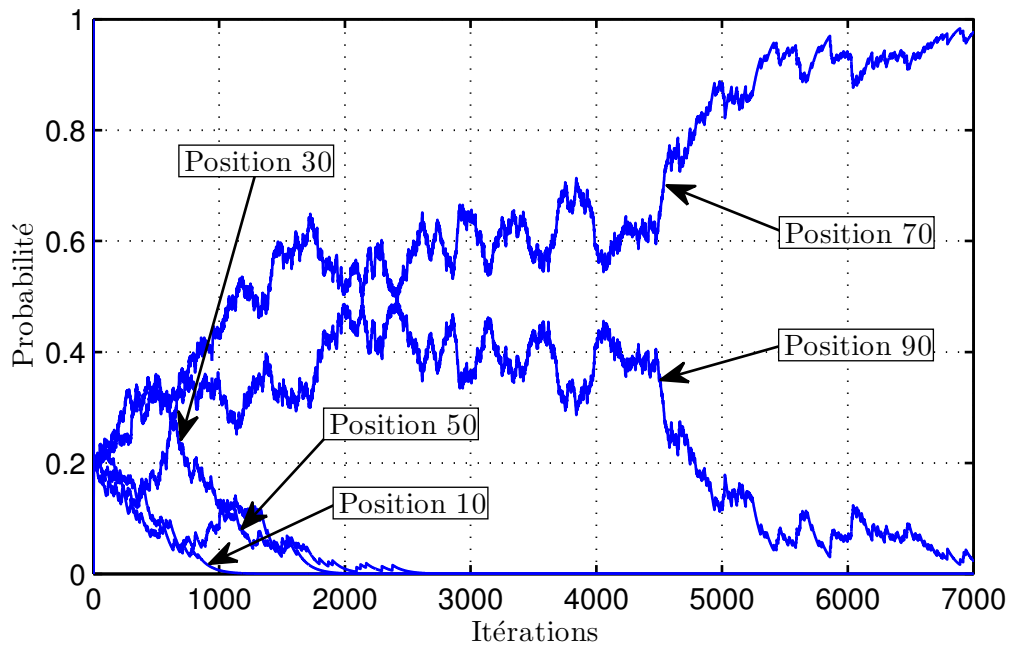
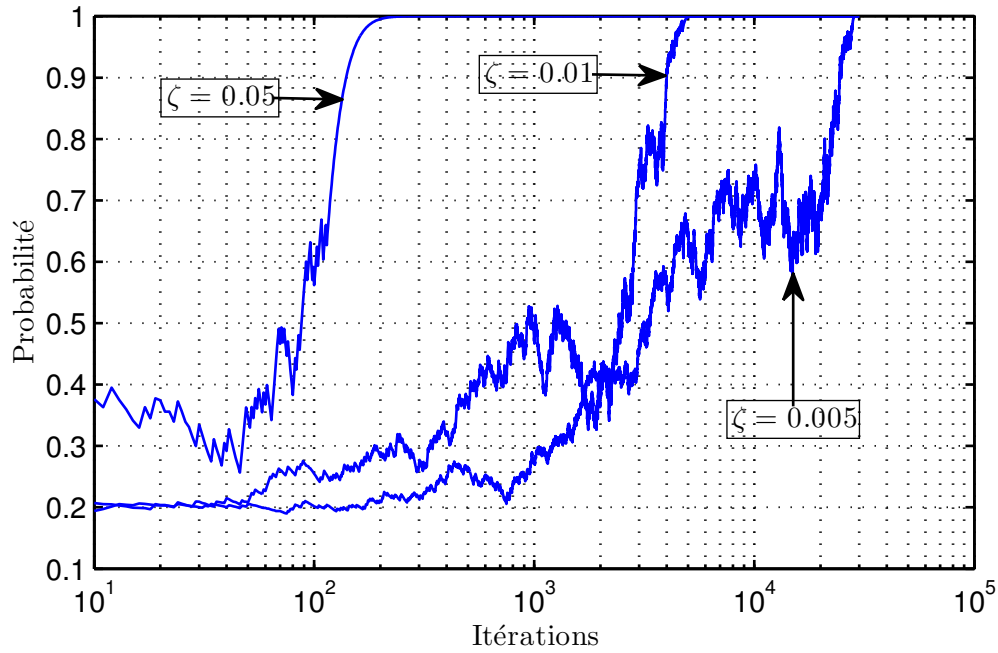


FIGURE 4.3 – Évolution des probabilités de positionnement du point d'accès 2 pour un apprentissage discret stochastique.

FIGURE 4.4 – Convergence des probabilités en fonction de la valeur du pas ζ .

La figure 4.4 illustre le temps de convergence de la probabilité d'une position en fonction de la valeur de ζ .

Le choix de ζ permet donc un compromis entre temps de convergence et probabilité de converger vers les positions souhaitées.

4.2 Jeu de satisfaction de contraintes de qualité de service

Dans la suite de ce chapitre, nous présentons un problème de satisfaction de Qualité De Service (QDS) pour lequel la méthode des meilleures réponses n'est pas assurée de converger vers un équilibre. Une variante de la meilleure réponse est donc proposée pour assurer la convergence vers un équilibre.

Un réseau décentralisé auto-configurable (RDAC) est un système de communication sans infrastructure dans lequel les terminaux mobiles choisissent leur propre configuration d'émission/réception de façon autonome, tout en garantissant des communications fiables. Typiquement, une configuration d'émission/réception peut être décrite en termes de stratégies d'allocation de puissance, de modulations, de stratégies d'ordonnancement, d'ordre de décodage, de bande de fréquence occupée, etc. Des exemples de RDAC classiques sont les réseaux de capteurs sans fil, les réseaux de faible portée dans les bandes de fréquence ISM (Wifi, Bluetooth, ZigBee...), les réseaux femto-cell et les réseaux ad hoc en général. Une constante des RDAC est que les émetteurs communiquent directement avec leur récepteur associé, sans l'in-

tervention d'un contrôleur central. Ainsi, pour que ces réseaux puissent fournir une QDS de communication suffisante, la plus grande contrainte est de limiter les interférences provenant des actions non-coordonnées des terminaux mobiles. Dans ce contexte, la garantie d'un niveau de QDS de communication impose de concevoir des règles de fonctionnement telles que les terminaux mobiles adaptent leur configuration de communication de manière à maintenir ce niveau de QDS dans une région souhaitée. En particulier, des capacités similaires d'adaptation de configuration ont déjà été mentionnées dans [Haykin 2005] dans le contexte de la radio cognitive.

De manière générale, la structure décentralisée des RDAC rend naturelle l'utilisation d'outils en provenance de la théorie des jeux [Lasaulce 2011, Han 2011]. Dans cette partie, nous nous concentrons sur une formulation particulière du problème de garantie de la QDS, à savoir les jeux de satisfaction [Perlaza 2012]. Plus spécifiquement, nous donnons des règles de fonctionnement qui permettent aux terminaux mobiles d'atteindre un équilibre de satisfaction efficace (ESE) dans les RDAC. La notion d'ESE, introduite dans [Perlaza 2012], définit un état du réseau pour lequel tous les terminaux mobiles satisfont leurs contraintes de QDS tout en produisant un effort minimal. Cette notion d'effort sous-entend que différentes configurations de communications n'ont pas les mêmes coûts pour les terminaux mobiles. Souvent, l'idée d'un effort important est associée à des configurations d'émission/réception coûteuses en énergie. Dans ce contexte, une des contributions principales de cette partie est l'introduction d'une règle de fonctionnement qui permet aux terminaux du système d'atteindre un ESE en utilisant uniquement des informations locales. Une autre contribution importante est de donner des conditions suffisantes pour que la règle proposée converge vers un ESE.

Afin de démontrer l'intérêt de nos contributions dans le contexte des RDAC, nous étudions en particulier un cas de contrôle de puissance sur les canaux montants d'un système cellulaire qui doit garantir un rapport signal à interférence plus bruit (RSIB) minimum. De manière intéressante, nous montrons que dans ce scénario particulier, la règle de fonctionnement proposée converge vers un ESE indépendamment de l'état de départ du réseau. Ce résultat contraste avec la littérature existante. Notamment, dans [Altman 2003], Altman et al. ont étudié ce problème dans le cadre général des ensembles d'actions sous forme de treillis compacts. Sous l'hypothèse qu'une solution au problème existe, ils ont établi que la DMR converge uniquement vers la solution à partir d'états de départ particuliers. Quand les ensembles de puissances d'émission sont continus, Yates et al. ont montré que la DMR convergeait à partir de n'importe quel état de départ dans [Yates 1995]. Enfin, dans le cas des ensembles d'actions discrets, un algorithme proche de la DMR est proposé dans [Wu 2001]. Cependant, il y a des conditions sur l'état de départ pour garantir la convergence de l'algorithme.

Cette partie est organisée de la façon suivante. Dans la partie 4.2.1, nous revisitons la notion d'équilibre de satisfaction (ES) et d'ESE et nous formulons un problème de contrôle de puissance pour respecter des contraintes de QDS. Dans la partie 4.2.2, nous étudions les limites de la convergence de la DMR vers un ESE. Dans la partie 4.2.3, nous présentons notre contribution principale : une règle de fonction-

nement qui permet aux RDAC de converger vers un ESE avec moins d'information que la DMR. Dans la partie 4.2.4, nous présentons les résultats numériques associés au problème de contrôle de puissance pour vérifier notre contribution théorique.

4.2.1 Formulation du jeu de satisfaction

Comme défini dans [Perlaza 2012], un jeu de satisfaction est pleinement décrit par le triplet :

$$\widehat{\mathcal{G}} = (\mathcal{K}, \{\mathcal{A}_k\}_{k \in \mathcal{K}}, \{F_k\}_{k \in \mathcal{K}}). \quad (4.21)$$

Dans ce triplet, $\mathcal{K}$ représente l'ensemble des joueurs, $\mathcal{A}_k$ est la stratégie du joueur $k \in \mathcal{K}$ et la correspondance F_k détermine l'ensemble des actions du joueur k qui lui permettent de satisfaire ses contraintes de QDS étant donné les actions des autres joueurs. Un profil de stratégie est noté $\underline{a} = (a_1, \dots, a_K) \in \mathcal{A}$.

Naturellement, un état d'intérêt d'un jeu de satisfaction est un état pour lequel tous les joueurs sont satisfaits, c'est-à-dire un équilibre de satisfaction (ES). La notion d'ES a été définie comme un point fixe dans [Perlaza 2012], de la manière suivante.

Définition 19 (Équilibre de satisfaction) *Un profil d'action $\underline{a}^+$ est un équilibre du jeu $\widehat{\mathcal{G}}$ si*

$$\forall k \in \mathcal{K}, \quad \underline{a}_k^+ \in F_k(\underline{a}_{-k}^+). \quad (4.22)$$

Comme nous le verrons, il n'y a souvent pas unicité de l'ES et il existe donc des ES aux propriétés plus intéressantes que d'autres. Dans la suite, nous introduisons la notion d'équilibre de satisfaction efficace (ESE). Dans cette optique, nous considérons une fonction de coût pour chaque joueur du jeu, pour modéliser la notion d'effort ou de coût associé à un choix d'action. Pour tout $k \in \mathcal{K}$, la fonction de coût $c_k: \mathcal{A}_k \rightarrow [0, 1]$ vérifie la condition suivante : $\forall (a_k, a'_k) \in \mathcal{A}_k^2$,

$$c_k(a_k) < c_k(a'_k), \quad (4.23)$$

si et seulement si a_k nécessite un effort moindre que l'action a'_k lorsqu'elle est jouée par le joueur k . En prenant en compte cette notion d'effort, l'ensemble des ES d'intérêt est ceux qui requièrent les efforts individuels les plus faibles. Cette notion d'équilibre est formalisée par la définition suivante.

Définition 20 (Équilibre de satisfaction efficace) *Un profil d'action $\underline{a}^*$ est un ESE du jeu $\widehat{\mathcal{G}} = (\mathcal{K}, \{\mathcal{A}_k\}_{k \in \mathcal{K}}, \{F_k\}_{k \in \mathcal{K}})$, avec la fonction de coût $\{c_k\}_{k \in \mathcal{K}}$, si*

$$\forall k \in \mathcal{K}, \quad \underline{a}_k^* \in F_k(\underline{a}_{-k}^*) \quad (4.24)$$

et

$$\forall k \in \mathcal{K}, \quad \forall a_k \in F_k(\underline{a}_{-k}^*), \quad c_k(a_k) \geq c_k(\underline{a}_k^*). \quad (4.25)$$

Notons que l'effort associé par chaque joueur à ses actions ne dépend pas du choix d'actions des autres joueurs. Ici, les joueurs peuvent choisir individuellement leur

fonction de coût, ce qui ajoute un autre degré de liberté à la modélisation du problème de QDS dans les RDAC.

Notons aussi qu'un jeu de satisfaction n'est pas un jeu dont les ensemble d'actions sont contraints, comme c'est le cas dans [Debreu 1952]. Ici l'ensemble d'actions d'un joueur est indépendant des actions des autres joueurs. La dépendance avec les actions des autres joueurs intervient uniquement à travers la satisfaction du joueur.

Jeu de contrôle de puissance

Nous considérons le cas du contrôle de puissance cellulaire sur la bande montante, tel que défini dans [Wu 2001, Yates 1995]. Dans une cellule, nous considérons K couples d'émetteurs/récepteurs indexés par $k \in \mathcal{K}$. Pour tout $k \in \mathcal{K}$, l'émetteur k utilise une puissance d'émission $p_k \in \mathcal{A}_k$, où $\mathcal{A}_k$ est un treillis compact. Pour chaque couple $k \in \mathcal{K}$, on note respectivement $p_k^{\min}$ et $p_k^{\max}$ le niveau de puissance minimale et le niveau maximal dans $\mathcal{A}_k$. Pour tout couple $(i, j) \in \mathcal{K}^2$, on note η_{ij} le gain du canal entre l'émetteur i et le récepteur j . La qualité de la transmission pour chaque couple $k \in \mathcal{K}$ peut être exprimée par la capacité de Shannon

$$u_k(p_k, \underline{p}_{-k}) = \log_2 \left(1 + \frac{p_k \eta_{kk}}{\sigma_k^2 + \sum_{j \neq k} p_j \eta_{jk}} \right) \quad [\text{bps/Hz}], \quad (4.26)$$

où σ_k^2 est la variance du bruit au niveau du récepteur k .

La contrainte de QDS pour chaque couple $k \in \mathcal{K}$ est que la capacité du canal $u_k(p_k, \underline{p}_{-k})$ soit supérieure à un seuil Λ_k , exprimé en bps/Hz. La correspondance du lien k est donc :

$$\begin{aligned} F_k(\underline{p}_{-k}) &= \{p_k \in \mathcal{A}_k \mid u_k(p_k, \underline{p}_{-k}) \geq \Lambda_k\} \\ &= \left\{ p_k \in \mathcal{A}_k \mid p_k \geq (2^{\Lambda_k} - 1) \frac{\sigma_k^2 + \sum_{j \neq k} p_j \eta_{jk}}{\eta_{kk}} \right\}. \end{aligned} \quad (4.27)$$

Un ESE du jeu de contrôle de puissance est un profil de puissance $\underline{p}^*$ tel que

$$\forall k \in \mathcal{K}, \quad p_k^* = \min \left\{ p_k \in \mathcal{A}_k \mid p_k \geq (2^{\Lambda_k} - 1) \frac{\sigma_k^2 + \sum_{j \neq k} p_j^* \eta_{jk}}{\eta_{kk}} \right\}. \quad (4.28)$$

Nous étudions à présent comment les couples émetteurs/récepteurs peuvent converger vers un tel équilibre.

4.2.2 Convergence de la DMR vers un ESE

Dans cette partie, nous donnons des conditions suffisantes pour la convergence de la DMR vers un ESE du jeu $\hat{\mathcal{G}} = (\mathcal{K}, \{\mathcal{A}_k\}_{k \in \mathcal{K}}, \{F_k\}_{k \in \mathcal{K}})$, avec les fonctions de coût $\{c_k\}_{k \in \mathcal{K}}$.

Dynamique des meilleures réponses

Dans le contexte d'un jeu de satisfaction $\widehat{\mathcal{G}} = (\mathcal{K}, \{\mathcal{A}_k\}_{k \in \mathcal{K}}, \{F_k\}_{k \in \mathcal{K}})$, nous définissons la meilleure réponse du joueur $k \in \mathcal{K}$ à un profil d'actions $\underline{a}_{-k}$ joué par les autres joueurs, par :

$$MR_k(\underline{a}_{-k}) = \arg \min_{a_k \in F_k(\underline{a}_{-k})} c_k(a_k). \quad (4.29)$$

Lorsque les joueurs mettent à jour leurs actions de manière séquentielle (méthode de Gauss-Seidel [Bertsekas 1989]), à l'itération $n + 1$, chacun des joueur met à jour son action selon la formule :

$$a_k^{(n+1)} = MR_k(a_1^{(n+1)}, \dots, a_{k-1}^{(n+1)}, a_{k+1}^{(n)}, \dots, a_K^{(n)}). \quad (4.30)$$

Dans le cas du jeu de contrôle de puissance avec un ensemble d'actions discret, la DMR peut être comparée à la version asynchrone de l'algorithme *Minimum Feasible Value Assignment* (MFVA), présenté dans [Wu 2001]. La différence est que dans [Wu 2001], les joueurs changent uniquement leur puissance lorsque le profil de puissance de l'étape précédente n'est pas satisfaisant. Avec la DMR, les joueurs choisissent la puissance minimale leur permettant d'être satisfaits indépendamment de leur état de satisfaction à l'étape précédente.

Convergence de la DMR

Pour étudier la convergence de la DMR, commençons par définir certaines notations. Soient $\underline{a} = (a_1, \dots, a_N)$ et $\underline{b} = (b_1, \dots, b_N)$ deux profils d'actions, notons $\underline{c} = \underline{a} \vee \underline{b}$ le maximum de $(\underline{a}, \underline{b})$ terme à terme, c'est-à-dire, $\underline{c} = (c_1, \dots, c_N)$ avec $c_n = \max(a_n, b_n) \forall n \in \{1, \dots, N\}$. De façon analogue, notons $\underline{a} \wedge \underline{b}$ le minimum de $(\underline{a}, \underline{b})$ terme à terme.

Définition 21 (Propriétés d'ascendance et de descendance) *La correspondance F_k vérifie la propriété d'ascendance si pour tout couple d'éléments $\underline{a}_{-k}$ et $\underline{a}'_{-k}$ dans l'ensemble $\mathcal{A}_{-k}$, $\underline{a}_{-k} = \underline{a}_{-k} \wedge \underline{a}'_{-k}$ implique que $\forall a_k \in F_k(\underline{a}_{-k})$ et $\forall a'_k \in F_k(\underline{a}'_{-k})$,*

$$\begin{cases} \min(a_k, a'_k) \in F_k(\underline{a}_{-k}), \\ \max(a_k, a'_k) \in F_k(\underline{a}'_{-k}). \end{cases} \quad (4.31)$$

La correspondance F_k vérifie la propriété de descendance si pour tout couple d'éléments $\underline{a}_{-k}$ et $\underline{a}'_{-k}$ dans l'ensemble $\mathcal{A}_{-k}$, $\underline{a}_{-k} = \underline{a}_{-k} \wedge \underline{a}'_{-k}$ implique que $\forall a_k \in F_k(\underline{a}_{-k})$ et $\forall a'_k \in F_k(\underline{a}'_{-k})$,

$$\begin{cases} \max(a_k, a'_k) \in F_k(\underline{a}_{-k}), \\ \min(a_k, a'_k) \in F_k(\underline{a}'_{-k}). \end{cases} \quad (4.32)$$

Une condition nécessaire pour qu'une fonction F_k soit ascendante ou descendante est que :

$$\forall \underline{a}_{-k} \in \mathcal{A}_{-k}, F_k(\underline{a}_{-k}) \neq \emptyset. \quad (4.33)$$

La définition d'une correspondance ascendante peut simplement se comprendre dans le contexte du contrôle de puissance distribué. Dans un tel contexte, la propriété d'ascendance signifie que si les autres joueurs augmentent leurs puissances d'émission, le joueur k doit également augmenter sa puissance pour rester satisfait. Notons également que si la propriété d'ascendance est vérifiée, il existe toujours au moins un niveau de puissance d'émission qui satisfasse le joueur k , quels que soient les niveaux de puissance joués par les autres joueurs. En particulier, si tous les joueurs utilisent leur puissance d'émission maximale, il existe une puissance d'émission satisfaisante pour le joueur k , ce qui est une hypothèse forte.

Proposition 5 *Pour tout $k \in \mathcal{K}$, si $F_k(\cdot)$ est non vide et compacte pour toute valeur de $\underline{a}_{-k}$, si $F_k(\cdot)$ possède la propriété d'ascendance ou de descendance et si $F_k(\cdot)$ est continue, alors*

- (i) *Un ESE existe.*
- (ii) *Si la DMR débute par le profil d'actions qui maximise ou minimise les coûts, pour tout $k \in \mathcal{K}$, alors la dynamique converge vers un ESE de manière monotone.*
- (iii) *Si la DMR débute d'un profil d'actions satisfaisant pour tous les joueurs, elle converge vers un ESE. Chacune des composantes du profil évolue de manière monotone.*
- (iv) *Dans un jeu comprenant deux joueurs, la DMR converge vers un ESE depuis n'importe quel profil d'actions de départ.*

La preuve de cette proposition peut se trouver dans [Mériaux 2012a].

DMR dans le jeu de contrôle de puissance

Comme indiqué précédemment, dans le cas général où les ensembles d'actions sont définis par des treillis compacts (qui incluent à la fois le cas des ensembles continus et des ensembles discrets), la DMR ne converge qu'à partir de profils d'actions particuliers. Cependant, dans le cas du jeu de contrôle de puissance avec des ensembles d'actions continus, il a été montré dans [Scutari 2006, Yates 1995] que la dynamique converge à partir de n'importe quel profil de puissance. C'est donc uniquement dans le cas des ensembles d'actions discrets que la convergence de la DMR n'est pas garantie. Pour convaincre le lecteur de ce problème, nous étudions le cas d'un jeu de contrôle de puissance à trois joueurs.

Exemple 1 *Soient $K = 3$ couples d'émetteurs/récepteurs. Pour tout $k \in \mathcal{K}$, l'émetteur k utilise la puissance d'émission $p_k \in \{p^{\min}, p^{\max}\}$. Étant données les contraintes énoncées dans la Section 4.2.1, nous choisissons arbitrairement les gains des canaux tels que*

$$\begin{aligned} F_1(p^{\min}, p^{\min}) &= F_3(p^{\min}, p^{\min}) = \{p^{\min}, p^{\max}\}, \\ F_1(p^{\min}, p^{\max}) &= F_3(p^{\min}, p^{\max}) = \{p^{\min}, p^{\max}\}, \\ F_1(p^{\max}, p^{\min}) &= F_3(p^{\max}, p^{\min}) = \{p^{\max}\}, \\ F_1(p^{\max}, p^{\max}) &= F_3(p^{\max}, p^{\max}) = \{p^{\max}\}, \end{aligned} \tag{4.34}$$

et

$$\begin{aligned} F_2(p^{\min}, p^{\min}) &= \{p^{\min}, p^{\max}\}, \\ F_2(p^{\min}, p^{\max}) &= \{p^{\max}\}, \\ F_2(p^{\max}, p^{\min}) &= \{p^{\min}, p^{\max}\}, \\ F_2(p^{\max}, p^{\max}) &= \{p^{\max}\}. \end{aligned} \quad (4.35)$$

Nous observons que F_k vérifie bien la propriété d'ascendance pour tout $k \in \mathcal{K}$. Ainsi défini, le jeu possède deux ESE :

- $(p^{\min}, p^{\min}, p^{\min})$ pour lequel chaque joueur émet à son plus faible niveau de puissance. Aucun joueur n'a intérêt à dévier de ce profil puisque tout autre puissance aurait un coût supérieur (bien que le joueur resterait satisfait).
- $(p^{\max}, p^{\max}, p^{\max})$ pour lequel chaque joueur émet à son plus haut niveau de puissance pour être satisfait. Si un joueur déviait de cette puissance maximale, il ne serait plus satisfait.

Bien que des ESE existent, en fonction du profil de puissance initial, il se peut que la DMR ne converge pas. Supposons par exemple que cette dynamique débute à partir du profil $\underline{p}^{(0)} = (p^{\max}, p^{\min}, p^{\max})$. À l'étape 1, le joueur 1 choisit la puissance minimale qui le satisfait, étant donné le profil de puissance des autres joueurs $\underline{p}_{-1}^{(0)} = (p^{\min}, p^{\max})$, c'est-à-dire,

$$p_1^{(1)} = BR_1(p^{\min}, p^{\max}) = p^{\min}. \quad (4.36)$$

Le joueur 2 choisit la puissance minimale qui le satisfait, étant donné le profil de puissance des autres joueurs $(p_1^{(1)}, \underline{p}_{-(1,2)}^{(0)}) = (p^{\min}, p^{\max})$, c'est-à-dire,

$$p_2^{(1)} = BR_2(p^{\min}, p^{\max}) = p^{\max}. \quad (4.37)$$

Le joueur 3 choisit la puissance minimale qui le satisfait, étant donné $(p_1^{(1)}, p_2^{(1)}) = (p^{\min}, p^{\max})$, c'est-à-dire,

$$p_3^{(1)} = BR_3(p^{\min}, p^{\max}) = p^{\min}. \quad (4.38)$$

À l'étape 2, le joueur 1 choisit la puissance minimale qui le satisfait, étant donné $\underline{p}_{-1}^{(1)} = (p^{\max}, p^{\min})$, c'est-à-dire,

$$p_1^{(2)} = BR_1(p^{\max}, p^{\min}) = p^{\max}. \quad (4.39)$$

Le joueur 2 choisit la puissance minimale qui le satisfait, étant donné $(p_1^{(2)}, \underline{p}_{-(1,2)}^{(1)}) = (p^{\max}, p^{\min})$, c'est-à-dire,

$$p_2^{(2)} = BR_2(p^{\max}, p^{\min}) = p^{\min}. \quad (4.40)$$

Le joueur 3 choisit la puissance minimale qui le satisfait, étant donné $(p_1^{(2)}, p_2^{(2)}) = (p^{\max}, p^{\min})$, c'est-à-dire,

$$p_3^{(2)} = BR_3(p^{\max}, p^{\min}) = p^{\max}. \quad (4.41)$$

La dynamique revient donc au profil de puissance initial et se trouve donc prise dans une boucle infinie.

4.2.3 Convergence de la dynamique de réponse robuste aveugle (DRRA) vers un ESE

La DMR présente plusieurs inconvénients. Premièrement, nous venons juste d'illustrer que dans un jeu avec plus de deux joueurs, il se peut que la dynamique ne converge pas vers un équilibre. Deuxièmement, pour obtenir sa meilleure réponse, chaque joueur $k \in \mathcal{K}$ doit connaître l'ensemble $F_k(\underline{a}_{-k}) \forall \underline{a}_{-k} \in \mathcal{A}_{-k}$, ce qui implique une connaissance forte du système. Pour pallier à ces inconvénients, nous proposons une dynamique qui nécessite moins d'information sur le système et qui converge cependant vers un ESE. Commençons par définir la réponse robuste aveugle (RRA) par $RRA_k: \mathcal{A} \rightarrow \mathcal{A}_k$, telle que :

$$(a_k, \underline{a}_{-k}) \rightarrow \begin{cases} a'_k, & \text{si } a'_k \in F_k(\underline{a}_{-k}), a_k \in F_k(\underline{a}_{-k}) \text{ et } c_k(a'_k) \leq c_k(a_k), \\ a'_k, & \text{si } a'_k \in F_k(\underline{a}_{-k}) \text{ et } a_k \notin F_k(\underline{a}_{-k}), \\ a_k, & \text{autrement,} \end{cases} \quad (4.42)$$

où l'action a'_k est choisie aléatoirement dans $\mathcal{A}_k$ selon une loi uniforme. Chaque fois qu'un joueur $k \in \mathcal{K}$ applique la RRA, une action de son ensemble d'action $\mathcal{A}_k$ est choisie, sans tenir compte des actions des autres joueurs. Chaque joueur $k \in \mathcal{K}$ a simplement besoin de savoir si sa nouvelle action et son action précédente le satisfont, puis de comparer leurs coûts respectifs. Si les deux actions permettent au joueur d'être satisfait, ce dernier choisit l'action avec le coût le plus faible. Si la nouvelle action permet au joueur d'être satisfait alors que la précédente ne le permet pas, le joueur choisit la nouvelle action. Dans le cas contraire, il maintient son action précédente. La dynamique de réponse robuste aveugle (DRRA) est définie par la situation où tous les joueurs utilisent la RRA de manière séquentielle, tel que $\forall k \in \mathcal{K}$

$$a_k^{(n+1)} = RRA_k(a_1^{(n+1)}, \dots, a_{k-1}^{(n+1)}, a_{k+1}^{(n)}, \dots, a_K^{(n)}), \quad (4.43)$$

Le théorème suivant s'applique à la DRRA.

Théorème 10 *Si pour tout $k \in \mathcal{K}$, $F_k(\cdot)$ est non-vide et compacte pour toute valeur de $\underline{a}_{-k}$, $F_k(\cdot)$ vérifie la propriété d'ascendance et est continue, et $c_k(\cdot)$ est strictement croissante. Alors les propriétés suivantes sont vérifiées :*

- (i) *Si la dynamique débute par un profil d'action satisfaisant pour tous les joueurs, la séquence de RRA converge vers un ESE. La séquence est décroissante pour toutes ses composantes.*
- (ii) *Si la dynamique débute par le profil d'actions les plus coûteuses, $\forall k \in \mathcal{K}$, la séquence de RRA converge monotonement vers un ESE.*
- (iii) *Dans un jeu à deux joueurs, la séquence de RRA converge vers un ESE à partir de n'importe quel profil de départ.*

A nouveau, la preuve n'est pas donnée dans cette partie mais peut se trouver dans [Mériaux 2012a].

En comparaison de la DMR, le principal avantage de la DRRA est de ne pas nécessiter la connaissance d'une expression explicite de F_k . En conséquence, seule

une connaissance locale du système est requise pour chaque joueur. Il suffit donc que chaque joueur ait la capacité de savoir s'il est satisfait ou non pour pouvoir utiliser la DRRA.

DRRA dans le jeu de contrôle de puissance

Dans le cas du jeu de contrôle de puissance avec des ensembles d'actions discrets, la DRRA vérifie une propriété de convergence extrêmement pratique (voir [Mériaux 2012a] pour la preuve).

Théorème 11 *Dans le jeu d'allocation de puissance défini dans la partie 4.2.1, avec des ensembles d'actions discrets, c'est-à-dire, $\forall k \in \mathcal{K}$, $\mathcal{A}_k = \{p_k^{(1)}, \dots, p_k^{N_k}\}$ avec N_k le nombre de niveaux de puissance dans l'ensemble d'actions $\mathcal{A}_k$, les DMR convergent vers un ESE depuis n'importe quel état de départ.*

Le théorème 11 assure la convergence vers un ESE lorsqu'un joueur rejoint ou quitte le jeu (ou lorsque les gains des canaux varient). En effet, supposons que K émetteurs se trouvent dans un ESE $\underline{p}^*$ et qu'un nouvel émetteur rejoigne le système. Un nouveau jeu débute alors avec $K+1$ joueurs. D'après le théorème 11, la convergence vers un nouvel ESE (s'il existe) est assurée à partir du profil $(\underline{p}^*, p_{K+1})$.

4.2.4 Résultats numériques

Dans cette partie, nous présentons des résultats numériques pour le jeu de contrôle de puissance avec des ensembles d'actions discrets, tel que défini en partie 4.2.1.

Dans la figure 4.5, nous montrons les séquences d'actions convergeant vers un ESE pour la DRRA dans un jeu à deux joueurs. La région colorée correspond à la région de satisfaction, c'est-à-dire, la région de puissance permettant aux deux joueurs d'être satisfaits. La coloration de cette région suit la somme des coûts de puissance pour chaque joueur. La DRRA converge d'abord vers la région de satisfaction, puis converge vers un ESE, tout en restant dans la région de satisfaction.

Dans les figures 4.6 et 4.7, nous présentons un scénario qui illustre la différence entre la DMR et la DRRA dans une cellule avec 3 émetteurs. Durant les 200 premières étapes, seuls les émetteurs 1 et 3 participent au jeu, puis l'émetteur 2 les rejoint pour les 200 étapes suivantes et finalement, l'émetteur 3 quitte la cellule durant les dernières 200 étapes. Sur chacune des deux figures, nous montrons la séquence des indices de niveau de puissance pour les 3 émetteurs, sachant que chaque ensemble d'actions est composé de $N_k = 32$ niveaux de puissances possibles allant de 10^{-6} W à 10^{-2} W. Nous montrons également l'état de satisfaction des trois émetteurs : pour chacune des étapes de la dynamique, si tous les émetteurs sont satisfaits, l'état de satisfaction vaut 1, autrement il vaut 0. La figure 4.6 et la figure 4.7 correspondent respectivement à la DMR et à la DRRA.

Les gains des canaux et les profils de puissance de départ des deux simulations sont les mêmes. Les gains des canaux valent $\eta_{22} = 10^{-5}$, $\eta_{11} = \eta_{33} = \eta_{13} = \eta_{21} =$

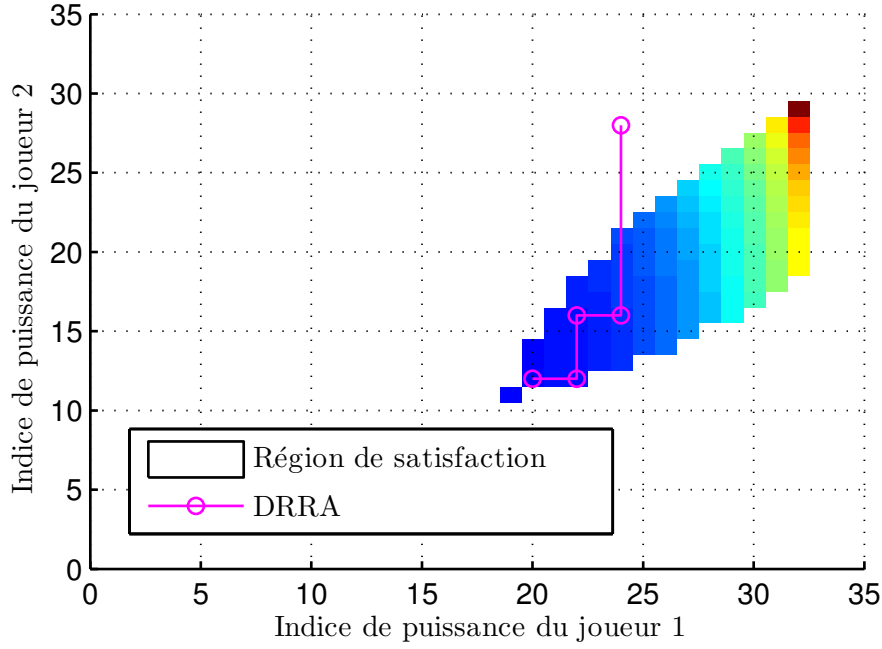


FIGURE 4.5 – Séquence des indices de puissance pour la DRRA dans le jeu de contrôle de puissance à deux joueurs. La région colorée est la région de satisfaction, c'est-à-dire, la région dans laquelle les deux joueurs satisfont mutuellement leurs contraintes de QDS.

$\eta_{32} = 10^{-6}$, $\eta_{12} = \eta_{23} = \eta_{31} = 10^{-7}$ et les émetteurs débutent respectivement à des puissances d'émission de 10^{-3} W, $10^{-5/2}$ W, et $10^{-9/4}$ W. Les seuils de satisfaction Λ_1 , Λ_2 , et Λ_3 sont fixés à 1.2 bps/Hz, 1.5 bps/Hz, et 1.2 bps/Hz, respectivement. La variance du bruit vaut 10^{-10} W pour tous les émetteurs. Il est intéressant de noter que la DMR converge vers l'ESE pendant la première et troisième phase, mais quand l'émetteur 2 entre dans la cellule durant la seconde phase, la convergence n'est pas assurée. En effet, durant cette phase, les émetteurs se trouvent dans une boucle et ne sont pas satisfaits. En ce qui concerne la DRRA, bien que le temps de convergence soit supérieur, elle présente l'avantage de converger durant les trois phases. Un autre point important est qu'avec cette dynamique, les émetteurs sont plus souvent dans un état de satisfaction qu'avec la DMR.

4.3 Conclusion

Dans ce chapitre, nous avons étudié deux problèmes distincts pour lesquels les équilibres de fonctionnement ne peuvent pas être formellement connus par les joueurs impliqués. Cependant, même sans connaître les équilibres, il est possible d'appliquer des dynamiques de mise à jour des actions des joueurs qui assurent la convergence vers un point d'équilibre.

Le modèle de jeu de positionnement proposé est très simple et demande à être

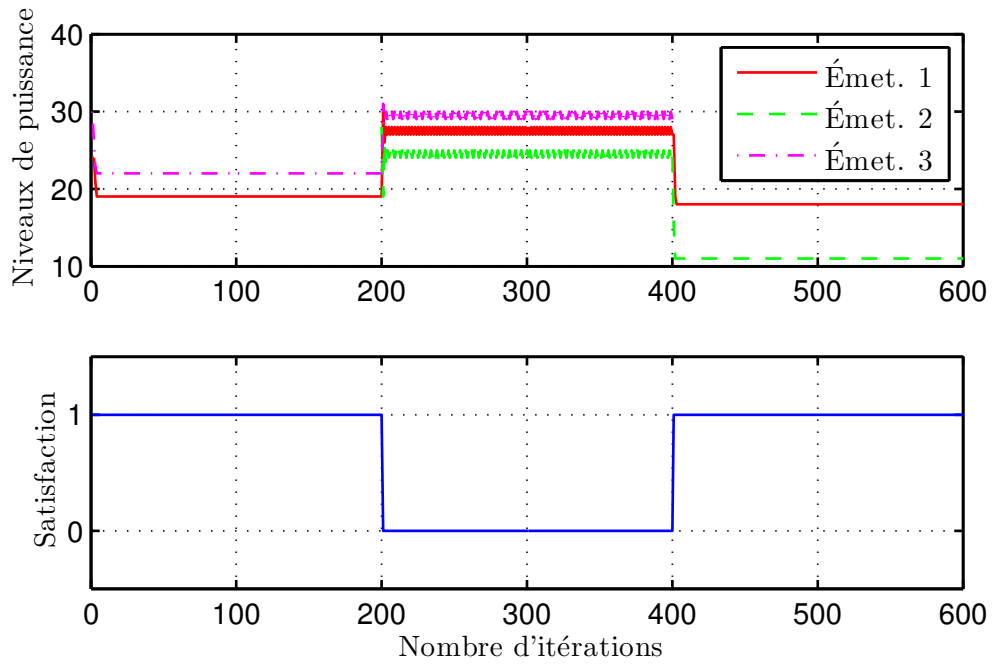


FIGURE 4.6 – Séquences des indices de puissance et états de satisfaction pour la DMR dans le jeu de contrôle de puissance avec trois joueurs.

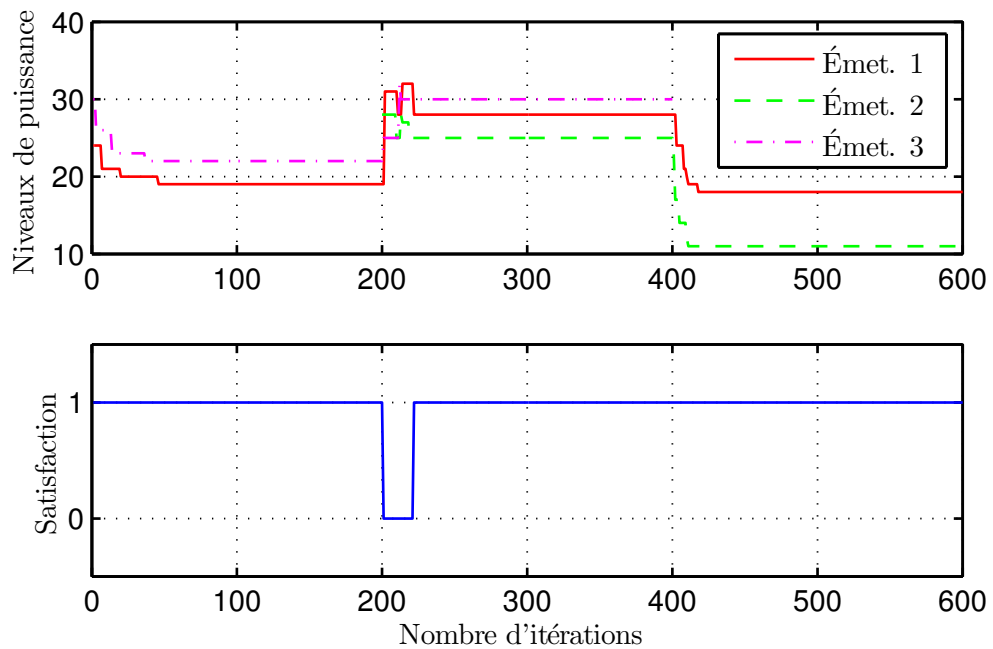


FIGURE 4.7 – Séquences des indices de puissance et états de satisfaction pour la DRRA dans le jeu de contrôle de puissance avec trois joueurs.

étendu pour offrir des résultats plus pertinents. Cependant, ce modèle présente d'ores et déjà l'avantage de permettre la caractérisation des positions d'équilibre de Nash et d'appliquer des dynamiques de convergence vers ces positions d'équilibre. Précisément, la convergence des DMR séquentielle et simultanée est prouvée et la convergence de l'apprentissage par renforcement est vérifiée expérimentalement. La perspective naturelle de ce travail serait d'étudier le jeu de positionnement avec des espaces d'actions en deux dimensions. Deux voies s'ouvrent dans cette direction. Premièrement, la version continue du problème multi-source de Fermat-Weber [Fekete 2005] peut être appliquée à notre scénario puisque ce problème consiste précisément à positionner des points de ventes dans un espace continu de clients en 2 dimensions. Deuxièmement, puisque dans notre scénario, le positionnement de chaque point d'accès crée une cellule de terminaux mobiles associés à ce point d'accès, les jeux dits de Voronoi [Dürr 2007] seraient également un contexte d'étude adapté à notre problème.

Dans la seconde partie, dans le cadre des jeux de satisfactions, nous avons proposé une règle de fonctionnement qui converge vers un ESE dans le cas général où les ensembles d'actions sont des treillis compacts. En comparaison de la DMR, la règle proposée nécessite moins d'information sur le système mais son temps de convergence est plus long. Dans le cas particulier du jeu de contrôle de puissance sur le canal montant avec des ensembles d'actions discrets, nous avons prouvé que la règle proposée converge vers un ESE à partir de n'importe quel profil d'actions de départ. Cette règle a donc la caractéristique particulière d'être robuste à l'entrée ou la sortie de terminaux dans le système. Cependant, une hypothèse forte de ce travail est que pour chaque joueur, il existe au moins une action qui le satisfasse quelles que soient les actions des autres joueurs. Dans le jeu de contrôle de puissance, il serait plus pertinent de considérer également des scénarios pour lesquels les niveaux de puissance d'émission des autres joueurs peuvent être trop élevés pour qu'un joueur donné puisse être satisfait. En conséquence, une perspective naturelle de ce travail est de relaxer cette hypothèse et d'étudier la convergence des dynamiques dans ce nouveau contexte.

Conclusion et perspectives

L'objectif principal de cette thèse était l'étude des réseaux sans fil décentralisés. Dans ce type de réseaux, la théorie des jeux s'impose comme un outil particulièrement adapté à la modélisation des interactions stratégiques entre terminaux. De plus, le volume croissant des échanges de données dans les réseaux sans fil incite à contenir le coût énergétique des transmissions de données afin de limiter les puissances d'émission des points d'accès et de préserver l'autonomie des terminaux mobiles. Dans ce contexte, nous avons choisi d'étudier l'efficacité énergétique dans les réseaux distribués puisque c'est une métrique de performance permettant naturellement d'optimiser l'énergie dépensée lors des transmissions de données. Le problème de contrôle de puissance efficace énergétiquement sur un canal à accès multiple est le cadre applicatif qui a donc été choisi comme point de départ de cette thèse. Les travaux que nous avons développé autour de cet axe couvrent un spectre théorique comprenant les jeux stochastiques, les jeux à champ moyen, la théorie du contrôle et la théorie de l'apprentissage automatique.

Parmi les contributions présentées dans ce manuscrit, nous avons étudié un problème de reports stratégiques de gain de canaux dans une cellule où la puissance est allouée aux terminaux mobiles, en fonction de leurs reports, pour maximiser l'efficacité énergétique de la cellule. Nous montrons que puisque les intérêts de la cellule et des terminaux mobiles ne sont pas alignés, les reports stratégiques peuvent dégrader l'efficacité énergétique de la cellule, bien que cet effet soit limité dans les cellules de petite dimension.

Dans la lignée des travaux de Le Treust et Lasaulce qui avaient précédemment étendu le modèle du problème de contrôle de puissance efficace énergétiquement à un jeu répété, nous avons proposé un modèle de jeu stochastique qui présente l'avantage de tenir compte des interactions stratégiques qui apparaissent entre les joueurs en raison de la répétition des actions, ainsi que des variations des canaux de transmission au cours du temps. Dans ce modèle, nous caractérisons la région d'efficacité énergétique des stratégies d'équilibre et nous établissons une stratégie de contrôle de puissance d'équilibre explicite pour laquelle les terminaux mobiles n'émettent pas lorsque leur gain de canal est trop faible. En termes d'efficacité énergétique, lorsque les terminaux mobiles ont des gains de canaux variant suivant les mêmes lois, nous montrons que cette stratégie est meilleure que les stratégies de contrôle de puissance issues du modèle de jeu statique et du modèle de jeu répété.

Puisqu'un de nos objectifs est d'optimiser la façon dont l'énergie est dépensée dans les réseaux décentralisés, un pendant naturel est de prendre en compte la quantité d'énergie disponible dans chaque terminal mobile. Pour ce faire, nous avons

introduit un modèle de jeu différentiel stochastique qui tient à la fois compte de l'évolution de l'énergie disponible dans les terminaux mobiles en fonction de leurs niveaux de puissance d'émission et de l'évolution des canaux au cours du temps. Pour ce jeu, une condition suffisante est donnée pour l'existence d'un équilibre de Nash. Cependant la caractérisation explicite d'une stratégie de contrôle de puissance optimale s'avère trop complexe et nous avons montré que l'étude du jeu à champ moyen issu du jeu différentiel stochastique (en tant que limite lorsque le nombre de joueurs et grand) permet en partie de pallier à cette complexité. En effet, dans le jeu à champ moyen, la stratégie optimale de contrôle de puissance est solution d'un système comprenant uniquement deux équations différentielles. Dans un scénario simple, nous illustrons des stratégies de contrôle de puissance optimales par des résultats numériques qui indiquent que les terminaux mobiles ont intérêt à tenir compte de l'énergie disponible dans leurs batteries pour choisir leurs niveaux de puissance d'émission.

Enfin, nous avons étudié des dynamiques d'apprentissage d'équilibre dans les réseaux décentralisés. Pour un jeu de positionnement de points d'accès, nous avons démontré la convergence des dynamiques de meilleures réponses vers des positions d'équilibre. Un algorithme d'apprentissage par renforcement a également été proposé, sans toutefois que sa convergence soit prouvée. Cette approche dynamique peut notamment être utile pour un opérateur mobile qui souhaite savoir comment activer un réseaux de petits points d'accès tels que des femtocells. Dans l'étude de la satisfaction de contraintes de qualité de service dans les réseaux décentralisés auto-configurables, nous avons proposé une dynamique proche des meilleures réponses qui ne nécessite de chaque terminal que des informations locales. D'une part, cette dynamique permet aux terminaux de se maintenir dans une zone de satisfaction de leurs contraintes. D'autre part, dans un jeu de contrôle de puissance avec des ensembles d'actions discrets, la dynamique proposée possède des propriétés de convergence plus intéressantes que la dynamique des meilleure réponses, alors qu'elle nécessite moins d'information sur l'ensemble du jeu. Elle permet notamment au système d'être robuste à l'entrée et la sortie de terminaux.

Les perspectives immédiates de cette thèse sont liées aux jeux à champ moyen. Tout d'abord, les conditions de convergence du jeu différentiel stochastique vers le jeu à champ moyen sont pour l'instant trop fortes, puisqu'elles consistent en une symétrie des conditions initiales des terminaux ainsi que des constantes d'évolution des canaux. De telles conditions ne sont donc pas applicables à un scénario pratique. En travaillant sur les conditions d'échangeabilité, nous pensons qu'il est possible d'obtenir des conditions de convergence moins contraignantes. Un autre axe de travail concerne la résolution du jeu à champ moyen. En effet, bien que moins complexe que la résolution du jeu différentiel stochastique, la résolution du jeu à champ moyen dans le cas général reste compliquée. Il est donc nécessaire d'approfondir la résolution numérique du jeu à champ moyen, dans la lignée de travaux tels que [Achdou 2010], afin d'obtenir des stratégies de contrôle de puissance dans des scénarios plus complets que celui présenté dans ce manuscrit.

Publications relatives aux jeux statiques et jeux stochastiques

A.1 IEEE-VALUETOOLS-2012

F. Mériaux, S. Valentin et S. Lasaulce. *An Energy-Efficient Power Allocation Game with Selfish Channel State Feedback in Cellular Networks*. In IEEE Conference on Performance Evaluation Methodologies and Tools (VALUETOOLS), pages 290-297, 2012.

An Energy-Efficient Power Allocation Game with Selfish Channel State Reporting in Cellular Networks

François Mériaux^{*†}, Stefan Valentin[†], Samson Lasaulce^{*} and Michel Kieffer^{*}

^{*}Laboratoire des Signaux et Systèmes – LSS (CNRS-SUPELEC-Paris Sud), Gif-sur-Yvette, France

Email: {meriaux, lasaulce, kieffer}@lss.supelec.fr

[†]Bell Labs, Alcatel-Lucent Deutschland AG, Stuttgart, Germany

Email: stefan.valentin@alcatel-lucent.com

Abstract—With energy-efficient resource allocation, mobile users and base station have different objectives. While the base station strives for an energy-efficient operation of the complete cell, each user aims to maximize its own data rate. To obtain this individual benefit, users may selfishly adjust their Channel State Information (CSI) reports, reducing the cell's energy efficiency. To analyze this conflict of interest, we formalize energy-efficient power allocation as a utility maximization problem and present a simple algorithm that performs close to the optimum. By formulating selfish CSI reporting as a game, we prove the existence of a unique equilibrium and characterize energy efficiency with true and selfish CSI in closed form. Our numerical results show that, surprisingly, energy-efficient power allocation in small cells is more robust against selfish CSI than cells with large transmit powers. This and further design rules show that our paper provides valuable theoretical insight to energy-efficient networks when CSI reports cannot be trusted.

Index Terms—Energy Efficiency, Power Allocation, Feedback, Game Theory

I. INTRODUCTION

Reducing the energy consumption of cellular networks is a challenging task for network operators and telecommunication equipment vendors. One relevant approach is the energy-efficient allocation of wireless resources such as bandwidth, time and transmit power. By carefully allocating these resources to the mobile users, a base station can reduce the energy consumption of the complete cell [12], [9].

However, such centralized form of energy-efficient resource allocation raises two problems. The first problem results from the time-variant nature of the wireless fading channel. To adapting to the users' varying channels, the base station has to update the resource allocation frequently, e.g., once per millisecond in most LTE systems [2, Fig. 5.1-1]. This renders computational complex approaches infeasible. The second problem results from fading and interference, causing most wireless channels to be non-reciprocal. Consequently, the channel state is only known at the mobile user and needs to be transferred to the base station. The accuracy of this CSI

feedback is crucial for the quality of the resource allocation decision. If users report inaccurate CSI, the resource allocation does not reflect the actual channel states and the performance of the overall network may degrade.

In this paper, we analyze energy-efficient resource allocation when users and base station have different interests. While the base station strives for energy efficiency of the whole cell, each users aims to maximize its individual Signal-to-Noise Ratio (SNR). Note that in a perfect cellular network the base station would simply overrule the users' interest. In practice, however, the base station controls the CSI format [1, Sec. 7.2] but not how mobile users generate the reported CSI values. Thus, mobile users can selfishly report CSI that provides them an individual benefit while the performance of the overall network suffers.

This makes it necessary to analyze the effect of selfishly chosen CSI on the performance of the wireless cell. To do so, we apply a game theoretical approach. Here, the mobile users are reflected as players who selfishly choose their CSI to maximize their individual performance. For all users, the base station computes the resource allocation by maximizing a utility function. This function expresses the energy efficiency of the complete cell. Limiting our analysis to power allocation allows us to formally compare centralized power allocation with true CSI to the allocation with selfishly chosen CSI. Consequently, our results provide significant insight to energy-efficient cellular networks in which the CSI accuracy cannot be trusted.

A. Contributions

In particular, we make the following contributions:

- 1) We analyze the optimization problem, where the base station maximizes the cell's energy efficiency under power constraints. We obtain an optimality condition and propose a simple power allocation algorithm whose energy efficiency is close to the optimum in general.
- 2) We prove that there exists a unique equilibrium for the users' selfish choice of the CSI reports. This allows studying the network in a stable state, where no user has an unilateral interest to report a different CSI.

François Mériaux has been supported by the *PhD@Bell Labs* internship program.

Michel Kieffer has been partly supported by the Institut Universitaire de France.

- 3) We compare the performance with true and selfish CSI reporting. For both types of reports, we provide a closed-form result for the cell's energy efficiency. Numerical results surprisingly show that small cells are more robust to selfish CSI than cells with large transmit power constraints.

All in all, our paper provides the theoretical insight to cope with selfish CSI reports in cellular networks with energy-efficient power allocation.

B. Paper Structure

Our paper is structured as follows. After discussing the novelty of our study with respect to related work in Sec. II, we describe the studied cellular system in Sec. III-A. Energy efficiency is formalized in terms of a constrained utility maximization problem in Sec. III-B. Using this function, we formulate the non-convex optimization problem in Sec. IV. Having discussed further properties of this problem, we provide an optimality condition and derive a power allocation algorithm. Sec. V is devoted to the game theoretical study of selfish CSI reports. Therein, we prove the existence of a unique equilibrium for the proposed energy-efficient power allocation. Finally, the numerical results in Sec. VI point out by how much a selfish choice of CSI reduces the energy efficiency of the system and by how much it increases the users' individual SNRs.

II. RELATED WORK AND NOVELTY

Our paper joins two fields of research. The first field is energy-efficient power control, where often transmit power is minimized under Quality of Service (QoS) constraints [5], [13]. We do not follow this common approach in our work. Instead, we maximize of the ratio between throughput and transmit power. This utility function was introduced in [6] and was used in [10], [4] for Code-Division-Multiple-Access (CDMA) systems. With CDMA, employed such utility function for power allocation results in a game where the players interact via the interference term. Our work differs, as the players interact via the power allocation algorithm of the base station. Such centralized resource allocation is common with cellular systems such as LTE and has not been studied with the utility functions from [6] so far.

The second related field is scheduling with non-cooperative mobile users. Typically, a centralized scheduler at the base station allocates wireless channel resources to maximize the instantaneous sum throughput of its cell. As such allocation is based on CSI reports, some users may not report their actual channel gains to obtain more resource from the base station. In [7], Kavitha et al. model this selfish choice of CSI as a signaling game [11]. Here, even the base station is considered as a player that knows which mobile users cooperate and which not. The same authors take a similar approach in [8] for α -fair scheduling. Here, a signaling game cannot be used since each resource allocation is affected by the scheduling history.

Although we focus on selfish CSI reporting, our work differs significantly. Unlike the above papers, we do not include the base station in the set of players. Instead, the base station performs centralized power allocation under its own general objective – energy efficiency per cell. This objective differs from the users' aim, which strive for maximizing their individual SNR. This conflict of interest between the user's individual objective and the base station's objective per cell has not been studied so far.

III. SYSTEM MODEL

A. Wireless Scenario

We study a single wireless cell where one base station allocates transmit power to K mobile users during the downlink. An arbitrary user is denoted by $k \in \mathcal{K} = \{1, \dots, K\}$ and the downlink transmit power is $p_k \in [0, P]$ Watts for each user. Time is divided into slots and power allocation is done once per slot. To focus on the effect of power allocation and to provide tractable results we ignore subband and time slot allocation. Consequently, each user has its own, fixed subband of bandwidth W and inter-cell interference is ignored.

For each subband, the wireless channel from the base station to the user is assumed to experience quasi-static, time-selective fading with channel coefficient h_k . The fading process is assumed to be i.i.d. Rayleigh, which leads to exponentially distributed channel gains $|h_k|^2$. For each mobile user $k \in \mathcal{K}$, we can write the instantaneous SNR as

$$\gamma(p_k, |h_k|^2) = \frac{p_k |h_k|^2}{\sigma^2}, \quad (1)$$

with σ^2 the variance of the noise for user k . To perform power allocation, the base station requires a CSI report from every mobile user per time slot. In this work, we require that this feedback is equivalent to the channel gain $|h_k|^2$, $\forall k \in \mathcal{K}$.

B. Energy-Efficient Utility of the Cell

The base station performs power allocation, taking into account the energy efficiency associated with each of the mobile users. There are many common uses of the term energy efficiency. Here, we precisely refer to the notion introduced by Goodman and Mandayam in [6], i.e., the energy efficiency is the ratio between the effective throughput of the mobile user and the transmit power spent to attain this throughput. Contrary to [6] in which power control is considered for the uplink, we perform power allocation in the downlink. Hence, the transmit power considered in the energy efficiency is not the power from the mobile user but the power allocated to the mobile user by the base station. We denote $u_{BS}^{(k)}$ the energy-efficient utility associated with the k^{th} mobile user. It writes

$$u_{BS}^{(k)}(p_k) = \frac{Rf(\gamma_k(p_k, |h_k|^2))}{p_k}, \quad (2)$$

with R being a fixed transmission rate of the base station in bit/s, and f is an S-shaped function taking its values in $[0, 1]$ which represents the packet success rate during transmission. This function depends on the expected SNR of each mobile

user $k \in \mathcal{K}$. The power p_k is expressed in Watt, the unit of the utility is hence bit/Joule. Similarly to Belmega and Lasaulce work in [3], for the particular case of SISO channel, the efficiency function f is defined as

$$\begin{aligned} f(\gamma_k(p_k, |h_k|^2)) &= 1 - P_{\text{out}}(\gamma_k(p_k, |h_k|^2), a), \\ &= \exp\left(-\frac{a}{\gamma_k(p_k, |h_k|^2)}\right), \end{aligned} \quad (3)$$

with $a = 2^{R/W} - 1$ as the threshold under which the SNR causes an outage. It is important to notice that there is an approximation here, as P_{out} should depend on the expectation of the channel gain, and not on the instantaneous value of the channel gain. The justification of this approximation is that power allocation is performed every time slot. Consequently, instantaneous channel gains are required to take into account the variations of the characteristics of the channels from one time slot to another.

Fig. 1 illustrates the typical shapes of this energy-efficient utility per user, for three different values of the channel gain. It represents the energy-efficient utility associated with one mobile user (in bit/Joule) with respect to the power allocated to that user. Energy-efficient utility maximization does not necessarily correspond to rate maximization. Indeed, contrary to a rate maximization, it occurs that it is not always optimal to use all the power available to maximize energy efficiency. It is also interesting to note that for the same transmit power, the energy-efficient utility per user increases with the channel gain. Regarding individual optimal power, we can check that the energy-efficient utility per user can be maximized with a lower power when the channel gain is higher. For a cellular energy-efficient power allocation perspective, it means that it is more interesting to allocate power to mobile users with a good channel gain, as these mobile users require less power and offer a better energy-efficient utility.

The energy-efficient utility per cell is the sum of all the energy-efficient utilities per user, i.e.,

$$u_{BS}(\mathbf{p}) = \sum_{k \in \mathcal{K}} u_{BS}^{(k)}(p_k), \quad (4)$$

where the vector $\mathbf{p}$ contains all power values allocated by the base station.

Applying these utility functions for optimal power allocation requires to understand their properties. For each user, energy efficiency is expressed by (2). This function $u_{BS}^{(k)}(p_k)$ is continuous with respect to p_k . Consequently, the sum (4) is continuous with respect to $\mathbf{p}$. Nonetheless, such similarity cannot be found for another property of the functions: although each of the individual functions $u_{BS}^{(k)}(p_k)$ is quasiconcave, their sum is neither concave nor quasiconcave.

IV. ENERGY-EFFICIENT POWER ALLOCATION

A. Optimization Problem

The objective of the base station is to allocate at most P Watts among the different mobile users in order to maximize

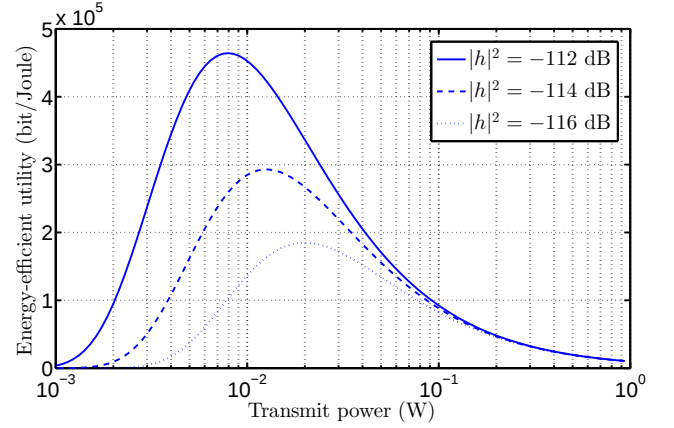


Fig. 1. Energy-efficient utility per user with respect to the allocated power for three different channel gains. Parameters are $a = 1$, and $\sigma^2 = 5 \times 10^{-14}$ W.

the energy-efficient utility per cell (4). The optimization problem is

$$\begin{aligned} &\underset{\mathbf{p} \in \mathbb{R}^K}{\text{maximize}} && \sum_{k \in \mathcal{K}} u_{BS}^{(k)}(p_k), \\ &\text{subject to} && \sum_{k \in \mathcal{K}} p_k \leq P, \\ &&& p_k \geq 0, \forall k \in \mathcal{K}. \end{aligned} \quad (5)$$

The problem is solved by choosing the optimization variables $p_1, \dots, p_K$ such that the $K + 1$ constraints hold. These constraints ensure that the optimization set is compact convex. The energy-efficient utility per cell u_{BS} being continuous with respect to $\mathbf{p}$, we know that u_{BS} has at least one maximum in the set delimited by the power constraints. Hence, there exists at least one optimal solution to this problem. But due to the utility u_{BS} being not quasiconcave, this optimization problem is hard to solve.

B. Optimality Conditions

Optimality conditions can be obtained by the study of the Lagrangian associated with (5). It writes

$$\mathcal{L}(\mathbf{p}, \lambda, \mu) = \sum_{k \in \mathcal{K}} u_{BS}^{(k)}(p_k) - \lambda \left(\sum_{k \in \mathcal{K}} p_k - P \right) + \sum_{k \in \mathcal{K}} \mu_k p_k. \quad (6)$$

The optimality conditions write $\forall k \in \mathcal{K}$

$$\frac{du_{BS}^{(k)}(\bar{p}_k)}{dp_k} = \lambda - \mu_k, \quad (7)$$

with $\bar{\mathbf{p}}$ an optimal power allocation. These optimality conditions lead to a partition of the set $\mathcal{K}$ in $\mathcal{K}'$ and $\mathcal{K}''$. The set $\mathcal{K}'$ is the subset of users for which power is allocated, leading to $\mu_k = 0$ and $\frac{du_{BS}^{(k)}(\bar{p}_k)}{dp_k} = \lambda$. The set $\mathcal{K}''$ is the subset of users with no power allocated ($\bar{p}_k = 0$), for which $\mu_k = \lambda$. In other words, in an optimal energy-efficient power allocation, there is a slope equality condition for a subset of the users, and the remaining users are given no power.

C. Algorithm Design Principle

Here we do not provide an optimal power allocation scheme to solve (5). Instead, we propose a simple suboptimal algorithm with a performance that is very close to the optimum. If the sum power constraint is not saturated, $\lambda = 0$, $\mathcal{K}' = \mathcal{K}$. Then, the proposed algorithm even provides the optimal allocation per cell.

The idea behind the proposed algorithm is that, without the sum power constraint, (5) can be divided into K simple quasiconcave optimization problems. For each of these K problems, the individual optimal power is known. As given in [3], $\forall k \in \mathcal{K}$, the power $p_k^*(|h_k|^2)$ that maximizes the energy-efficient utility of mobile user k is

$$p_k^*(|h_k|^2) = \arg \max_{p_k} \frac{f(\gamma_k(p_k, |h_k|^2))}{p_k}, \quad (8)$$

$$= \min \left\{ \frac{\sigma^2 a}{|h_k|^2}, P \right\}.$$

This power is called individual optimal power. If the sum of all these individual optimal powers is less than the sum power constraint P , expression (8) provides the solution of the optimization problem. If the sum exceeds P , the base station cannot allocate the individual optimal power p_k^* for each mobile user. Then it has to choose which mobile users to serve and which users to exclude in order to maximize the energy-efficient utility per cell while satisfying the sum power constraint. How to make this choice is justified by Lemma 1.

Lemma 1: $\forall i, j \in \mathcal{K}$, $i \neq j$, $|h_i|^2 \geq |h_j|^2$ is equivalent to

$$p_i^*(|h_i|^2) \leq p_j^*(|h_j|^2), \quad (9)$$

and

$$u_{BS}^{(i)}(p_i^*(|h_i|^2), |h_i|^2) \geq u_{BS}^{(j)}(p_j^*(|h_j|^2), |h_j|^2). \quad (10)$$

This means that a high individual optimal power offers a poor outcome in terms of energy-efficient utility, whereas a low individual optimal power offers a good outcome. Hence, from the base station perspective, it is more interesting to allocate the power budget for the users with the lower individual optimal power values. Note that contrary to water-filling, the individual optimal power associated with a mobile user with a good channel gain is less than the individual optimal power associated with a mobile user with a low channel gain. This is due to the fact that energy efficiency is maximized instead of sum capacity.

D. Algorithm

Based on the observations of Sec. IV-C, we propose the following algorithm to allocate power in the cell, assuming that the coefficients provided by (8) ($p_1^*(|h_1|^2), \dots, p_K^*(|h_K|^2)$) are in increasing order (to simplify the notation, we only write $(p_1^*, \dots, p_K^*)$ in what follows). The power allocated to mobile user k by the base station with that particular algorithm is denoted by $\tilde{p}_k$.

The algorithm is designed the following way. First, allocated power values are initialized to 0 (line 1). While allocated

Algorithm 1

Require: $(p_1^*, \dots, p_K^*)$, P .

```

1:  $(\tilde{p}_1, \dots, \tilde{p}_K) = (0, \dots, 0)$ 
2:  $i \leftarrow 1$ 
3: while  $\sum_{k=1}^K \tilde{p}_k < P$  do
4:   if  $p_i^* < p_{i+1}^*$  then
5:     if  $\sum_{k=1}^i p_k^* < P$  then
6:        $\tilde{p}_i \leftarrow p_i^*$ 
7:     else
8:        $\tilde{p}_i \leftarrow P - \sum_{k=1}^{i-1} \tilde{p}_k$ 
9:     end if
10:     $i \leftarrow i + 1$ 
11:  else
12:     $j \leftarrow i$ 
13:    while  $p_i^* == p_{i+1}^*$  do
14:       $i \leftarrow i + 1$ 
15:    end while
16:    if  $\sum_{k=1}^{j-1} \tilde{p}_k + \sum_{k=j}^i p_k^* < P$  then
17:       $\forall k \in \{j, \dots, i\}$ ,  $\tilde{p}_k \leftarrow p_k^*$ 
18:    else
19:       $\forall k \in \{j, \dots, i\}$ ,  $\tilde{p}_k \leftarrow \frac{P - \sum_{k=1}^{j-1} \tilde{p}_k}{i - j + 1}$ 
20:    end if
21:     $i \leftarrow i + 1$ 
22:  end if
23: end while
24: return  $(\tilde{p}_1, \dots, \tilde{p}_K)$ 

```

power does not exceed P , the allocation of power continues (line 3). If the individual optimal power of mobile user i is strictly less than the individual optimal power of the next mobile user (line 4),

- if the sum of the individual optimal power of mobile user i to the power values already allocated does not exceed P , this individual optimal power is allocated (line 6),
- if the sum of the individual optimal power of mobile user i to the power values already allocated exceeds P , only the remaining power is allocated (line 8).

If the individual optimal power of mobile user i is equal to the individual optimal power of the next mobile user (line 11), the number of successive equal individual optimal power values is counted (line 13),

- if the sum of these optimal power values to the previously allocated power values does not exceed the constraint, these optimal power values are allocated (line 17),
- if the sum of these optimal power values to the previously allocated power values exceeds the constraint, the remaining power is fairly shared (line 19).

Note that the algorithm takes into account the case in which several individual optimal power values are equal. Such an event occurs with almost null probability when the channel gains are considered to be continuous and follow an exponential distribution law.

For the same sum power constraint, Fig. 2 illustrates the performance gap between the outcome of the proposed

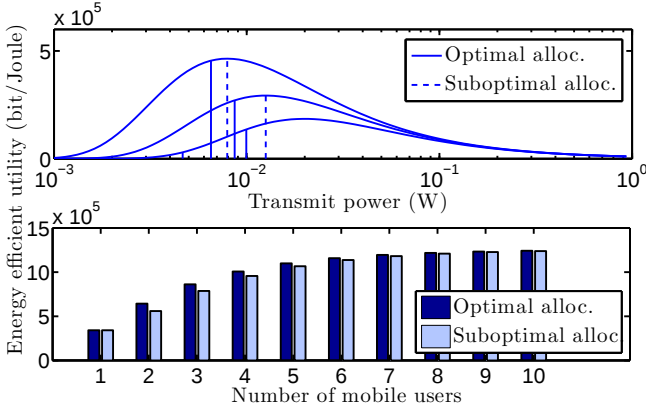


Fig. 2. Energy-efficient performance comparison of the optimal power allocation and the proposed allocation.

algorithm and the optimal power allocation. The figure on top compares the optimal power allocation and the proposed allocation in a cell with three mobile users. For each user, the associated energy-efficient utility is represented, and the power allocations are given for the optimal and suboptimal cases. While both allocations saturate the sum power constraint, the optimal allocation verifies the slopes equality, in accordance with Sec. IV-B, whereas the proposed allocation gives the individual optimal power values to the two mobile users with the best channel gains, and only the remaining power for the third mobile user. The figure below represents the energy efficiency in the cellular cell for the two allocations, for various amount of mobile users in the cell. For each number of mobile users, the channel gains and the sum power constraint are chosen such that the two allocations differ the most. It can be considered as a worst-case scenario for the proposed algorithm.

V. SELFISH CHANNEL STATE REPORTING

A. Definition of a game

This section focuses on the behavior of the mobile users. The main difference with the base station behavior is that mobile users are not concerned with energy efficiency in downlink as they do not provide transmit power themselves. Instead, they are only concerned about their individual SNR. From an operator perspective, the energy consumed at the base station dominates the one needed by the mobile users. Therefore, it makes sense to consider that only the base station is concerned by energy efficiency while the mobile users are only concerned about their SNR. For each mobile user, this SNR is proportional to the power allocated by the base station. Sec. IV shows that this power depends on the channel gain the considered user reports to the base station, but also on the channel gains reported by of all the other users. Considering that each mobile user prefers to have a high allocated power, it can try to twist the channel gain it reports to the base station in order to get higher allocated power. Hence, we assume that the mobile users have the freedom of sending whatever channel

gain feedback they want to the base station, and we use game theory to study what are the outcomes of such a scenario. For each mobile user $k \in \mathcal{K}$, two values of the channel gain are important:

- its actual channel gain $|h_k|^2$, as the actual performance of the mobile user depends on this gain;
- the value it reports to the base station that is denoted by $g_k \in [0, G]$, with G the maximum gain a mobile user can report. Power is allocated to the mobile user based on this value.

Both terms appear in the utility of each mobile user, which is the SNR after power allocation by the base station. $\forall k \in \mathcal{K}$,

$$u_k(g_k, g_{-k}) = \frac{\tilde{p}_k(g_k, g_{-k})|h_k|^2}{\sigma^2}. \quad (11)$$

With a slight abuse of notation (g_k, g_{-k}) emphasizes the feedback of mobile user k compared to the CSI reports of all the other mobile users. We can now properly define the game under study.

Definition 1: The channel feedback game is defined by the triplet $\mathcal{G} = (\mathcal{K}, \mathbb{R}, \{u_k\}_{k \in \mathcal{K}})$ in which

- $\mathcal{K}$ is the set of players of the game, which represents the mobile users.
- $\mathcal{A}_k = [0, G]$ is the set of actions for player k . In this game, an action $g_k \in \mathcal{A}_k$ is the report to the base station. The set of actions profiles is denoted by $\mathcal{A} = \times_{k=1}^K \mathcal{A}_k$, and the cardinal product of the actions sets of all players except player k is denoted by $\mathcal{A}_{-k} = \times_{i \neq k} \mathcal{A}_i$.
- The utility of player k is

$$u_k(g_k, g_{-k}) = \frac{\tilde{p}_k(g_k, g_{-k})|h_k|^2}{\sigma^2}.$$

It is its SNR given the power allocated by the base station knowing all the reported channel gains.

B. Characterization of the Nash Equilibrium

Generally, an important concept to study the outcome of a game is the Nash equilibrium.

Definition 2: In the game $\mathcal{G}$, a Nash equilibrium is an action profile $(g_1^*, \dots, g_K^*)$ such that $\forall k \in \mathcal{K}$

$$\forall g_k \neq g_k^*, u_k(g_k, g_{-k}^*) \leq u_k(g_k^*, g_{-k}^*). \quad (12)$$

In other words, it is an action profile from which no player has interest to deviate unilaterally. It is a selfish equilibrium. Interestingly, in the game $\mathcal{G}$, we can prove that there exists at least a pure Nash equilibrium.

Proposition 1: There is a unique Nash equilibrium in the game $\mathcal{G}$. It is $\mathbf{g}^* = (g_1^*, g_2^*, \dots, g_K^*)$ with $g^* = \frac{K\alpha\sigma^2}{P}$ such that

$$\arg \max_p \frac{f(\gamma(p, g^*))}{p} = \frac{P}{K}. \quad (13)$$

At this equilibrium, power is uniformly allocated among the mobile users and the SNR of player $k \in \mathcal{K}$ is

$$u_k(\mathbf{g}^*) = \frac{P|h_k|^2}{K\sigma^2}. \quad (14)$$

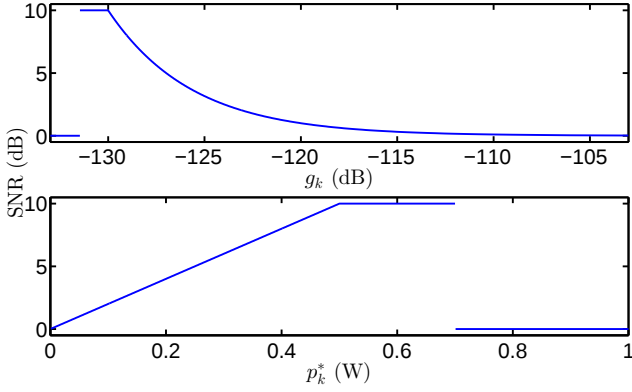


Fig. 3. SNR of player k with respect to the reported channel gain and its associated individual optimal power.

It is interesting to note that this equilibrium depends only on the number of mobile users in the cell, and the sum power constraint. The actual channels gains are absolutely not involved in the expression of this equilibrium.

Proof: First, we show that the action profile $(g^*, g^*, \dots, g^*)$ is a Nash equilibrium. Assume that player k deviates from the action g^* , it chooses an action $g_k \neq g^*$. The action profile is then (g_k, g_{-k}^*) .

- If $g_k < g^*$ then $p_k^* > \frac{P}{K}$, and

$$\sum_{i \in \mathcal{K}} p_i^* = (K-1) \frac{P}{K} + p_k^* > P. \quad (15)$$

Hence, according to the allocation algorithm, $\tilde{p}_k$ is set to $P - (K-1) \frac{P}{K} = \frac{P}{K}$, i.e., the SNR of player k remains unchanged.

- If $g_k > g^*$ then $p_k^* < \frac{P}{K}$, then

$$\sum_{i \in \mathcal{K}} p_i^* = (K-1) \frac{P}{K} + p_k^* < P. \quad (16)$$

In this case, all the players are given the individual optimal power. Hence player k gets $\tilde{p}_k = p_k^* < \frac{P}{K}$, and its SNR decreases.

We have proven that $\forall k \in \mathcal{K}$, player k has no interest in deviating from the power profile $(g^*, g^*, \dots, g^*)$, hence it is a Nash equilibrium. We now prove that this equilibrium is unique.

- If the sum of individual optimal power values is below the sum power constraint, i.e.,

$$\sum_{k \in \mathcal{K}} p_k^* < P, \quad (17)$$

then if one player k decreases g_k , it gets a higher $\tilde{p}_k$. Hence there can be no equilibrium when the maximum total power constraint is not active.

- If the sum of individual optimal power values is higher than the sum power constraint, and $\exists k \in \mathcal{K}$ such that $\tilde{p}_k = 0$. Then player k can increase its report g_k in order

to have some power allocated. Hence player k has interest in deviating.

- If the sum of individual optimal power values is higher than the sum power constraint, if no player gets zero power allocated, and $\exists (i, j) \in \mathcal{K}^2$ such that $\tilde{p}_i < \tilde{p}_j$, then player i can slightly decrease g_i in order to get more power from the base station.
- If the sum of individual optimal power value is higher than the sum power constraint, and $\forall (i, j) \in \mathcal{K}^2$, $\tilde{p}_i = \tilde{p}_j$, then $\forall i \in \mathcal{K}$, $\tilde{p}_i = \frac{P}{K}$. If $\exists i \in \mathcal{K}$ such that $p_i^* > \frac{P}{K}$, then any other player $j \neq i$ can report g_j such that $p_j^* \in [\frac{P}{K}, p_i^*]$. In that case, player j receives more power from the base station.

We have proven that there cannot be any other equilibrium than the case for which the sum of individual optimal power values saturates the sum power constraint, and $\forall i \in \mathcal{K}$, $\tilde{p}_i = p_i^* = \frac{P}{K}$. ■

Given that at this equilibrium, mobile users do not report their actual channel gains, there are two energy-efficient utilities per cell of interest. First, the energy-efficient utility per cell the base station is convinced to have, given the equilibrium reports of the mobile users. This is not the actual value of the energy-efficient utility per cell. It writes

$$\begin{aligned} u_{BS}(\frac{P}{K}, g^*) &= \sum_{k \in \mathcal{K}} \frac{K}{P} \exp(-\frac{aK\sigma^2}{Pg^*}), \\ &= \frac{K^2}{P} \exp(-1). \end{aligned} \quad (18)$$

Once again, we can check that this utility depends only on the total number of mobile users in the cell and the power constraint. The second energy-efficient utility per cell of interest is of course the actual energy efficiency of the cell, which takes into account the true channel gains. This utility writes

$$u_{BS}(\frac{P}{K}, |h_k|^2) = \frac{K}{P} \sum_{k \in \mathcal{K}} \exp(-\frac{g^*}{|h_k|^2}). \quad (19)$$

Hence, the ratio between the actual energy-efficient performance of the cell and the believed performance is

$$\frac{u_{BS}(\frac{P}{K}, |h_k|^2)}{u_{BS}(\frac{P}{K}, g^*)} = \frac{1}{K} \sum_{k \in \mathcal{K}} \exp(1 - \frac{g^*}{|h_k|^2}). \quad (20)$$

VI. NUMERICAL RESULTS

In Fig. 4, a scenario with two mobile users is considered. For both of them, the channel gains are set to -112 dB, and we study how the energy-efficient utility per cell varies for all the possible combinations of allocated power values. Hence, the energy-efficient utility per cell is represented with respect to SNR of each of the mobile users. The constraint on the total transmit power is $P = 1W$. On the one hand, we can check that the energy efficiency of the cell is maximized when mobile users report their actual channel gain. But at this point, the sum power constraint is not saturated. Hence, it is not Pareto-optimal for the mobile users. On the other hand, when the mobile users report the equilibrium channel gains, a Pareto

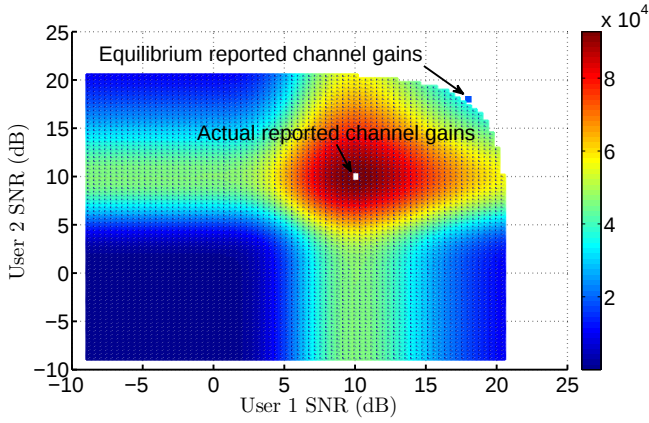


Fig. 4. Energy-efficient utility of the cell with respect to the SNR of users 1 and 2. The color bar represents the cell energy-efficient utility in bit/Joule. Parameters are $\alpha = 10$, $|h_1|^2 = |h_2|^2 = -112$ dB, $P = 1$ W, and $\sigma^2 = 5 \times 10^{-14}$ W.

optimal point is reached. But in this configuration, the energy efficiency of the cell is not a maximum.

In Fig. 5, the energy-efficient utility of the cell is represented as a function of the number of mobile users in the cell. The variance of the noise, σ^2 is set to 5×10^{-14} W. The parameter α is set to 6. Channel gains are assumed to be exponentially distributed. For each number of mobile users, 10^4 realizations are computed, and the presented results are the mean over these realizations. Similarly to Fig. 4, we compare the case in which mobile users actually report their channel gains, and the case in which they twist their reports in order to maximize their own utilities. In addition, two power constraints are considered $P = 0.1$ W, and $P = 1$ W. With no surprise, for both power constraints, the energy-efficient utility of the cell is higher when mobile users report their actual channel gains. This utility increases with the number of mobile users simply because it is a sum over the mobile users. Interestingly, when mobile users report their actual channel gains, there is no significant gap between the two power constraints. There are two explanations for this phenomenon. First, when the sum power constraint is not saturated, the power allocation is exactly the same whatever the power constraint. Second, we recall that mobile users with bad channel gains are not interesting in terms of energy efficiency as their individual optimal power is high and the associated energy efficiency is low. Hence, allocating power or not to these mobile users does not make a significant difference in terms of energy efficiency. Another interesting phenomenon is when mobile users report equilibrium channel gains, the energy-efficient utility of the cell is worse for $P = 1$ W than for $P = 0.1$ W. We recall that at the equilibrium, the sum power constraint is saturated. Hence, there is more power allocated to the mobile users in the case with $P = 1$ W. But after a certain threshold, increasing power decreases the energy efficiency of the overall system.

In Fig. 6, the mean SNR of a mobile user is represented with regard to the number of mobile users for the same scenario as Fig. 5. The SNR of the mobile user is given in dB, and

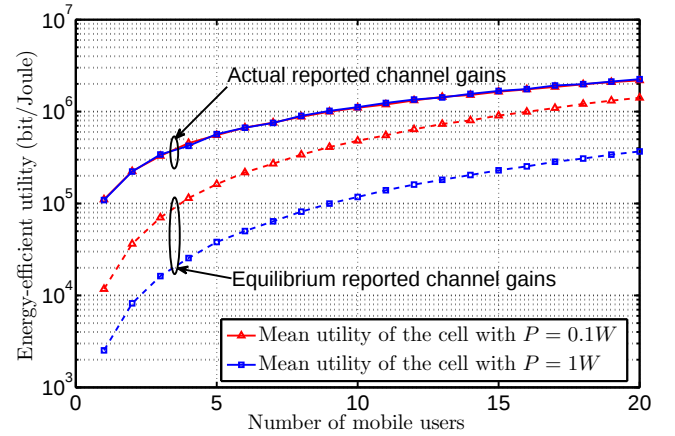


Fig. 5. Energy-efficient utility of the cell with respect to the number of mobile users.

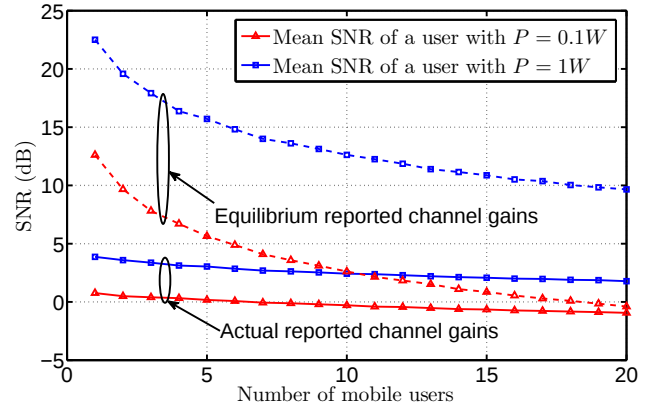


Fig. 6. Mean SNR of one mobile user with respect to the number of mobile users.

for both power constraints, this SNR is higher when mobile users report equilibrium channel gains. As the SNR of the mobile user at the equilibrium is proportional to the sum power constraint, we can check there is a gain between the two sum power constraints. When mobile users report their actual channel gains, their SNR is lower. But a higher sum power constraint allows the base station to serve more users. Thus the mean SNR of the mobile users with actual channel gains increases with the sum power constraint.

VII. CONCLUSION

In this paper we studied the conflict of interest between (i) mobile users that selfishly choose their CSI reports to increase their SNR and (ii) a base station that aims to maximize the cell's energy efficiency by power allocation. We formulated power allocation as a non-convex optimization problem, stated its optimality conditions, and derived a low complexity algorithm. Formulating selfish CSI reporting as a game allowed us to prove the existence of a unique equilibrium from which no user has interest to deviate. Consequently, we obtained a stable system for which we derived energy efficiency with and

without true CSI in closed form.

We illustrate this powerful theoretical framework by numerical results. Naturally, these results show that selfish CSI reports allow mobile users to increase their own SNR while the cell's energy efficiency decreases. Interestingly, this degradation diminishes for smaller sum power constraints and for a larger number of users. Consequently, small cells are more robust to selfish CSI reports than cells with large power constraints and few users.

As future work, we aim to extend this framework by an optimal power allocation algorithm, energy-efficient utilities not only for the base station but for the mobile users, as well as by a practical, discrete set of CSI values.

REFERENCES

- [1] 3GPP. E-UTRA physical layer procedures (release 9). Technical Specification TS 36.213 V9.0.1, 3GPP, 2009.
- [2] 3GPP. E-UTRA and E-UTRAN overall description (stage 2, release 9). Technical Specification TS 36.300 V9.2.0, 3GPP, 2010.
- [3] E. V. Belmega and S. Lasaulce. Energy-efficient precoding for multiple-antenna terminals. *IEEE Trans. on Signal Processing*, 59(1):329–340, January 2011.
- [4] S. Buzzi and D. Saturnino. A game-theoretic approach to energy-efficient power control and receiver design in cognitive CDMA wireless networks. *Journal of Selected Topics in Signal Processing*, 5(1):137–150, 2011.
- [5] G. J. Foschini and Z. Miljanic. A simple distributed autonomous power control algorithm and its convergence. *IEEE Trans. on Vehicular Technology*, 42(4):641–646, 1993.
- [6] D. J. Goodman and N. B. Mandayam. Power control for wireless data. *IEEE Person. Comm.*, 7:48–54, 2000.
- [7] V. Kavitha, E. Altman, R. E. Azouzi, and R. Sundaresan. Opportunistic scheduling in cellular systems in the presence of noncooperative mobiles. *IEEE Transactions on Information Theory*, 58(3):1757–1773, 2012.
- [8] V. Kavitha, E. Altman, R. El-Azouzi, and R. Sundaresan. Fair scheduling in cellular systems in the presence of noncooperative mobiles. In *Proceedings of the 29th conference on Information communications, INFOCOM'10*, pages 2615–2623, Piscataway, NJ, USA, 2010. IEEE Press.
- [9] I. Koutsopoulos, S. Stanczak, and A. Feistel. Transmit rate control for energy-efficient estimation in wireless sensor networks. In *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, June 2010.
- [10] F. Meshkati, M. Chiang, H. V. Poor, and S. C. Schwartz. A game-theoretic approach to energy-efficient power control in multi-carrier CDMA systems. *IEEE Journal on Selected Areas in Communications*, 24(6):1115–1129, 2006.
- [11] J. Sobel. Signaling games. In *Encyclopedia of Complexity and Systems Science*, pages 8125–8139. 2009.
- [12] C. W. Tan, D. Palomar, and M. Chiang. Energy-robustness tradeoff in cellular network power control. *IEEE/ACM Trans. Networking*, 17(3):912–925, May 2009.
- [13] R. D. Yates. A framework for uplink power control in cellular radio systems. *IEEE J. Sel. Areas Comm.*, 13(7):1341–1347, Sep. 1995.

A.2 IEEE-DSP-2011

F. Mériaux, M. Le Treust, S. Lasaulce et M. Kieffer. *A Stochastic Game Formulation of Energy-Efficient Power Control : Equilibrium Utilities and Practical Strategies*. In IEEE International Conference on Digital Signal Processing (DSP), pages 665-671, 2011.

A STOCHASTIC GAME FORMULATION OF ENERGY-EFFICIENT POWER CONTROL: EQUILIBRIUM UTILITIES AND PRACTICAL STRATEGIES

François Mériaux, Maël Le Treust, Samson Lasaulce, Michel Kieffer

L2S - CNRS - SUPELEC - Univ Paris-Sud
3 rue Joliot-Curie
F-91192 Gif-sur-Yvette, France
{meriaux,letreust,lasaulce,kieffer}@lss.supelec.fr

ABSTRACT

Frequency non-selective time-selective multiple access channels in which transmitters can freely choose their power control policy are considered. The individual objective of the transmitters is to maximize their averaged energy-efficiency. For this purpose, a transmitter has to choose a power control policy that is, a sequence of power levels adapted to the channel variations. This problem can be formulated as a stochastic game with discounting for which there exists a theorem characterizing all the equilibrium utilities (equilibrium utility region). As in its general formulation, this theorem relies on global channel state information (CSI), it is shown that some points of the utility region can be reached with individual CSI. Interestingly, time-sharing based solutions, which are usually considered for centralized policies, appear to be part of the equilibrium solutions. This analysis is illustrated by numerical results providing further insights to the problem under investigation.

Index Terms— Distributed power control, Stochastic games, Folk theorem, Nash equilibrium, Game Theory.

1. INTRODUCTION

In the past decade, new types of wireless networks have appeared. Just to name a few, when inspecting the wireless literature, we find ad hoc networks, networks of wireless devices operating in unlicensed bands, wireless networks with cognitive radios, small cells based cellular networks. The decentralized nature of such networks make game-theoretic analyses relevant [1]. Interestingly, game theory offers a large set of tools and concepts to better understand important problems such as resources allocation and power control in distributed networks. By distributed, it is meant that the allocation or/and control policy is left to the terminal itself. In this paper, the problem under consideration is precisely the one of distributed power control. The assumed network model is a multiple access channel (MAC), which, by definition, includes several transmitters and one common receiver. A brief overview of previous works about power allocation for MACs is presented in [2]. In our framework, based on a certain knowledge which includes his individual channel state information, each transmitter has to tune his power level for each data block. In particular, it can ignore some specified centralized policies. The assumed performance criterion is the one introduced by [3] that is to say, that transmitters aim at maximizing the energy efficiency associated with the transmit radio-frequency signal, which is measured as a number of correctly decoded information bits per Joule consumed at the transmitter. The assumed channel model is the same as [3] that is, links are assumed to be fre-

quency non-selective but the channel gains can vary from block to block. The authors of [3] have formalized the energy-efficient power control problem as a one-shot/static non-cooperative game: on each block, the players (the transmitters) play a one-shot game by choosing their action/move (their power level) in order to maximize their utility (their individual energy-efficiency). The main drawback of such a formulation is that it leads to an outcome (Nash equilibrium) which is not efficient. Indeed, it can be checked that, for each block, there exists a vector of power levels (an action profile) which allows all the players to have better utilities than those obtained at the Nash equilibrium; the latter is said to be Pareto-dominated.

Motivated by the existence of power profiles which Pareto-dominates the one-shot game Nash equilibrium solution, several authors proposed solutions which are both efficient and compatible with the framework of decentralized decisions. For instance, pricing is proposed in [4], a Stackelberg formulation is introduced in [5], and a repeated game formulation is exploited in [6]. The framework adopted in this paper is a more general framework than the one chosen in [6]. Indeed, although the repeated game model in [6] takes into account the fact that transmitters interact several/many times, the work in [6] has an important weakness: this is the need for a normalized game which does not depend on the realization channels. By definition (see e.g., [7]), a repeated game consists in repeating the same one-shot game. The consequence of such a modeling choice is a loss in terms of optimality in terms of expected utilities (averaged over the channel realizations). The main purpose of this paper is to propose a more general dynamic game model namely, a stochastic game. Based on this choice, the contributions of this paper are essentially as follows: (i) after providing the signal model (Sec. 2) and reviewing the one-shot game model (Sec. 3) we present the stochastic game model used and show how to exploit the recent game-theoretic results by [8][9] to obtain a Folk theorem for the power control game under investigation: this Folk theorem allows one to fully characterize equilibrium utilities when public signals are available to the transmitters and channel stats are quantized (Sec. 4); (ii) some simple equilibrium power control strategies relying on individual channel state information and possibly a recommendation of the receiver are proposed (Sec. 5); (iii) a numerical study is conducted to assess the performance of the proposed strategies and give more insights on tuning the relevant parameters of the problem (Sec. 6).

2. SIGNAL MODEL

We consider a decentralized MAC with $K \geq 1$ transmitters. The network is said to be decentralized as the receiver (e.g., a base

station) does not dictate to the transmitters (e.g., mobile stations) their power control policy. Rather, all the transmitters choose their policy by themselves and want to selfishly maximize their energy-efficiency; in particular, they can ignore some specified centralized policies. We assume that the users transmit their data over quasi-static channels, at the same time and frequency band and without any beamforming [10]. Note that a block is defined as a sequence of $M \geq 1$ consecutive symbols which contains a training sequence: a specific symbols sequence used to estimate the channel (or other related quantities) associated with a given block. A block has therefore a duration less than the channel coherence time. The signal model used corresponds to the information-theoretic channel model used for studying MAC, see [2] for more comments on the multiple access technique involved. This model is both simple to be presented and captures the different aspects of the problem. It can be readily applied to specific systems such as CDMA systems [3][5] or multi-carrier CDMA systems [11]. The equivalent baseband signal received by the base station can be written as

$$y(n) = \sum_{i=1}^K g_i(n)x_i(n) + z(n) \quad (1)$$

where $i \in \mathcal{K}$, $\mathcal{K} = \{1, \dots, K\}$, $x_i(n)$ represents the symbol transmitted by transmitter i at time n , $\mathbb{E}|x_i|^2 = p_i$, the noise z is assumed to be distributed according to a zero-mean Gaussian random variable with variance σ^2 and each channel gain g_i varies over time but is assumed to be constant over each block. For each transmitter i , the channel gain modulus is assumed to lie in a compact set $|g_i|^2 \in [\eta_i^{\min}, \eta_i^{\max}]$. This assumption models the finite receiver sensitivity and the existence of a minimum distance between the transmitter and receiver. At last, the receiver is assumed to implement single-user decoding.

At a given instant, the SINR at receiver $i \in \mathcal{K}$ writes as:

$$\text{SINR}_i = \frac{p_i \eta_i}{\sum_{j \neq i} p_j \eta_j + \sigma^2} \quad (2)$$

where p_i is the power level for transmitter i and $\eta_i = |g_i|^2$.

3. ONE-SHOT POWER CONTROL GAME

In this section, the one-shot game model of [3] is reviewed since it both allows one to build the stochastic game model of Sec. 4 and serves as a reference for performance comparison. A useful (non-equilibrium) operating point in this game is also defined, as a basis for the proposed control strategies in the stochastic game model.

Definition 1 (One-shot power control game) *The strategic form of the one-shot power control game is a triplet $\mathcal{G} = (\mathcal{K}, \{\mathcal{A}_i\}_{i \in \mathcal{K}}, \{u_i\}_{i \in \mathcal{K}})$ where $\mathcal{K}$ is the set of players, $\mathcal{A}_1, \dots, \mathcal{A}_K$ are the corresponding sets of actions, $\mathcal{A}_i = [0, P_i^{\max}]$, $P_i^{\max}$ is the maximum transmit power for player i , and $u_1, \dots, u_K$ are the utilities of the different players which are defined by:*

$$u_i(p_1, \dots, p_K) = \frac{R_i f(\text{SINR}_i)}{p_i} \text{ [bit/J]}. \quad (3)$$

We denote by R_i the transmission information rate (in bps) for user i and f an efficiency function representing the block success rate. The numerator of the utility is thus the rate of bits successfully received at the base station. f is assumed to be sigmoidal and

identical for all the users; the sigmoidness assumption is a reasonable assumption, which is well justified in [12][13]. Recently, [14] has shown that this assumption is also justified from an information-theoretic standpoint.

In this game with complete information ($\mathcal{G}$ is known to every player) and rational players (every player does the best for himself and knows the others do so and so on), a major game solution concept is the Nash equilibrium (i.e. a point from which no player has interest in unilaterally deviating). When it exists, the non-saturated¹ Nash equilibrium of this game can be obtained by setting $\frac{\partial u_i}{\partial p_i}$ to zero $\forall i \in \mathcal{K}$ which gives an equivalent condition on the SINR: the best SINR in terms of energy-efficiency for transmitter i has to be a solution of $x f'(x) - f(x) = 0$ (this solution is independent of the player index since a common efficiency function is assumed, see [13] for more details). This leads to:

$$\forall i \in \{1, \dots, K\}, p_i^* = \frac{\sigma^2}{\eta_i} \frac{\beta^*}{1 - (K-1)\beta^*} \quad (4)$$

where β^* is the unique solution of the equation $x f'(x) - f(x) = 0$. An important property of the Nash equilibrium given by (4) is that transmitters only need to know their individual channel gain η_i to play their equilibrium strategy. Another interesting property is that the product $p_i^* \eta_i$ (instantaneous received power) is constant for all the players.

Interestingly, the authors of [6] propose to study the power profile obtained when imposing the received signals to have the same instantaneous power. The idea is to solve $\frac{\partial u_i}{\partial p_i}(p) = 0$ under the aforementioned constraint. This leads to the following system:

$$\forall (i, j) \in \mathcal{K}^2, \begin{cases} p_i \eta_i = p_j \eta_j \\ \frac{\partial u_i}{\partial p_i}(p) = 0 \end{cases} \quad (5)$$

The unique solution of (5), called *Operating point* can be checked to be:

$$\forall i \in \mathcal{K}, \bar{p}_i = \frac{\sigma^2}{\eta_i} \frac{\tilde{\gamma}_K}{1 - (K-1)\tilde{\gamma}_K} \quad (6)$$

where $\tilde{\gamma}_K$ is the unique solution of $x[1 - (K-1) \cdot x]f'(x) - f(x) = 0$.

The difference between this operating point and the Nash equilibrium can be explained by the fact that $\frac{\partial u_i}{\partial p_i}(p) = \frac{\partial u_i}{\partial p_i}(p_i)$ when adding the instantaneous power equality constraint. This operating point can be proved to always Pareto-dominate the Nash equilibrium and reach the Pareto frontier for each channel realization; additionally, only individual CSI is needed to operate at the corresponding power levels. This point will serve as a basis of the power control strategies proposed in Sec. 5.

4. STOCHASTIC POWER CONTROL GAME

With the one-shot non-cooperative game model, transmitters are assumed to play once for each block and independently from block to block. The goal here is to take into account the fact that transmitters generally interact over several blocks, which is likely to change their behavior w.r.t. the one-shot interaction model even if

¹By using the term “non-saturated Nash equilibrium” we mean that the maximum transmit power for each user, denoted by $P_i^{\max}$, is assumed to be sufficiently high not to be reached at the equilibrium i.e. each user maximizes his energy-efficiency for a value less than $P_i^{\max}$ (see [5] for more details about the saturated case).

they are always assumed to be selfish. As channel gains are time-varying, the most natural model is the one of stochastic games [15]. In such a model, important differences w.r.t. the one-shot game model are that averaged utilities are considered, the channel state $\underline{\eta}(t) = (\eta_1(t), \dots, \eta_K(t)) \in \Gamma$ and $\Gamma = \Gamma_1 \times \Gamma_2 \times \dots \times \Gamma_K$ may vary according to a certain evolution law (the i.i.d. block fading case is the most simple of them), and the state can depend on the played actions (in conventional wireless settings this is however not the case). Stochastic game stages correspond to instants at which players can choose their actions. From one stage to another, the channel state $\underline{\eta}(t)$ is assumed to be discrete (e.g., resulting from quantization effects) and stochastically varies according to the transition probability distribution π . This distribution is said to be an irreducible transition probability if for any pair of channel states $\underline{\eta}$ and $\underline{\eta}'$ we have $\pi(\underline{\eta}'|\underline{\eta}) > 0$. For example, this irreducibility condition is met for i.i.d channels. The second important assumption we do to obtain a Folk theorem is to assume that a public signal is available to all the transmitters. Two special cases of interest are: (a) Every transmitter knows the power of the received signal that is, $\sum_{i=1}^K \eta_i p_i + \sigma^2$ is known; (b) Every transmitter has global CSI and perfectly observes the action profiles that is, $(\underline{\eta}, \underline{p})$ is known.

The game Course. The game starts at the first stage with a channel state $\underline{\eta}(1)$ known by the players. The transmitters simultaneously choose their power level $\underline{p}(1) = (p_1(1), \dots, p_K(1))$ and are assumed to receive a public signal, denoted $\theta \in \Theta$. At stage t , the channel states $\underline{\eta}(t)$ are drawn from the transition probability $\pi(\cdot|\underline{\eta}(t-1)) \in \Delta(\Gamma)$ and the players observe the public signal

$$\phi : \begin{cases} \Gamma \times \mathcal{A} & \rightarrow \Theta \\ (\underline{\eta}, \underline{p}) & \mapsto \theta \end{cases} \quad (7)$$

where $\mathcal{A} = \mathcal{A}_1 \times \dots \times \mathcal{A}_K$. The sequence of past signals $\underline{h}(t) = (\theta(1), \dots, \theta(t-1), \underline{\eta}(t))$ is the common history of the players.

Definition 2 (Players' strategies) A strategy for player $i \in \mathcal{K}$ is a sequence of functions $(\tau_{i,t})_{t \geq 1}$ with

$$\tau_{i,t} : \begin{cases} \Theta^t & \rightarrow \mathcal{A}_i \\ \underline{h}_t & \mapsto p_i(t). \end{cases} \quad (8)$$

The strategy of player i will therefore be denoted by τ_i while the vector of strategies $\underline{\tau} = (\tau_1, \dots, \tau_K)$ will be referred to a joint strategy. A joint strategy $\underline{\tau}$ induces in a natural way a unique sequence of action plans $(\underline{p}(t))_{t \geq 1}$ and a unique sequence of public signals $(\theta(t))_{t \geq 1}$. The averaged utility for player i is defined as follows.

Definition 3 (Players' utilities) Let $\underline{\tau} = (\tau_1, \dots, \tau_K)$ be a joint strategy. The utility for player $i \in \mathcal{K}$ if the initial channel state is $\underline{\eta}(1)$, is defined by:

$$v_i(\underline{\tau}, \underline{\eta}(1)) = \sum_{t \geq 1} \lambda(1 - \lambda)^{t-1} \mathbb{E}_{\underline{\tau}, \pi} [u_i(\underline{p}(t), \underline{\eta}(t)) | \underline{\eta}(1)] \quad (9)$$

where $(\underline{p}(t))_{t \geq 1}$ is the sequence of power profiles induced by the joint strategy $\underline{\tau}$.

The parameter λ is the discount factor, which can model various effects such as the probability the game stops, the fact that players evaluate short-term and long-term benefits differently, etc. We now present the definition of a stochastic game.

Definition 4 (Stochastic game) A stochastic game with public monitoring is defined as a tuple:

$$\mathcal{G} = (\mathcal{K}, (\mathcal{T})_{i \in \mathcal{K}}, (v_i)_{i \in \mathcal{K}}, (\Gamma_i)_{i \in \mathcal{K}}, \pi, \Theta, \phi), \quad (10)$$

where $\mathcal{T}_i$ is the set of strategies for player i and v_i his long-term utility function.

Equilibrium concept. Let us define the Nash equilibrium of a stochastic game starting with the channel state $\underline{\eta}(1)$.

Definition 5 (Equilibrium Strategies) A strategy $\underline{\tau}$ supports an equilibrium of the stochastic game with initial channel state $\underline{\eta}(1)$ if

$$\forall i \in \mathcal{K}, \forall \tau'_i, v_i(\underline{\tau}, \underline{\eta}(1)) \geq v_i(\tau'_i, \underline{\tau}_{-i}, \underline{\eta}(1)) \quad (11)$$

where $-i$ is the standard notation to refer to the set $\mathcal{K} \setminus \{i\}$, $\underline{\tau}_{-i} = (\tau_1, \dots, \tau_{i-1}, \tau_{i+1}, \dots, \tau_K)$. Denote $E_\lambda(\underline{\eta}(1))$ the set of equilibrium utilities with initial channel state $\underline{\eta}(1)$.

We now characterize the equilibrium utility region for the stochastic game under study. For this, define the min-max level $\tilde{v}_i$ of player $i \in \mathcal{K}$ as the most severe punishment level for player $i \in \mathcal{K}$. The feasible utility region is denoted by $F_\lambda(\underline{\eta}(1))$. The result of Dutta [16] states that if the transition probability π is irreducible, then the min-max levels, the feasible utility region and the equilibrium utility region are independent of the initial state $\underline{\eta}(1)$.

$$\lim_{\lambda \rightarrow 0} \min_{\tau_{-i}} \max_{\tau_i} \tilde{v}_i(\tau_i, \tau_{-i}, \underline{\eta}(1)) = \tilde{v}_i, \forall \underline{\eta}(1) \quad (12)$$

$$\lim_{\lambda \rightarrow 0} F_\lambda(\underline{\eta}(1)) = F, \forall \underline{\eta}(1) \quad (13)$$

$$\lim_{\lambda \rightarrow 0} E_\lambda(\underline{\eta}(1)) = E, \forall \underline{\eta}(1) \quad (14)$$

As already mentioned, irreducibility is verified under the common assumption of i.i.d. channels. As a consequence, the equilibrium utility region is independent of the initial channel state for a stochastic game with a public signal.

Definition 6 We define the set of asymptotic feasible and individually rational payoff by:

$$F^* = \{x \in F | x_i \geq \tilde{v}_i, \forall i \in \mathcal{K}\} \quad (15)$$

Theorem 7 Suppose that the players see the same public signal. Then, for each utility vector $\underline{u} \in F^*$, there exists a λ_0 such that for all $\lambda < \lambda_0$, there exists a perfect public equilibrium strategy in the stochastic power control game, such that the long-term utility equals $\underline{u} \in F^*$.

The proof is not given here and is based on Theorem 2 of [8]. Note that such a characterization is very powerful. Indeed, the brute-force technique to find the feasible utility region would be to look at all possible action plans for the players and compute the corresponding utilities. This would be intractable even when every player could only choose two power levels for a finite stochastic games with 100 stages (each player could then choose between 2^{100} action plans). The Folk theorem characterizes the equilibrium utilities from quantities far much easier to evaluate (namely $E, F, \tilde{v}_i$). Additionally, as shown by Dutta [16], there is no loss of optimality by restricting the set of strategies to Markov strategies (which only depend on the current channel state). This will be exploited in Sec. 6

5. STRATEGIES FOR K-PLAYER GAMES

5.1. Best user selection (BUS)

The strategy we propose is based on the operating point presented in Sec. 3. When channels gains vary from stage to stage, if every transmitter plays at the operating point, the network is not socially

optimal (in contrast with the case where the channel state would be constant). It turns out that we get better results in terms of social welfare if the set of players playing the operating point at each stage is shrunk. We name this approach the *best user selection* scheme.

At each stage t of the game, the receiver sets $\mathcal{K}'^t \subset \mathcal{K}$, the optimal set of players playing the *Operating point* in terms of social welfare. For player $i \in \mathcal{K}$: • If $i \in \mathcal{K}'^t$, he is recommended to play the *Operating point* at stage t ; • If $i \notin \mathcal{K}'^t$, he should not transmit at this stage. To ensure the equilibrium of this strategy, if one of the player deviates from the strategy, the other players punish him by playing the one-shot *Nash equilibrium* for the remaining of the game.

5.2. Threshold-based user selection (T-US)

Note that in the former strategy, the set of players playing the *Operating point* is decided at each stage but one can also imagine a simpler strategy with a threshold $\alpha \in [0, 1]$ set for the entire game such that for player $i \in \mathcal{K}$: • If $\eta_i^t \geq \alpha \eta_{\max}^t$, he is recommended to play the *Operating point*; • If $\eta_i^t < \alpha \eta_{\max}^t$, he should not transmit at this stage, where $\eta_{\max}^t$ is the best channel gain realization at stage t and η_i^t the channel gain realization of player i at stage t . As before, if one of the players deviates from the strategy, the other players punish him by playing the one-shot *Nash equilibrium* for the remaining of the game.

5.3. Properties of Best User Selection Scheme

Although one might think that the BUS scheme requires complex computations at the base receiver, it happens that the former does not have to compare all possible combinations of players. Indeed, the BUS scheme always includes the player with the best channel gain and a given number of other players with the following channel gains in decreasing order. Thus, the base station just has to order the channels gains in decreasing order and decides which gain level is the minimum to separate players allowed to play from the others, which is far less complex, especially when the number of players is very large. In term of complexity, for K players, instead of considering 2^K possible combinations, the base station has just to consider K combinations.

Theorem 8 (BUS scheme for k players) *At equal transmitting rate, the best user selection of k players, $k \in \mathcal{K}$, playing together at the Operating point is the set of the k players with the best channels gains.*

For the strategy to be an equilibrium of the stochastic game, it is needed that no player has interest in deviating from the given plan. This condition is expressed by the fact that the cost of the punishment must always be higher than what a player can get when deviating at one stage. It results in the following theorem

Theorem 9 (Equilibrium strategy) *The BUS strategy is an equilibrium of the stochastic game if $\forall i \in \mathcal{K}$*

$$\lambda \leq \frac{\mathbb{E}[u_i(\underline{p}^{bus}, \underline{\eta})] - \mathbb{E}[u_i(\underline{p}^*, \underline{\eta})]}{\frac{R\eta_{\max}}{\sigma^2} \frac{f(\beta^*)}{\beta^*} + \mathbb{E}[u_i(\underline{p}^{bus}, \underline{\eta})] - \mathbb{E}[u_i(\underline{p}^*, \underline{\eta})]} \quad (16)$$

The expected utility of the BUS scheme can be compared to the strategy based on the one-shot *Nash equilibrium*, the strategy based on pure *Time-sharing*² and the *Operating point* strategy.

²At each stage, only the player with the best channel gain can play.

Theorem 10 (Dominance) *For equal transmission rates, for i.i.d. channel states among the players, we have*

$$\forall i \in \mathcal{K}, \mathbb{E}[u_i(\underline{p}^{bus}, \underline{h})] \geq \mathbb{E}[u_i(\underline{p}^*, \underline{h})]$$

$$\forall i \in \mathcal{K}, \mathbb{E}[u_i(\underline{p}^{bus}, \underline{h})] \geq \mathbb{E}[u_i(\underline{\tilde{p}}, \underline{h})]$$

$$\forall i \in \mathcal{K}, \mathbb{E}[u_i(\underline{p}^{bus}, \underline{h})] \geq \mathbb{E}[u_i(\underline{p}^{ts}, \underline{h})]$$

Simulations based on this strategy are discussed in Sec 6.3.

5.4. Information assumptions

For the BUS and the TUS schemes, players adapt their transmit power at the *Operating point* if they are recommended to play by the Base station. Thus, they need to know whether they are recommended to play. If so, given (5), they need to know the number of players transmitting with them and the state of their own channel. The recommendation and the number of recommended players are sent by the Base station whereas the channel state is observed by the transmitter, all these signals are modeled by θ . The last necessary piece of side information is due to the equilibrium condition of the strategy: if one player deviates from the plan, he is punished by the other players for the remaining of the game. It implies that players must be able to detect a deviation, what can be done if they know their SINR at each stage (see [6] for more details). From one strategy to another, the amount of information required may vary considerably. For instance, if one wants players to play the *Social Optimum* strategy, players would have to know channels gains of every other player whereas if they play the one-shot *Nash equilibrium* they would only need to know their own channel gain and the total number of players. To have a clear view about the amount of side information needed to implement the discussed strategies, we provide a simplified comparison for the various strategies under study in Fig. 1. By “deviation alarm” we mean a signal allowing players to detect single deviations from the cooperative action plan.

6. NUMERICAL RESULTS

6.1. Simulation Parameters

To obtain numerical results, we work with the efficiency function $f(\gamma) = e^{-\frac{a}{\gamma}}$ with $a = 2^R - 1$, see [2] for more details about this efficiency function. All our results are obtained from games with 10^5 stages.

6.2. Two-state channel K -player game

Fig. 2 is obtained for a 10-player game where channels gains can only reach two states $\eta_{\min}$ and $\eta_{\max}$ with probability $\{\frac{1}{2}, \frac{1}{2}\}$ for each player and $a = 0.1$. It is interesting in the sense that it clearly highlights the fact that the idea of the proposed strategy is interesting when channels gains of players are sufficiently different. Indeed, we can observe that for a channel gain ratio $\frac{\eta_{\max}}{\eta_{\min}} = 1$, the proposed strategy is equal to the classical *Operating point* strategy. For $\frac{\eta_{\max}}{\eta_{\min}} > 1$, the proposed strategy becomes more efficient. By working with random distributions for the channels gains, the case of realizations with very different channels gains often occurs, which is interesting for the proposed strategy.

	CSI	Recommendation signal	Nb of players	Deviation alarm
Pure time-sharing	individual	needed	not needed	not needed
One-shot Nash	individual	not needed	needed	not needed
Operating Point	individual	not needed	needed	SINR
T-US scheme	individual	needed	needed	SINR
BUS scheme	individual	needed	needed	SINR
Social optimum	global	not needed	not needed	SINR

Fig. 1. Information required for various stochastic strategies

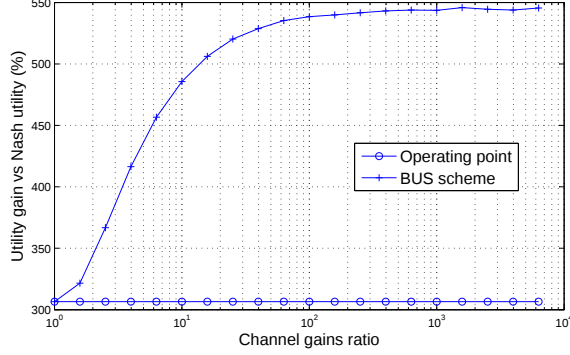


Fig. 2. Comparison of BUS utilities versus Nash utilities depending on the ratio between the good channel state and the bad channel state for a 10-player game.

Fig. 3 shows the achievable utility region for a 2-player, 2-states game (with $\frac{\eta_{\max}}{\eta_{\min}} = 4$ and $a = 0.5$) when considering all the possible strategies. The *minmax* line delimits the equilibrium region. The mean utilities of BUS, *Operating point* and static *Nash equilibrium* are also represented in this region. It is clear that BUS strategy is closer to the Pareto frontier.

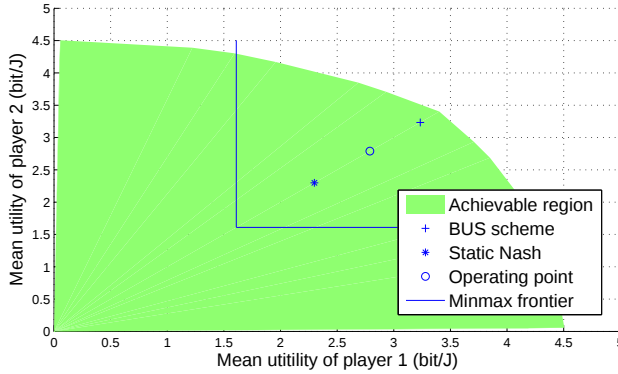


Fig. 3. Achievable region with expected utilities of various strategies

6.3. Truncated Rayleigh distribution for K -player game

In Fig. 4, we consider a K -player game with K varying from 1 to 10 and a fixed to 0.1. In each game, channels states follow the same truncated Rayleigh distribution law for every player. Four strategies are compared on this figure: one-shot *Nash equilibrium* (Sec. 3), pure *Time-sharing*, *Operating point* (equation (6)), T-US with $\alpha = 0.5$ (Sec. 5.2) and BUS (Sec. 5.1). There are several points to notice. First, for all the studied strategies, as the number of players increases, the mean utility decreases for each player. This is due to the fact that players see each other signal as interference: the more

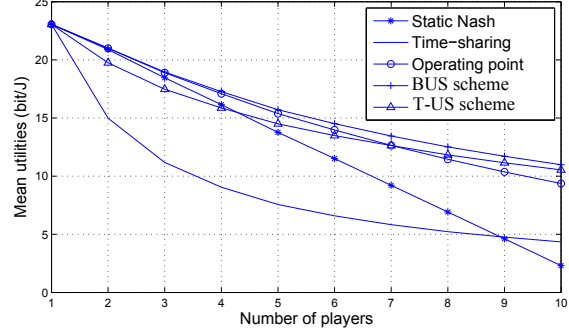


Fig. 4. Comparison of strategies utilities depending on the number of players

players in the game, the more they have to share resources. Second, it is clear that the *Operating point* and BUS strategies are more efficient than one-shot *Nash equilibrium* or *Time-sharing* strategies. As the number of players increases, this gap becomes even larger.

Fig. 5 shows a graphic representation of the different configurations $\mathcal{H}_1^i(k)$ and $\mathcal{H}_2^i(k)$ a player i can meet at each stage of the stochastic game (owing to the symmetry in the channels distribution law, the configurations probabilities are the same for every player). $\mathcal{H}_1^i(k) \subset \Gamma$ is the set of channels realizations where k players are recommended to play and player i is part of these players. $\mathcal{H}_2^i(k) \subset \Gamma$ is the set of channels realizations where player i is not one of the k players recommended to play. The simulation is made with 5 players and $a = 0.2$.

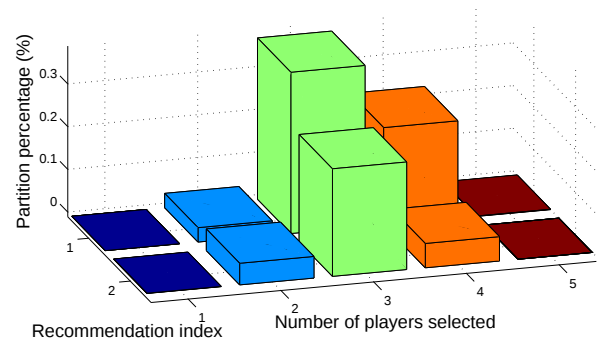


Fig. 5. Partition of $\mathcal{H}_1^i$ and $\mathcal{H}_2^i$ for a 5-player game with BUS strategy

Fig. 6 refers to theorem 9. It represents the maximum value the discount factor λ can have for the BUS strategy to be an equilibrium of the stochastic game. As the gap between the expected utility of

Best combination strategy and the expected utility of the one-shot Nash equilibrium increases with the number of players, the maximum value of the discount factor increases as well.

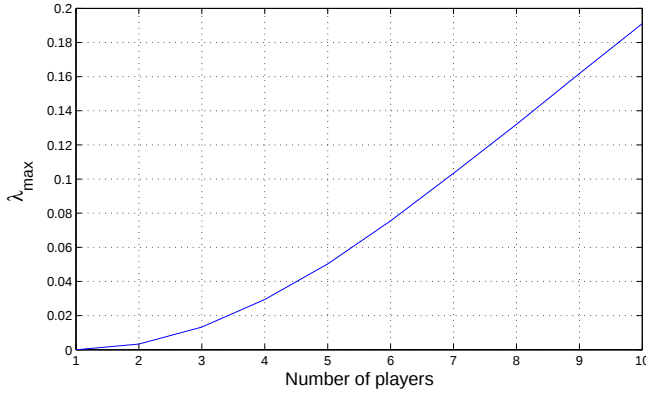


Fig. 6. Maximum value of discount factor λ for the BUS strategy to be an equilibrium.

7. CONCLUSION

Conventionally, for i.i.d channels, power control schemes are designed such that the power levels are chosen in an independent manner from block to block. In distributed networks with selfish transmitters, the point of view has to be re-considered even if the channels are i.i.d. due to the fact that long-term interaction may change the behavior of selfish transmitters. In order to take into account this effect and the fact that channel gains may vary from block to block, the model of stochastic games is proposed. When transmitters observe a public signal (e.g., the sum of received signals), a recent game-theoretic result allows one to fully characterize the equilibrium utility region. It is shown how to reach some points of this region by assuming individual CSI only. Both analytical and simulation results show potential gains in terms of energy-efficiency induced by the proposed model. In particular, because of long-term interaction, transmitters may have interest of shutting down for some blocks, leading therefore to legitimate time-sharing based control policies. Further investigations on the proposed approach are needed. In particular, it is relevant to characterize which part of the equilibrium utility region can be reached under the individual CSI assumption, which is relevant to include fairness issues. Additionally, typical features of modern wireless networks such as finite size buffers, Markovian evolution law for the channel state, should be accounted for to make the proposed framework more applicable.

8. REFERENCES

- [1] A. Mackenzie, *Game Theory for Wireless Engineers (Synthesis Lectures on Communications)*, Morgan & Claypool Publishers, 1 edition, May 2006.
- [2] E. V. Belmega, S. Lasaulce, and M. Debbah, "Power allocation games for mimo multiple access channels with coordination," *Trans. Wireless. Comm.*, vol. 8, no. 6, pp. 3182–3192, 2009.
- [3] D. J. Goodman and N. B. Mandayam, "Power control for wireless data," *IEEE Person. Comm.*, vol. 7, pp. 48–54, 2000.
- [4] C. U. Saraydar, N. B. Mandayam, and D. J. Goodman, "Efficient power control via pricing in wireless data networks," *IEEE Trans. on Communications*, vol. 50, no. 2, pp. 291–303, 2002.
- [5] S. Lasaulce, Y. Hayel, R. El Azouzi, and M. Debbah, "Introducing hierarchy in energy games," *IEEE Trans. on Wireless Comm.*, vol. 8, no. 7, pp. 3833–3843, 2009.
- [6] M. Le Treust and S. Lasaulce, "A repeated game formulation of energy-efficient decentralized power control," *IEEE Trans. on Wireless Commun.*, 2010.
- [7] S. Sorin, *Repeated Games with Complete Information*, in *Handbook of Game Theory with Economic Applications*, vol. 1, Elsevier Science Publishers, 1992.
- [8] J. Hörner, T. Sugaya, S. Takahashi, and N. Vieille, "Recursive methods in discounted stochastic games : An algorithm for delta approaching 1 and a folk theorem," Cowles Foundation Discussion Papers 1742, Cowles Foundation for Research in Economics, Yale University, Dec. 2009.
- [9] D. Fudenberg and Y. Yamamoto, "The folk theorem for irreducible stochastic games with imperfect public monitoring," *Journal of Economic Theory*, vol. In Press, Corrected Proof, 2011.
- [10] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE Signal Processing Magazine*, vol. 5, no. 2, pp. 4–24, April 1988.
- [11] F. Meshkati, M. Chiang, H. V. Poor, and S. C. Schwartz, "A game-theoretic approach to energy-efficient power control in multi-carrier cdma systems," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 6, pp. 1115–1129, 2006.
- [12] V. Rodriguez, F. Jondral, and R. Mathar, "Power allocation for social benefit through price-taking behaviour on a cdma reverse link shared by energy-constrained and energy-sufficient data terminals," in *ISWCS'09: Proceedings of the 6th international conference on Symposium on Wireless Communication Systems*, Piscataway, NJ, USA, 2009, pp. 56–60, IEEE Press.
- [13] F. Meshkati, H. V. Poor, S. C. Schwartz, and B. M. Narayan, "An energy-efficient approach to power control and receiver design in wireless data networks," *IEEE Transactions on Communications*, vol. 53, pp. 1885–1894, 2005.
- [14] E. V. Belmega and S. Lasaulce, "An information-theoretic look at mimo energy-efficient communications," *ACM Proc. of the Intl. Conf. on Performance Evaluation Methodologies and Tools (VALUETOOLS)*, 2009.
- [15] L. Shapley, "Stochastic games," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 39, no. 10, pp. 1095–1100, 1953.
- [16] P. K. Dutta, "A folk theorem for stochastic games," *Journal of Economic Theory*, vol. 66, no. 1, pp. 1 – 32, 1995.

Publications relatives aux jeux à champ moyen

B.1 SPRINGER-DGAA-2013

F. Mériaux, S. Lasaulce et H. Tembine. *Stochastic Differential Games and Energy-Efficient Power Control*. Dynamic Games and Applications, vol. 3, no. 1, pages 3-23, 2013.

Stochastic Differential Games and Energy-Efficient Power Control

François Mériaux · Samson Lasaulce ·
Hamidou Tembine

Received: date / Accepted: date

Abstract One of the contributions of this work is to formulate the problem of energy-efficient power control in multiple access channels (namely, channels which comprise several transmitters and one receiver) as a stochastic differential game. The players are the transmitters who adapt their power level to the quality of their time-varying link with the receiver, their battery level, and the strategy updates of the others. The proposed model not only allows one to take into account long-term strategic interactions but also long-term energy constraints. A simple sufficient condition for the existence of a Nash equilibrium in this game is provided and shown to be verified in a typical scenario. As the uniqueness and determination of equilibria are difficult issues in general, especially when the number of players goes large, we move to two special cases: the single player case which gives us some useful insights of practical interest and allows one to make connections with the case of large number of players. The latter case is treated with a mean-field game approach for which reasonable sufficient conditions for convergence and uniqueness are provided. Remarkably, this recent approach for large system analysis shows how scalability can be dealt with in large games and only relies on the individual state information assumption.

Keywords Differential games · Energy efficiency · Mean-field games · Power control · Wireless networks

François Mériaux
Laboratoire des Signaux et Systèmes
E-mail: meriaux@lss.supelec.fr

Samson Lasaulce
Laboratoire des Signaux et Systèmes
E-mail: lasaulce@lss.supelec.fr

Hamidou Tembine
Department of Telecommunications,
École Supérieure d'Électricité (SUPELEC)
E-mail: tembine@ieee.org

1 Introduction

Power control has always been recognized as an important problem for multiuser communications (Foschini and Miljanic, 1993; Yates, 1995). With the appearance of new paradigms such as ad hoc networks (Gupta and Kumar, 1997), unlicensed band communications, and cognitive radio (Fette, 2006; Mitola and Maguire, 1999), the study of distributed power control policies has become especially relevant; in such networks, terminals can freely choose their power control policies and do not need to follow orders from central nodes. The work reported in this paper precisely falls into this framework that is, the design of distributed power control policies in multiuser networks. More precisely, the assumed network model is a multiple access channel (MAC), which, by definition, includes several transmitters and one common receiver. A brief overview of previous works about power allocation for MACs is presented by Belmega et al (2009). In our framework, based on a certain knowledge which includes his individual channel state information, each transmitter has to tune his power level at each time instance. The literature of power control is vast and here we will only refer to the two closest bodies of related works. In the first body, the goal is to minimize the transmit power under constraints (data rate constraints typically). While energy minimization is sought, energy-efficiency is not necessarily high when measured in terms of a benefit to cost ratio (as it is done in Physics or Economics). Clearly, energy minimization and energy-efficiency maximization are two different approaches whose relevance depends on the context under consideration (see Goodman and Mandayam (2000) and related works for more justifications) and cannot be compared in general. The results provided in this paper concern the second body of works, in which the goal is to maximize energy-efficiency which is measured as an average number of successfully decoded bits per Joule consumed by the transmitter.

In the original formulation proposed by Goodman and Mandayam (2000) and re-used in most related works (e.g., Meshkati et al, 2006; Bonneau et al, 2008; Lasaulce et al, 2009; Buzzi and Saturnino, 2011), the problem of energy-efficient power control is modeled by a sequence of static games which are played independently from stage to stage. One implicit motivation behind this choice is that, in scenarios in which the channel state (i.e., the quality of the transmitter-receiver link) corresponds to i.i.d. realizations of a given random variable, correlating the power levels from block to block is a priori not relevant. But, when there exists a strategic interaction, this approach may be very suboptimal and the main drawback of the formulation of Goodman et al is precisely that it generally leads to an outcome (Nash equilibrium) which is not efficient. Motivated by this observation, Le Treust and Lasaulce (2010) proposed a repeated game formulation of the problem. One of the strong features of their formulation w.r.t. the famous pricing approach from Saraydar et al (2002) (which also aims at improving the efficiency of the game outcome on each block) is that each transmitter only needs to have individual channel state information. Although the repeated game model by Le Treust and Lasaulce

(2010) takes into account the fact that transmitters interact several/many times, the corresponding work has one major weakness: there is a need for a normalized stage game which does not depend on the realization channels. This normalization is valid only if no player has his power constraint active and even in this case, there is a loss of optimality in terms of expected utilities (note that this optimality loss is also undergone by the static game formulation with pricing proposed by Saraydar et al (2002)). This is one of the main reasons why we propose a different formulation which is based on stochastic games. One of the goals of the present paper is to study the influence of long-term strategic interaction in a game with states and long-term energy constraints (e.g., the limited battery life typically) on energy-efficient power control. Indeed, in the work of Goodman and Mandayam (2000) and related references (e.g. Meshkati et al, 2006; Bonneau et al, 2008; Lasaulce et al, 2009; Buzzi and Saturnino, 2011), the terminals always transmit, which amounts to considering no constraint on the available energy. More specifically, the energy-efficient power control problem in MAC under long-term energy constraints is modeled by a stochastic differential game (SDG) in which the existence of a Nash equilibrium is proven. But the problem of characterizing the performance of distributed networks modeled by SDG becomes hard and even impossible when the number of players becomes large. The same statement holds for determining individually optimal control strategies. For instance, in a previous work from Mériaux et al (2011b), equilibrium control strategies are proposed but they are not optimal strategies. This is where mean-field games come into play. Mean-field games (Lasry and Lions, 2007) represent a way of approximating a stochastic differential (or difference) game, by a much more tractable model. Under the assumption of individual state information, the idea is precisely to exploit as an opportunity the fact that the number of players is large to simplify the analysis. Typically, instead of depending on the actions and states of all the players, the mean-field utility of a player only depends on his own action and state, and depends on the others through an mean-field. It seems that the most relevant work in which mean-field games have been used for power control is given by Tembine et al (2010). Compared to the latter reference, the present work is characterized by a different utility function (no linear quadratic control assumptions is made here), a more realistic channel evolution law, and the fact that the battery level of a transmitter is considered as part of a terminal state.

The remaining of the paper is organized as follows. Section 2 gives a brief review of the static game formulation of the energy-efficient power control problem. Section 3 introduces the stochastic dynamic game modeling the energy-efficient power control game under long-term energy constraints. In Section 4, two particular cases of the game are studied. The single player case highlights the impact of a long-term energy constraint on the transmitter power policy. Then the large system case is modeled by a mean-field game. Section 5 concludes this work.

Notations: In the following, scalars and vectors are respectively denoted by lower case symbols and underlined lower case symbols. The vector

$\underline{a}_{-k} = (a_1, \dots, a_{k-1}, a_{k+1}, \dots, a_K)$ denotes the vector obtained by dropping the k -th component of the vector $\underline{a}$. With a slight abuse of notation, the vector $\underline{a}$ can be written $(a_k, \underline{a}_{-k})$, in order to emphasize the influence of its k -th component. $\nabla_{\underline{x}} f$ and $\Delta_{\underline{x}} f$ respectively represent the gradient and the Laplacian of the function f w.r.t. the vector $\underline{x}$. $\text{div}_{\underline{x}}$ is the divergence operator w.r.t. the vector $\underline{x}$. $\langle \cdot, \cdot \rangle_{\mathbb{E}}$ represents the scalar product in the space $\mathbb{E}$.

2 Review of the static game formulation of the energy-efficient power control problem

The purpose of this section is to provide a brief review of how Goodman and Mandayam (2000) formulated the power control problem. The motivation for this is twofold. It allows us to have a reference for comparison and also allows us to build in a clearer manner the SDG formulation. The communication scenario is a multiple access channel (Cover and Thomas, 1991). There are $K \geq 1$ transmitters and one receiver. Each transmitter sends a signal to a common receiver and has to choose the power level of the transmitted signal. In order to optimize his individual energy-efficiency, i.e., the ratio of his throughput to its transmit power, each transmitter has not only to adapt his power level to the quality of the channel or link between him and the receiver but also to the power levels chosen by the others. The static game formulation of this problem is as follows.

Definition 1 (Static game model of the power control problem)

The strategic form of the static power control game is a triplet

$\bar{\mathcal{G}} = (\mathcal{K}, \{\mathcal{P}_i\}_{i \in \mathcal{K}}, \{\bar{u}_i\}_{i \in \mathcal{K}})$ where:

- $\mathcal{K} = \{1, 2, \dots, K\}$ is the set of transmitters;
- $\mathcal{P}_i = [0, P_i^{\max}]$ is the action space of player $i \in \mathcal{K}$;
- the utility function of player $i \in \mathcal{K}$ is given by

$$\bar{u}_i(p_1, p_2, \dots, p_K) = \frac{Rf(\gamma_i(p_1, p_2, \dots, p_K))}{p_i} \text{ [bit/J]}, \quad (1)$$

where:

—

$$\gamma_i(p_1, p_2, \dots, p_K) = \frac{p_i |\underline{h}_i|^2}{\sum_{j \in \mathcal{K}, j \neq i} p_j |\underline{h}_j|^2 + \sigma^2}; \quad (2)$$

- $\forall i \in \mathcal{K}$, $\underline{h}_i \in \mathbb{R}^2$ is a vector of parameters which represent the quality of the channel between transmitter i and the receiver;
- σ^2 is a constant which models the communication noise effects at the receiver;
- R is a constant in [bit/s] which corresponds to the communication data rate (see Goodman and Mandayam, 2000);

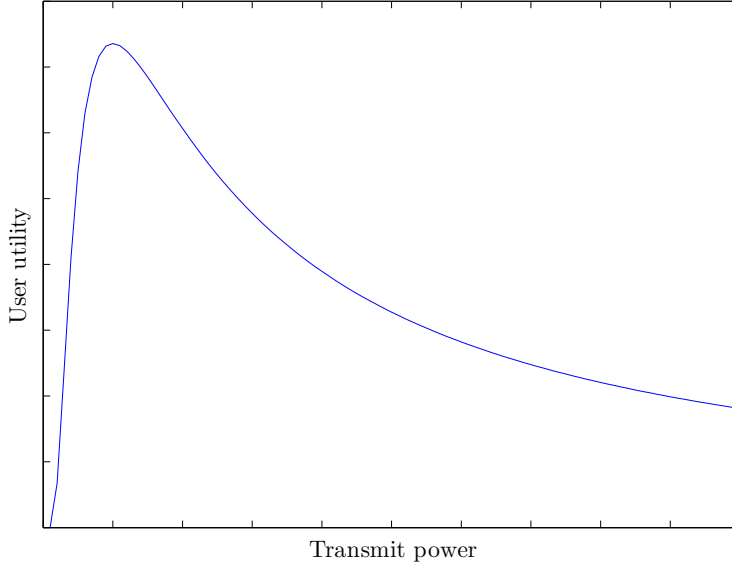


Fig. 1 Typical form of the utility function regarding to the transmit power.

- the function $f : \mathbb{R}^+ \rightarrow [0, 1]$ is a sigmoidal or S-shaped function which represents the packet success rate; recall that a sigmoidal function is convex up to a point and then concave from this point. Additionally, f is assumed to be sufficiently regular so that $\bar{u}_i$ is differentiable on $\mathcal{P}_i$. See the work of e.g. Rodriguez (2003); Meshkati et al (2005); Belmega and Lasaulce (2011) for a justification.

The above performance metric can be seen as a tradeoff between the data rate conveyed over the air and the electromagnetic pollution (since the radiated power is concerned) in the corresponding region. Fig.1 illustrates the typical shape of the utility function with regard to transmit power. A consequence of the use of this utility function is that the optimal transmit power generally does not make the maximum power constraint active. Precisely, this static game can be shown to be quasi-concave and has therefore a pure NE (see e.g. Lasaulce and Tembine, 2011). Additionally, this NE is unique and is given by:

$$\forall i \in \{1, \dots, K\}, p_i^* = \frac{\sigma^2}{|h_i|^2} \frac{\beta^*}{1 - (K-1)\beta^*} \quad (3)$$

where β^* is the unique solution of the equation

$$xf'(x) - f(x) = 0. \quad (4)$$

It has been shown by Goodman and Mandayam (2000), Meshkati et al (2006), that this equilibrium is not Pareto-optimal. Assuming that the chan-

nel coefficients are i.i.d., one could have thought that the study the one-shot energy-efficient power control problem was sufficient to understand problems of energy-efficient power control over time. But even with i.i.d. channel coefficients, there is a new phenomenon when the game is played over time which is not taken into account by the one-shot game: strategic interactions over time.

Moving to repeated game is a way to take into account these interactions and Le Treust and Lasaulce (2010) have shown that more efficient solutions can be obtained in such a framework. However, to account for channel variations and varying energy level in the battery, using repeated games is not sufficient. Stochastic games become necessary and are useful to further improve efficiency. The main problem is that both the analysis and algorithm design become complicated. This is why we propose to use mean-field games. Therefore, the merit of the proposed approach is to apply stochastic games in the energy-efficient power control framework to obtain efficient solutions and simplify both the analysis and design of algorithms when the system is large. In particular, only the individual channel state information, battery state, and the interference level at the receiver are needed to implement the proposed power control policy. This is an attractive practical feature which is not available for competitive approaches based on stochastic games with finite number of players.

3 A stochastic differential game formulation

In this section, we present an SDG which is built from the static game described in the previous section. Time is assumed be continuous that is, $t \in \mathbb{R}$. This assumption has been discussed in other works on power control (see e.g. Foschini and Miljanic, 1993; Olama et al, 2006; Tembine et al, 2010). In particular, it is relevant in scenarios in which the channels are subject to fast fading or interpreted as a limiting case for slow fading channels. From now on, we will make appear explicitly the channel states as arguments of the instantaneous utility function which will be denoted by u_i instead of $\bar{u}_i$. The time horizon of the game is finite, it is the interval ranging from T to T' . However the methodology can be extended to infinite horizon (with time average payoff) because the underlying processes are ergodic (the proposed channel model is ergodic and the remaining energy dynamics is ergodic).

Definition 2 (SDG model of the power control problem)

The stochastic differential power control game is defined by the 5-tuple $\mathcal{G} = (\mathcal{K}, \{\mathcal{P}_i\}_{i \in \mathcal{K}}, \{\mathcal{X}_i\}_{i \in \mathcal{K}}, \{\mathcal{S}_i\}_{i \in \mathcal{K}}, \{U_i\}_{i \in \mathcal{K}})$ where:

- $\mathcal{K}$ and $\mathcal{P}_i$ are defined by Def. 1;
- $\mathcal{X}_i$ is the state space of player i . The game state at time t is defined by $\underline{X}_i(t) = [E_i(t), \underline{h}_i(t)]^T$ and follows the evolution law:

$$d\underline{X}_i(t) = \begin{bmatrix} -p_i(t) \\ \frac{1}{2}(\underline{\mu} - \underline{h}_i(t)) \end{bmatrix} dt + \begin{bmatrix} 0 \\ \eta \end{bmatrix} d\underline{W}_i(t); \quad (5)$$

- $E_i(t)$ is the energy available for player i at time t ;
- $\forall i \in \mathcal{K}$, $\mathbb{W}_i(t)$ are mutually independent Wiener processes of dimension 2;
- $\underline{\mu} \geq 0$ and $0 \leq \eta < +\infty$ are constants which physical interpretation is given in Prop. 1;
- $\mathcal{S}_i$ is the set of feedback control policies for player i (see e.g. Basar and Olsder, 1999). A control policy will be denoted by $p_i(T \rightarrow T')$ with $p_i(t) = p_i(t, \underline{X}(t))$, and T, T' two reals such that $T' \geq T$;
- the average utility function U_i is defined by:

$$U_i(p(T \rightarrow T')) = \mathbb{E} \left[\int_T^{T'} u_i(p(t), \underline{X}(t)) dt + q(\underline{X}(T')) \right], \quad (6)$$

where $p(T \rightarrow T') = (p_1(T \rightarrow T'), p_2(T \rightarrow T'), \dots, p_K(T \rightarrow T'))$ is the control strategy profile, $\underline{X}(t) = [\underline{X}_1(t), \underline{X}_2(t), \dots, \underline{X}_K(t)]^T$ is the state profile, $q(\underline{X}(T'))$ is the utility at the final state, and u_i is the instantaneous utility.

The motivations for selecting the proposed dynamics for the state $\mathcal{X}_i$ are as follows. The term $dE_i(t) = -p_i(t)dt$ means that the variation of energy during dt is proportional to the consumed power for the transmission. Indeed, the proposed model accounts for a cost when transmitting. This is fully relevant for transmitters having a finite amount of energy at disposal like cell phones, unplugged laptops, small base stations which have to be autonomous energetically speaking, etc. Such terminals have a battery with a finite amount of energy over a certain period of time and need to be recharged when empty. Our model holds over a horizon which lies strictly between two recharging instants. Over such a horizon, the available energy is a non-increasing function of the time. Although the assumed evolution law can be used for both fast and slow fading, in practice, when implementing a discrete-time version of the control policy, the energy decrease from sample to sample is stronger in the second case. Taking into account the energy of the battery in the game model changes the outcome of the game. For instance, if the battery of a transmitter is empty, the optimal power level has to be zero and cannot be the power levels recommended by the static game approach such as those given by Goodman and Mandayam (2000) or Saraydar et al (2002).

As far as time horizons are concerned, T and T' can be chosen arbitrarily provided that $T' \geq T$. Of course, the statistics of the evolution law are required to be stationary on this interval. In particular, $\underline{\mu}$ and η have to be fixed. As for the order of horizon measured in second, it depends on the targeted application. Note that, in practice, if those parameters need to be known and updated, appropriate estimation schemes (with well-chosen time windows) are implemented. Let us consider two cases, fast and slow power control. For fast power control (the statistics are given by the path loss which is fixed), if the evolution approximates a discrete-time power control problem for which the (fast fading) channel is i.i.d. from block to block (in this case, η is large), $T' - T$ is of the order of the second (a packet duration is typically 1 ms in cellular systems). In practice, most often, power levels are updated from block to block (meaning that typical updating frequency ranges from 100 Hz to a few kHz).

As shown by Foschini and Miljanic (1993), deriving continuous-time power control policies are still useful since it is possible to build practical discrete-time algorithms from them. Additionally, when one studies the convergence of these discrete time algorithms, often, it amounts to studying continuous-time dynamics. For slow power control, if the evolution law represents the variation of the pathloss/shadowing/slow fading (as assumed by Olama et al (2006) and related works), the time horizon is much larger, typically of the order of a minute or more (this depends on the mobile velocity of course). For both fast and slow power control, the number of samples to approximate the integrals can be of the same order since the updating frequencies are different.

Although being simple, the dynamics for the channel gain $\underline{h}_i(t)$ capture several typical effects in wireless communications. Before commenting on these effects, let us state a property of the random process $\underline{h}_i(t)$.

Proposition 1 (Channel dynamics property)

Let $\underline{h}_i(t) = (x_i(t), y_i(t))$ be governed by the dynamics defined in the SDG $\mathcal{G}$, then we have that:

$$\begin{aligned} \lim_{t \rightarrow +\infty} \mathbb{E}[\underline{h}_i(t)] &= \underline{\mu}, \\ \lim_{t \rightarrow +\infty} \mathbb{E}[|\underline{h}_i(t)|^2] - \mathbb{E}[|\underline{h}_i(t)|]^2 &= 2\eta^2. \end{aligned} \quad (7)$$

The stationary probability density functions $m_x : \mathbb{R} \rightarrow \mathbb{P}(\mathbb{R})$, $m_y : \mathbb{R} \rightarrow \mathbb{P}(\mathbb{R})$ of the two components x_i , y_i of h_i are:

$$\begin{cases} m_x(x_i) = \frac{1}{\eta\sqrt{2\pi}} e^{-\frac{(x_i - \mu_x)^2}{2\eta^2}}, \\ m_y(y_i) = \frac{1}{\eta\sqrt{2\pi}} e^{-\frac{(y_i - \mu_y)^2}{2\eta^2}}. \end{cases} \quad (8)$$

The proof of this result is simple and provided in Appendix A. This shows that the proposed dynamics allow one to model Rician channels namely, channels with zero-mean gains; this is possible by tuning $|\underline{\mu}| \geq 0$ which represents the Rice component. Also, by choosing the variance $2\eta^2$ in an appropriate manner, one can account for the fading effects. As a relevant comment, note that the assumed channel dynamics can also be seen as a limiting case of the important Gauss-Markov discrete-time model used to model time correlation for the channel gains (Agarwal and Honig, 2012). To conclude on the choice of these dynamics, we will see in Sec. 4 that they also possess an interesting property for the mean-field dynamics under investigation.

At this point, we can define a Nash equilibrium of the SDG $\mathcal{G}$ and state our existence result.

Definition 3 (Nash equilibrium of $\mathcal{G}$)

A control strategy profile $\underline{p}^*(t, \underline{X}(t)) = (p_1^*(t, \underline{X}(t)), \dots, p_K^*(t, \underline{X}(t)))$ is a feedback Nash equilibrium of the SDG if and only if $\forall i \in \mathcal{K}$, p_i^* is a solution of the control problem

$$\sup_{p_i(T \rightarrow T')} \mathbb{E} \left[\int_T^{T'} u_i \left(p_i(t, \underline{X}(t)), \underline{p}_{-i}^*(t, \underline{X}(t)), \underline{X}(t) \right) dt + q(\underline{X}(T')) \right], \quad (9)$$

subject to

$$d\mathbf{X}(t) = \begin{bmatrix} -[p_i(t), \underline{p}_{-i}^*(t)]^T \\ \frac{1}{2}(\mathbf{1} \otimes \underline{\mu} - \underline{h}(t)) \end{bmatrix} dt + \begin{bmatrix} 0 \\ \eta \end{bmatrix} d\mathbf{W}(t), \quad (10)$$

where $\mathbf{1} = (1, 1, \dots, 1) \in \mathbb{R}^K$, $\underline{h}(t) = [\underline{h}_1(t), \underline{h}_2(t), \dots, \underline{h}_K(t)]^T$, $\otimes$ stands for the Kronecker product, and $\mathbf{W}(t) = [\mathbf{W}_1(t), \mathbf{W}_2(t), \dots, \mathbf{W}_K(t)]^T$.

Regarding to the existence of a Nash equilibrium, one can state the following proposition (the function f is defined by Def. 1).

Proposition 2 (Existence of a Nash equilibrium in $\mathcal{G}$)

A sufficient condition for the existence of a Nash Equilibrium in $\mathcal{G}$ is that for all (θ_0, γ_0) such that $f'(\gamma_0)\gamma_0 - f(\gamma_0) = \theta_0\gamma_0^2$, we have that $2\theta_0 - f''(\gamma_0) \neq 0$.

For the proof of this result and for clarifying some points in the sequel, we will use the (auxiliary) Bellman function, which is defined by:

$$v_i(T, \mathbf{X}(T)) = \sup_{p_i(T \rightarrow T')} \mathbb{E} \left[\int_T^{T'} u_i(\underline{p}(t), \mathbf{X}(t)) dt + q(\mathbf{X}(T')) \right]. \quad (11)$$

Proof According to Bressan (2010), a sufficient condition for the existence of a Nash equilibrium for the SDG is the existence of a solution to the Hamilton-Jacobi-Bellman-Fleming (Fleming and Soner, 1993) equation for each transmitter

$$\begin{aligned} 0 = & \sup_{p_i(T \rightarrow T')} \left[u_i(\mathbf{X}(t), \underline{p}(t, \mathbf{X}(t))) - p_i(t, \mathbf{X}(t)) \frac{\partial v_i(t, \mathbf{X}(t))}{\partial E_i} \right] \\ & + \frac{1}{2} \langle \underline{\mu} - \underline{h}_i(t), \nabla_{\underline{h}_i} v_i(t, \mathbf{X}(t)) \rangle_{\mathbb{R}^2} + \frac{\partial v_i(t, \mathbf{X}(t))}{\partial t} + \frac{\eta^2}{2} \Delta_{\underline{h}_i} v_i(t, \mathbf{X}(t)). \end{aligned} \quad (12)$$

There exists a solution if the function

$$\begin{aligned} H \left(\mathbf{X}(t), \underline{p}_{-i}(t, \mathbf{X}(t)), \frac{\partial v_i(t, \mathbf{X}(t))}{\partial E_i} \right) = \\ \sup_{p_i(T \rightarrow T')} \left[u_i(\mathbf{X}(t), \underline{p}(t, \mathbf{X}(t))) - p_i(t, \mathbf{X}(t)) \frac{\partial v_i(t, \mathbf{X}(t))}{\partial E_i} \right] \end{aligned} \quad (13)$$

is smooth (see Evans (2010) for more details). And with a similar reasoning as Mériiaux et al (2011a), we can show that finding optimal power control

$$\underline{p}^*(T \rightarrow T') = (p_1^*(T \rightarrow T'), \dots, p_n^*(T \rightarrow T')) \quad (14)$$

such that $\forall i \in \mathcal{K}$, $p_i^*(T \rightarrow T') \in$

$$\arg \max_{p_i(T \rightarrow T')} \left[u_i \left(\mathbf{X}(t), p_i(t, \mathbf{X}(t)), \underline{p}_{-i}^*(t, \mathbf{X}(t)) \right) - p_i(t, \mathbf{X}(t)) \frac{\partial v_i(t, \mathbf{X}(t))}{\partial E_i} \right] \quad (15)$$

amounts to solving $\forall i \in \mathcal{K}, \forall t \in [T, T']$

$$f'(\gamma_i(t))\gamma_i(t) - f(\gamma_i(t)) = \frac{\gamma_i(t)^2}{R} \frac{\partial v_i(t, \underline{X}(t))}{\partial E_i} \left(\frac{\sigma^2 + \sum_{j \neq i}^n |\underline{h}_j(t)|^2 p_j^*(t)}{|\underline{h}_i(t)|^2} \right)^2. \quad (16)$$

Note that we consider that $\frac{\partial v_i(t, \underline{X}(t))}{\partial E_i} \geq 0$, otherwise the optimal power $p_i^*(t) \rightarrow \infty$. The existence of a non-zero solution depends on the term

$$\theta_i = \frac{1}{R} \frac{\partial v_i(t, \underline{X}(t))}{\partial E_i} \left(\frac{\sigma^2 + \sum_{j \neq i}^n |\underline{h}_j(t)|^2 p_j^*(t)}{|\underline{h}_i(t)|^2} \right)^2. \quad (17)$$

It can be checked that there exists a threshold $\theta_{\max}$ such that if $\theta_i < \theta_{\max}$, there exists a unique global maximizer γ^* different from 0 and if $\theta_i \geq \theta_{\max}$, 0 is the global maximizer.

We call for the implicit function theorem to state smoothness of

$H\left(\underline{X}(t), \underline{p}_{-i}(t, \underline{X}(t)), \frac{\partial v_i(t, \underline{X}(t))}{\partial E_i}\right)$, we define

$$\begin{aligned} g : [0, \theta_{\max}[\times \mathbb{R}^+ &\rightarrow \mathbb{R} \\ (\theta, \gamma) &\rightarrow f'(\gamma)\gamma - f(\gamma) - \theta\gamma^2. \end{aligned} \quad (18)$$

g is C^∞ , then if $g(\theta_0, \gamma_0) = 0$, there exists $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ such that $\gamma_0 = \varphi(\theta_0)$. φ is C^∞ and

$$\frac{\partial \varphi}{\partial \theta}(\theta_0) = - \frac{\frac{\partial g(\theta_0, \gamma_0)}{\partial \theta}}{\frac{\partial g(\theta_0, \gamma_0)}{\partial \gamma}}. \quad (19)$$

In our case, it writes

$$\frac{\partial \varphi}{\partial \theta}(\theta_0) = \frac{\gamma_0}{2\theta_0 - f''(\gamma_0)}. \quad (20)$$

If $2\theta_0 - f''(\gamma_0) \neq 0$, then smoothness is ensured. $\blacksquare$

Remarkably, the proposed sufficient condition holds for all particular choices of efficiency function made in the literature. In particular, it holds for the information-theoretic choice of Belmega and Lasaulce (2011). Indeed, if $f(x) = e^{-\frac{a}{x}}$, $a \geq 0$ we have that:

$$\begin{cases} e^{-\frac{a}{\gamma_0}} \left(\frac{a}{\gamma_0^2} - 1 \right) = \theta_0 \gamma_0^2, \\ 2\theta_0 - \left(\frac{a^2}{\gamma_0^4} - \frac{2a}{\gamma_0^3} \right) e^{-\frac{a}{\gamma_0}} \neq 0, \end{cases} \quad (21)$$

which gives

$$2\theta_0 - \left(\frac{a}{\gamma_0} - 2 \right) \frac{a\theta_0}{a - \gamma_0} \neq 0. \quad (22)$$

While Nash equilibrium uniqueness is an attractive property of the static game $\mathcal{G}$, this property is not easy to be verified for the SDG $\mathcal{G}$. Rather, this type of games has generally a large number of equilibria. Concerning the explicit determination of Nash equilibrium power control policies, it has to be mentioned that this task is also not easy a priori. Precisely, as written in

(12), the determination of a Nash equilibrium requires to solve a system of K Hamilton-Jacobi-Bellman-Fleming equations, coupled by the state $\underline{X}(t)$. Interestingly, there are two limiting cases of $\mathcal{G}$ for which both uniqueness and existence issues are much easier. The first special case is when there is only one player (note that the corresponding optimization problem has not been studied in the literature). The second case is when the number of players is large, making the mean-field game analysis fully relevant. These two cases are the purpose of the next section.

4 Two relevant special cases of $\mathcal{G}$: $K = 1$ and $K \rightarrow +\infty$

4.1 The single-player case ($K = 1$)

One of the interests in analyzing the single-user case is to separate the effects due to the long-term energy constraint from those due to interaction between players (two effects can incite the transmitter to be off, namely a bad channel state, and high interference level). Indeed, in the single-player case only the former effects appear. Below, it is proven through simple equations that the transmitter is not always on. The fraction of time during which the transmitter has to be off is approximated, which is of practical interest. When moving to the case of several players, players can also have an interest in not transmitting (as observed by Mériiaux et al (2011a)), making appropriate time-sharing policies natural equilibria. Summing up, both reducing the interference between the transmitters and long-term energy constraint can incite a transmitter to be off.

In this section, we therefore study the special case of the game $\mathcal{G}$ in which there is only one player. In this context, there is obviously no interaction between players and the main interest of this case is to show the influence of the long-term energy constraint. From the preceding section, it can be seen that determining an optimal control policy amounts to solving the following equation in $\gamma_1(t)$:

$$f'(\gamma_1(t))\gamma_1(t) - f(\gamma_1(t)) = \frac{\gamma_1(t)^2}{R} \frac{\partial v_1(t, \underline{X}_1(t))}{\partial E_1} \frac{\sigma^4}{|\underline{h}_1(t)|^4}. \quad (23)$$

By denoting $\gamma_1^*(t)$ the largest solution of the above equation, an optimal control policy follows (see (2)):

$$p_1^*(t) = \frac{\sigma^2}{|\underline{h}_1(t)|^2} \gamma_1^*(t), \quad (24)$$

with $\gamma_1^*(t) \geq 0$. Interestingly, $\gamma_1^*(t) = 0$ can also occur at given time instants. In particular, this depends on the channel quality $|\underline{h}_1(t)|$. If the latter is too low, transmitting is not energy-efficient, leading to a vanishing transmit power. With the same reasoning as Mériiaux et al (2011a), the fraction of the time

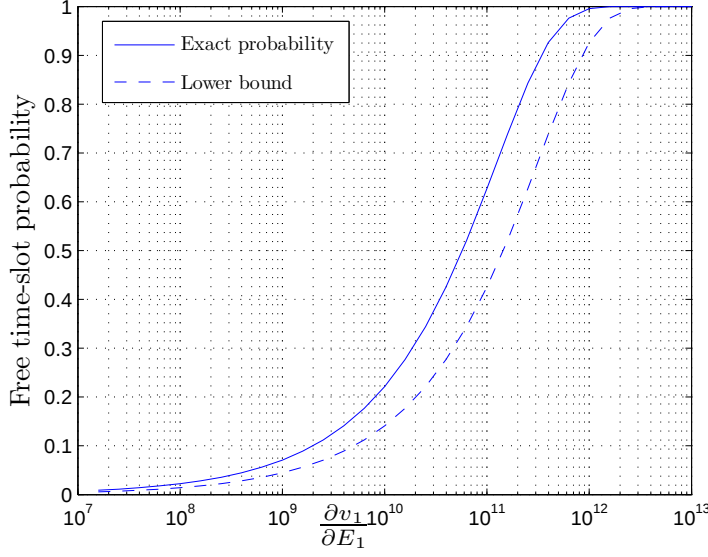


Fig. 2 Exact probability of the transmitter being off and lower bound of this probability depending on $\frac{\partial v_1(t, \underline{X}_1(t))}{\partial E_1}$.

during which the transmitter is off can be assessed by using a simple lower bound on the probability that $\gamma_1^*(t) = 0$:

$$\Pr \left[\max f'' \leq 2 \frac{\partial v_1(t, \underline{X}_1(t))}{\partial E_1} \frac{\sigma^4}{R |\underline{h}_1(t)|^4} \right] \leq \Pr[\gamma_1^*(t) = 0]. \quad (25)$$

Many Monte Carlo simulations have shown that this lower bound is reasonably tight, one of them is provided in figure 2. This result is of practical interest since it allows one to quantify the impact of a long-term energy constraint on power control policies, which is one of the goals of this paper.

4.2 The mean-field game analysis ($K \rightarrow +\infty$)

4.2.1 Convergence to the mean-field game

We have mentioned in the preceding sections that the SDG becomes more and more difficult to analyze when the number of players increases. However, our problem has a special structure which can be exploited and simplifies the problem when K is large. Indeed, from a given player standpoint, what matters in terms of utility is a weighted sum of the played actions. The relevant quantity which affects the utility of player i is the following quantity:

$$I_i(t) = \frac{1}{K} \sum_{j \in \mathcal{K}, j \neq i} p_j(t) |\underline{h}_j(t)|^2. \quad (26)$$

The quantity $I_i(t)$ is called the interference in communication networks. Here, we have normalized this quantity. There are many justifications of practical interest for this normalization (see e.g. Tulino and Verdú, 2004). For example, it is fully justified in random CDMA systems (Meshkati et al, 2006). If it can be proven that if $I_i(t)$ converges, then the game $\mathcal{G}$ converges to a mean-field game. To justify the convergence here, the conventional weak law of large numbers is not applicable since the random processes $p_j(t)|\underline{h}_j(t)|^2$ are interdependent. However, there still exist some conditions under which convergence is ensured. One of them is the exchangeability or indistinguishability property which is defined next.

Definition 4 (Exchangeability)

The states $X_1, X_2, \dots, X_K$ are said to be exchangeable in law under the feedback strategies α if they generate a joint law which is invariant by permuting the players indices, i.e.,

$$\forall K, \mathcal{L}(\underline{X}_1, \dots, \underline{X}_K \mid \alpha) = \mathcal{L}(\underline{X}_{\pi(1)}, \dots, \underline{X}_{\pi(K)} \mid \alpha), \quad (27)$$

for any bijection π (one-to-one mapping) defined over $\{1, \dots, K\}$.

To guarantee this property for the game under study we make the following assumptions:

- each player only knows his individual state;
- each player implements an homogeneous admissible control:

$$p_i(t) = \alpha(t, \underline{X}_i(t)); \quad (28)$$

$$- \mathbb{E} \left[\int_T^{T'} \alpha(t, \underline{X}_i(t))^2 dt \right] < +\infty.$$

As a consequence of the exchangeability property in $\mathcal{G}$, the game now comprises generic players. From now on, we call $\underline{s}(t) = [E(t), \underline{h}(t)]^T$ the generic individual state of a player. The state dynamics of a generic player is given by the following stochastic differential equations (SDEs):

$$\begin{cases} dE(t) = -\alpha(t, \underline{s}(t))dt, \\ d\underline{h}(t) = \frac{1}{2}(\underline{\mu} - \underline{h}(t))dt + \eta d\underline{W}(t). \end{cases} \quad (29)$$

where $\underline{\mu}, \eta$ are time-independent.

Before discussing the convergence of $\mathcal{G}$ to a mean-field game, we introduce some notations to define the mean-field game concept properly. Let $(\Omega, \mathbb{P}, \mathcal{F})$ be a complete filtered probability space, on which a one-dimensional standard Brownian motion $\mathcal{W}$ is defined with $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$, being its natural filtration augmented by all the $\mathbb{P}$ -null sets (sets of measure-zero with the respect $\mathbb{P}$.) The filtration $\mathcal{F}_t$ will be combined with the one generated by the initial state of the players. Note that at time t the trajectory generated by state dynamics (29) is in $\mathcal{F}_t$. The state given by the channel gains and the battery levels follows a certain distribution which evolves over time: this distribution is called the mean field.

Proposition 3 (Convergence to the mean-field game)

If the states $(\underline{X}_i(t))_{i,t}$ and the admissible controls $(\alpha(t, \underline{X}_i(t)))_{i,t}$ preserve the indistinguishability property and the $(\underline{X}_i(0))_i$ are indistinguishable, then the stochastic differential game converges to a mean-field game.

Proof

$$\begin{aligned} I_i(t) &= \frac{1}{K} \left[\sum_{j=1}^K \alpha(t, \underline{X}_j(t)) |\underline{h}_j(t)|^2 - \alpha(t, \underline{X}_i(t)) |\underline{h}_i(t)|^2 \right], \\ &= \int_{\underline{s}} |\underline{h}|^2 \alpha(t, \underline{s}) M_t^K(d\underline{s}) - \frac{\alpha(t, \underline{X}_i(t)) |\underline{h}_i(t)|^2}{K}, \end{aligned} \quad (30)$$

with

$$M_t^K = \frac{1}{K} \sum_{j=1}^K \delta_{\underline{s}_j(t)}, \quad (31)$$

where $\delta_{\underline{s}_j(t)}$ is the Dirac measure concentrated at $\underline{s}_j(t)$.

If the number of transmitters becomes very large ($K \rightarrow \infty$), we can consider that we have a continuum of transmitters. The convergence of the interference term when $K \rightarrow \infty$ needs to be proven. Using admissible control, $\mathbb{E}[\alpha(t, \underline{X}_i(t)) |\underline{h}_i(t)|^2] < \infty$, then

$$\lim_{K \rightarrow \infty} \frac{\alpha(t, \underline{X}_i(t)) |\underline{h}_i(t)|^2}{K} = 0. \quad (32)$$

To prove $I_i(t)$ converges weakly, it suffices to prove $\int_{\underline{s}} |\underline{h}|^2 \alpha(t, \underline{s}) M_t^K(d\underline{s})$ converges weakly. A sufficient condition is the weak convergence of the process M_t^K . Since we have chosen the control law that preserves the indistinguishability property, we can use the work by Tembine and Huang (2011) which states that there exists a distribution m_t such that

$$m_t = \lim_{K \rightarrow \infty} M_t^K. \quad (33)$$

This distribution m_t is the mean-field. In our case, the evolution of the state of the transmitters does not depend on the index of the transmitters and each transmitter state law satisfies the system (29). Thus, the indistinguishability property holds. $\blacksquare$

4.2.2 Solution to the mean-field response problem

We are now in position to define the solution concept of the SDE (29).

Definition 5 Let $T, T' > 0$ such that $[T, T']$ is the horizon of the game. We say that the state distribution $m_t(\underline{s})$ is a weak solution to the state dynamics (29) if m_t is integrable over $[T, T']$ and for any infinitely continuously differentiable function ϕ over $(T, T') \times \mathbb{R}^3$ with compact support (test function), one has

$$\mathbb{E}_{m_T}[\phi_T(\cdot)] + \int_T^{T'} \mathbb{E}_{m_t} \left[\partial_t \phi_t(\cdot) - \langle \nabla_{\underline{s}} \phi_t, D^*(t, \underline{s}) \rangle_{\mathbb{R}^2} + \frac{\eta^2}{2} \Delta_{\underline{s}} \phi_t \right] = 0, \quad (34)$$

where $\mathbb{E}_{m_t}$ is the expectation with the respect to m_t and $D^*(t, \underline{s})$ is the drift vector $[-\alpha(t, \underline{s}), \frac{1}{2}(\underline{\mu} - \underline{h}(t))]^T$.

Assuming that D^* is sufficiently regular in time t and in state $\underline{s}$, we examine the existence of a solution $\underline{s}(\cdot)$ which is $\mathcal{F}$ -adapted,

$$\mathbb{E} \left[\sup_{t \in [T, T']} |\underline{s}(t)|^2 \right] < +\infty, \quad (35)$$

and $\underline{s}(\cdot)$ has continuous paths.

A consequence of Itô's formula (see e.g. Karatzas and Shreve, 1991) states that the law of the SDE starting from distribution m_T is a weak solution of the partial differential equation (29). We deduce, from the Definition 5, the equation satisfied by the distribution of the states, i.e., *Fokker-Planck-Kolmogorov forward equation* is given by:

$$\frac{\partial m_t}{\partial t} - \frac{\partial}{\partial E}(m_t \alpha) + \text{div}_{\underline{h}}(m_t \frac{1}{2}(\underline{\mu} - \underline{h})) - \frac{\eta^2}{2} \Delta_{\underline{h}} m_t = 0. \quad (36)$$

With the new parameters of the game defined as:

$$\hat{I}(t, m_t) = \int_{\underline{s}} |\underline{h}|^2 \alpha(t, \underline{s}) m_t(d\underline{s}), \quad (37)$$

$$\hat{\gamma}(\underline{s}(t), m_t) = \frac{p(t) |\underline{h}(t)|^2}{\sigma^2 + \hat{I}(t, m_t)}, \quad (38)$$

$$\hat{u}(t) = \frac{Rf(\hat{\gamma}(\underline{s}(t), m_t))}{p(t)} =: \hat{r}(\underline{s}(t), p(t), m_t), \quad (39)$$

we can formulate the mean-field response problem in which each generic user best-responds to the mean-field:

$$\hat{v}_T = \sup_{p(T \rightarrow T')} \mathbb{E} \left[q(\underline{s}(T')) + \int_T^{T'} \hat{r}(\underline{s}(t), p(t), m_t^*) dt \right], \quad (40)$$

where m_t^* is the mean-field optimal trajectory and m_T is assumed to be given. A solution of the mean-field response problem is a solution of

$$\begin{cases} \frac{\partial \hat{v}_t}{\partial t} + \tilde{H}(\underline{s}(t), \frac{\partial \hat{v}_t}{\partial E}, m_t) + \frac{1}{2} \langle \underline{\mu} - \underline{h}, \nabla_{\underline{h}} \hat{v}_t \rangle_{\mathbb{R}^2} + \frac{\eta^2}{2} \Delta_{\underline{h}} \hat{v}_t = 0, \\ \frac{\partial m_t}{\partial t} + \frac{\partial}{\partial E}(m_t \frac{\partial}{\partial u} \tilde{H}(\underline{s}(t), \frac{\partial \hat{v}_t}{\partial E}, m_t)) + \text{div}_{\underline{h}}(m_t \frac{1}{2}(\underline{\mu} - \underline{h})) = \frac{\eta^2}{2} \Delta_{\underline{h}} m_t, \end{cases} \quad (41)$$

with $\hat{v}_{T'} = q(\underline{s}(T'))$, m_T known and

$$\tilde{H}(\underline{s}, u', m) = \sup_p \{ \hat{r}(\underline{s}, p, m) - p \cdot u' \}. \quad (42)$$

As for the SDG case, the first equation is a Hamilton-Jacobi-Bellman-Fleming equation. But it is now coupled with a Fokker-Planck-Kolmogorov equation. The former one is a backward equation whereas the latter one is a forward equation. The other main difference between the SDG and the MFG is that

in the former, each transmitter needs full knowledge of the channel states and the transmit powers of the other players to compute the outcome of the game, whereas in the latter only the knowledge of individual state and the mean-field is required to compute the outcome. In our model, although this mean-field cannot be directly known by the transmitters or the common receiver, it aggregates in the interference term which can be known by the receiver. Consequently, once the game is solved, given the interference at the receiver (which can be broadcast to every transmitters), its channel state and its battery state, each transmitter knows the power value it should use. The solution provides every transmitter a function of interference, channel state and battery state which output is the stable power control policy.

4.2.3 Uniqueness of the solution

Interestingly, a sufficient condition can be given for the solution of the mean-field response problem to be unique. First, we recall the definition of positiveness for an operator.

Definition 6 (Positiveness of an operator)

We say that the operator $\mathcal{O} : \mathcal{E} \rightarrow \mathcal{E}$ is positive (denoted by $\mathcal{O} \succ 0$) if

$$\forall x \neq 0_{\mathcal{E}}, \langle x, \mathcal{O}x \rangle_{\mathcal{E}} > 0,$$

where $0_{\mathcal{E}}$ is the neutral element of $\mathcal{E}$.

Proposition 4 (Uniqueness of the mean-field response problem solution) *A sufficient condition for the uniqueness of the solution to the mean-field response problem is for all triplet $(\underline{s}, u', m) \in \mathbb{R}^3 \times \mathbb{R} \times \mathbb{P}(\mathbb{R}^3)$,*

$$\begin{aligned} -\frac{1}{m} \frac{\partial}{\partial m} \tilde{H}(\underline{s}, u', m) &\succ 0, \\ \frac{\partial^2}{\partial u'^2} \tilde{H}(\underline{s}, u', m) &> 0, \\ \frac{\partial^2}{\partial m \partial u'} \tilde{H}(\underline{s}, u', m) - \frac{1}{2} \frac{(\frac{\partial^2}{\partial m \partial u'} \tilde{H}(\underline{s}, u', m))^2}{\frac{\partial^2}{\partial u'^2} \tilde{H}(\underline{s}, u', m)} &\succ 0. \end{aligned} \tag{43}$$

The proof of this proposition is given in Appendix B.

5 Conclusion

This paper provides a stochastic differential game formulation of the energy-efficient power control problem initially introduced in Goodman and Mandayam (2000). This formulation allows one to better optimize the global efficiency of the network, account for long-term energy constraints, and take into account propagation effects such as time correlation for the channel gains. The problem is that this model becomes intractable when the number of transmitters

becomes large. Instead of seeing large networks as a curse, they can be seen as a blessing since under the large system assumption, the game can be approximated by a mean-field game. Under the assumption of individual state information, the idea is precisely to exploit the large number of players to simplify the analysis. The authors believe the present paper provides several interesting results to go into this direction but must admit that the numerical analysis and the design aspect still require a lot of efforts to make this approach more practical (in the same way as random matrix theory was initially introduced in the wireless literature by Tse and Hanly (1999) and shown to be of practical interest later on by e.g., Dumont et al (2010)). Nonetheless, this paper provides several interesting results on the mean-field game approach. Under the exchangeability assumption, the stochastic differential game is shown to converge to a mean-field game as the number of players increases. This new game simplifies and even makes possible the equilibrium analysis since the equilibrium derivation only requires the knowledge of the individual state and the mean-field to solve a system of two equations. To be more precise, each transmitter needs to know its channel state, its battery state and the instantaneous interference it undergoes. For this, the receiver only needs to feed back the instantaneous interference. In the mean-field model, this instantaneous interference is the same for all the transmitters. Hence, the required signal is a broadcast to all the transmitters. Remarkably, this signal is fully scalable since, in theory (up to quantization effects), the amount of signalling does not depend of the number of transmitters. This framework allows us to derive simple sufficient conditions for the existence and uniqueness of an equilibrium power control policy.

A Proof of Proposition 1

Both results can be proven by using Ito's formula (see e.g. Karatzas and Shreve, 1991). For the mean, from (5), one has

$$d\mathbb{E}[\underline{h}_i(t)] = \frac{1}{2}(\underline{\mu} - \mathbb{E}[\underline{h}_i(t)])dt, \quad (44)$$

then $\mathbb{E}[\underline{h}_i(t)] = \underline{\mu}(1 - e^{-\frac{t}{2}}) + \underline{h}_i(0)e^{-\frac{t}{2}}$. The limit when t goes to $+\infty$ writes

$$\lim_{t \rightarrow +\infty} \mathbb{E}[\underline{h}_i] = \underline{\mu}. \quad (45)$$

For the variance, assume that for the two components x_i and y_i of h_i :

$$\begin{aligned} dx_i(t) &= g_x(t)dt + \eta d\mathbb{W}_x, \\ dy_i(t) &= g_y(t)dt + \eta d\mathbb{W}_y, \end{aligned} \quad (46)$$

with $\mathbb{W}_x$ and $\mathbb{W}_y$ two independent Wiener processes of dimension 1. Then

$$dx_i(t)^2 = (2x_i(t)g_x(t) + \eta^2)dt + 2x_i(t)\eta d\mathbb{W}_x, \quad (47)$$

and

$$d\mathbb{E}[x_i(t)^2] = \mathbb{E}[(2x_i(t)g_x(t) + \eta^2)]dt. \quad (48)$$

If $g_x(t) = 0$, $\mathbb{E}[x_i(t)^2] = x_i(0)^2 + \eta^2 t$ and $\lim_{t \rightarrow \infty} \mathbb{E}[x_i(t)^2] = \infty$. That is the reason why a deterministic term is needed in (5). With $g_x(t) = \frac{1}{2}(\mu_x - x_i(t))$, one has

$$d\mathbb{E}[x_i(t)^2] = \mathbb{E}[2x_i(t)\frac{1}{2}(\mu_x - x_i(t)) + \eta^2]dt, \quad (49)$$

then

$$\frac{d\mathbb{E}[x_i(t)^2]}{dt} = -\mathbb{E}[x_i(t)^2] + \mu_x \mathbb{E}[x_i(t)] + \eta^2. \quad (50)$$

The solution of this differential equation has the form

$$\begin{aligned} \mathbb{E}[x_i(t)^2] &= \left(x_i(0)^2 + \int_0^t (\mu_x \mathbb{E}[x_i(t')] + \eta^2) e^{t'} dt' \right) e^{-t}, \\ &= \left[(\mu_x^2 + \eta^2) e^{t'} + 2\mu_x (x_i(0) - \mu_x) e^{\frac{t'}{2}} \right]_0^t e^{-t}, \\ &= (\mu_x^2 + \eta^2)(1 - e^{-t}) + 2\mu_x (x_i(0) - \mu_x)(e^{-\frac{t}{2}} - e^{-t}), \end{aligned} \quad (51)$$

thus

$$\lim_{t \rightarrow +\infty} \mathbb{E}[x_i(t)^2] = \mu_x^2 + \eta^2. \quad (52)$$

The analogous is true for y_i . Hence we have

$$\lim_{t \rightarrow +\infty} \mathbb{E}[|h_i(t)|^2] = |\underline{\mu}|^2 + 2\eta^2, \quad (53)$$

and finally

$$\lim_{t \rightarrow +\infty} \mathbb{E}[|h_i(t)|^2] - \mathbb{E}[|h_i(t)|]^2 = 2\eta^2. \quad (54)$$

Regarding to the probability density functions, applying the Kolmogorov forward equation to the state $\underline{h}_i$ with the dynamics given in (5), one has for the component x_i

$$\begin{aligned} \frac{\partial m_x(x_i, t)}{\partial t} &= -\frac{\partial}{\partial x_i} \left[m_x(x_i, t) \frac{1}{2}(\mu_x - x_i) \right] + \frac{\eta^2}{2} \frac{\partial^2 m_x(x_i, t)}{\partial x_i^2}, \\ &= \frac{1}{2} m_x(x_i, t) - \frac{1}{2}(\mu_x - x_i) \frac{\partial m_x(x_i, t)}{\partial x_i} + \frac{\eta^2}{2} \frac{\partial^2 m_x(x_i, t)}{\partial x_i^2}. \end{aligned} \quad (55)$$

The stationary case gives

$$0 = \frac{1}{2} m_x(x_i) - \frac{1}{2}(\mu_x - x_i) \frac{\partial m_x(x_i)}{\partial x_i} + \frac{\eta^2}{2} \frac{\partial^2 m_x(x_i)}{\partial x_i^2}. \quad (56)$$

One can check that $\hat{m}_x(x_i) = \frac{1}{\eta\sqrt{2\pi}} e^{-\frac{(x_i - \mu_x)^2}{2\eta^2}}$ is a solution of (56). This is the stationary density of x_i . The analogous can also be written for y_i : $\hat{m}_y(y_i) = \frac{1}{\eta\sqrt{2\pi}} e^{-\frac{(y_i - \mu_y)^2}{2\eta^2}}$. ■

B Proof of Proposition 4

The proof follows the the same principle as in chapter *Risk-sensitive mean-field games* in the notes *Mean-field stochastic games* by Tembine. Only the sketch of the proof is given here. To prove the uniqueness of the solution, we suppose that there exists two solutions $(v_{1,t}, m_{1,t}), (v_{2,t}, m_{2,t})$ of the above system. We want to find a sufficient condition under which the quantity $\int_{\underline{s}} (v_{2,t}(\underline{s}) - v_{1,t}(\underline{s}))(m_{2,t}(\underline{s}) - m_{1,t}(\underline{s})) d\underline{s}$ is monotone in time, which is not possible.

$$\begin{cases} m_{1,T}(\underline{s}) = m_{2,T}(\underline{s}), \\ v_{1,T'}(\underline{s}) = v_{2,T'}(\underline{s}), \end{cases} \quad (57)$$

Compute the time derivative

$$\begin{aligned}
S_t &= \frac{d}{dt} \left(\int_{\underline{s}} (v_{2,t}(\underline{s}) - v_{1,t}(\underline{s}))(m_{2,t}(\underline{s}) - m_{1,t}(\underline{s})) d\underline{s} \right), \\
&= \int_{\underline{s}} \left(\frac{\partial v_{2,t}}{\partial t}(\underline{s}) - \frac{\partial v_{1,t}}{\partial t}(\underline{s}) \right) (m_{2,t}(\underline{s}) - m_{1,t}(\underline{s})) d\underline{s}, \\
&\quad + \int_{\underline{s}} (v_{2,t}(\underline{s}) - v_{1,t}(\underline{s})) \left(\frac{\partial m_{2,t}}{\partial t}(\underline{s}) - \frac{\partial m_{1,t}}{\partial t}(\underline{s}) \right) d\underline{s}.
\end{aligned} \tag{58}$$

To express the first term, the difference between the two HJBF equations is taken and multiplied by $m_{2,t} - m_{1,t}$:

$$\begin{aligned}
&\int_{\underline{s}} \left(\frac{\partial v_{2,t}}{\partial t} - \frac{\partial v_{1,t}}{\partial t} \right) (m_{2,t} - m_{1,t}) d\underline{s} = \\
&\int_{\underline{s}} \tilde{H}(\underline{s}(t), \frac{\partial v_{1,t}}{\partial E}, m_{1,t}) (m_{2,t} - m_{1,t}) d\underline{s} - \int_{\underline{s}} \tilde{H}(\underline{s}(t), \frac{\partial v_{2,t}}{\partial E}, m_{2,t}) (m_{2,t} - m_{1,t}) d\underline{s} \\
&+ \int_{\underline{s}} \frac{\eta^2}{2} \frac{\partial^2 v_{1,t}}{\partial h^2} (m_{2,t} - m_{1,t}) d\underline{s} - \int_{\underline{s}} \frac{\eta^2}{2} \frac{\partial^2 v_{2,t}}{\partial h^2} (m_{2,t} - m_{1,t}) d\underline{s} \\
&+ \frac{1}{2} \int_{\underline{s}} \langle \underline{\mu} - \underline{h}(t), \nabla_{\underline{h}} v_{1,t} \rangle_{\mathbb{R}^2} (m_{2,t} - m_{1,t}) d\underline{s} - \frac{1}{2} \int_{\underline{s}} \langle \underline{\mu} - \underline{h}(t), \nabla_{\underline{h}} v_{2,t} \rangle_{\mathbb{R}^2} (m_{2,t} - m_{1,t}) d\underline{s}.
\end{aligned} \tag{59}$$

For the second term, the difference between the two FPK equations is taken and multiplied by $v_{2,t} - v_{1,t}$:

$$\begin{aligned}
&\int_{\underline{s}} \left(\frac{\partial m_{2,t}}{\partial t} - \frac{\partial m_{1,t}}{\partial t} \right) (v_{2,t} - v_{1,t}) d\underline{s} = \\
&- \int_{\underline{s}} \frac{\partial}{\partial E} (m_{2,t} \frac{\partial}{\partial u'} \tilde{H}(\underline{s}(t), \frac{\partial v_{2,t}}{\partial E}, m_{2,t})) (v_{2,t} - v_{1,t}) d\underline{s} \\
&+ \int_{\underline{s}} \frac{\partial}{\partial E} (m_{1,t} \frac{\partial}{\partial u'} \tilde{H}(\underline{s}(t), \frac{\partial v_{1,t}}{\partial E}, m_{1,t})) (v_{2,t} - v_{1,t}) d\underline{s} \\
&+ \int_{\underline{s}} \frac{\eta^2}{2} \frac{\partial^2 m_{2,t}}{\partial h^2} (v_{2,t} - v_{1,t}) d\underline{s} - \int_{\underline{s}} \frac{\eta^2}{2} \frac{\partial^2 m_{1,t}}{\partial h^2} (v_{2,t} - v_{1,t}) d\underline{s} \\
&+ \frac{1}{2} \int_{\underline{s}} \text{div}_{\underline{h}} (m_{2,t} (\underline{\mu} - \underline{h}(t))) (v_{2,t} - v_{1,t}) d\underline{s} - \frac{1}{2} \int_{\underline{s}} \text{div}_{\underline{h}} (m_{1,t} (\underline{\mu} - \underline{h}(t))) (v_{2,t} - v_{1,t}) d\underline{s}.
\end{aligned} \tag{60}$$

By integration by parts

$$\int_{\underline{s}} \text{div}_{\underline{s}}(k) \phi d\underline{s} = - \int_{\underline{s}} k \text{div}_{\underline{s}}(\phi) d\underline{s}, \tag{61}$$

then

$$\begin{aligned}
&\int_{\underline{s}} \left(\frac{\partial m_{2,t}}{\partial t} - \frac{\partial m_{1,t}}{\partial t} \right) (v_{2,t} - v_{1,t}) d\underline{s} = \\
&\int_{\underline{s}} (m_{2,t} \frac{\partial}{\partial u'} \tilde{H}(\underline{s}(t), \frac{\partial v_{2,t}}{\partial E}, m_{2,t})) \left(\frac{\partial v_{2,t}}{\partial E} - \frac{\partial v_{1,t}}{\partial E} \right) d\underline{s} \\
&- \int_{\underline{s}} (m_{1,t} \frac{\partial}{\partial u'} \tilde{H}(\underline{s}(t), \frac{\partial v_{1,t}}{\partial E}, m_{1,t})) \left(\frac{\partial v_{2,t}}{\partial E} - \frac{\partial v_{1,t}}{\partial E} \right) d\underline{s} \\
&+ \int_{\underline{s}} \frac{\eta^2}{2} \frac{\partial^2 m_{2,t}}{\partial h^2} (v_{2,t} - v_{1,t}) d\underline{s} - \int_{\underline{s}} \frac{\eta^2}{2} \frac{\partial^2 m_{1,t}}{\partial h^2} (v_{2,t} - v_{1,t}) d\underline{s} \\
&+ \frac{1}{2} \int_{\underline{s}} m_{2,t} \langle \underline{\mu} - \underline{h}(t), \nabla_{\underline{h}} (v_{2,t} - v_{1,t}) \rangle_{\mathbb{R}^2} d\underline{s} - \frac{1}{2} \int_{\underline{s}} m_{1,t} \langle \underline{\mu} - \underline{h}(t), \nabla_{\underline{h}} (v_{2,t} - v_{1,t}) \rangle_{\mathbb{R}^2} d\underline{s}.
\end{aligned} \tag{62}$$

The full derivative writes

$$\begin{aligned}
S_t &= \int_{\underline{s}} \tilde{H}(\underline{s}(t), \frac{\partial v_{1,t}}{\partial E}, m_{1,t})(m_{2,t} - m_{1,t}) d\underline{s} - \int_{\underline{s}} \tilde{H}(\underline{s}(t), \frac{\partial v_{2,t}}{\partial E}, m_{2,t})(m_{2,t} - m_{1,t}) d\underline{s} \\
&\quad - \int_{\underline{s}} m_{1,t} \frac{\partial}{\partial u'} \tilde{H}(\underline{s}(t), \frac{\partial v_{1,t}}{\partial E}, m_{1,t}) (\frac{\partial v_{2,t}}{\partial E} - \frac{\partial v_{1,t}}{\partial E}) d\underline{s} \\
&\quad + \int_{\underline{s}} m_{2,t} \frac{\partial}{\partial u'} \tilde{H}(\underline{s}(t), \frac{\partial v_{2,t}}{\partial E}, m_{2,t}) (\frac{\partial v_{2,t}}{\partial E} - \frac{\partial v_{1,t}}{\partial E}) d\underline{s}.
\end{aligned} \tag{63}$$

We now introduce

$$m_{\lambda,t} = (1 - \lambda)m_{1,t} + \lambda m_{2,t} = m_{1,t} + \lambda(m_{2,t} - m_{1,t}), \tag{64}$$

and in the same way

$$v_{\lambda,t} = (1 - \lambda)v_{1,t} + \lambda v_{2,t}. \tag{65}$$

We study the auxiliary integral

$$\begin{aligned}
C_\lambda &= \int_{\underline{s}} \tilde{H}(\underline{s}(t), \frac{\partial v_{1,t}}{\partial E}, m_{1,t})(m_{\lambda,t} - m_{1,t}) d\underline{s} - \int_{\underline{s}} \tilde{H}(\underline{s}(t), \frac{\partial v_{\lambda,t}}{\partial E}, m_{\lambda,t})(m_{\lambda,t} - m_{1,t}) d\underline{s} \\
&\quad - \int_{\underline{s}} m_{1,t} \frac{\partial}{\partial u'} \tilde{H}(\underline{s}(t), \frac{\partial v_{1,t}}{\partial E}, m_{1,t}) (\frac{\partial v_{\lambda,t}}{\partial E} - \frac{\partial v_{1,t}}{\partial E}) d\underline{s} \\
&\quad + \int_{\underline{s}} m_{\lambda,t} \frac{\partial}{\partial u'} \tilde{H}(\underline{s}(t), \frac{\partial v_{\lambda,t}}{\partial E}, m_{\lambda,t}) (\frac{\partial v_{\lambda,t}}{\partial E} - \frac{\partial v_{1,t}}{\partial E}) d\underline{s},
\end{aligned} \tag{66}$$

which derivative is:

$$\begin{aligned}
\frac{d}{d\lambda} \left(\frac{C_\lambda}{\lambda} \right) &= - \int_{\underline{s}} \frac{\partial}{\partial m} \tilde{H}(\underline{s}(t), \frac{\partial v_{\lambda,t}}{\partial E}, m_{\lambda,t})(m_{2,t} - m_{1,t})^2 d\underline{s} \\
&\quad + \int_{\underline{s}} m_{\lambda,t} \frac{\partial^2}{\partial u'^2} \tilde{H}(\underline{s}(t), \frac{\partial v_{\lambda,t}}{\partial E}, m_{\lambda,t}) (\frac{\partial v_{2,t}}{\partial E} - \frac{\partial v_{1,t}}{\partial E})^2 d\underline{s} \\
&\quad + \int_{\underline{s}} m_{\lambda,t} \frac{\partial^2}{\partial m \partial u'} \tilde{H}(\underline{s}(t), \frac{\partial v_{\lambda,t}}{\partial E}, m_{\lambda,t})(m_{2,t} - m_{1,t}) (\frac{\partial v_{2,t}}{\partial E} - \frac{\partial v_{1,t}}{\partial E}) d\underline{s}.
\end{aligned} \tag{67}$$

Note that $\frac{\partial}{\partial m} \tilde{H}(\cdot)$ and $\frac{\partial^2}{\partial m \partial u'} \tilde{H}(\cdot)$ are functional derivatives. They are defined such that for all $m' \in \mathbb{P}(\mathbb{R}^3)$

$$\begin{aligned}
\langle \frac{\partial}{\partial m} \tilde{H}(\underline{s}(t), \frac{\partial v_{\lambda,t}}{\partial E}, m_{\lambda,t}), m' - m_{\lambda,t} \rangle_{\mathbb{P}(\mathbb{R}^3)} &= \\
\lim_{t \rightarrow 0} \frac{\tilde{H}(\underline{s}(t), \frac{\partial v_{\lambda,t}}{\partial E}, m_{\lambda,t} + t(m' - m_{\lambda,t})) - \tilde{H}(\underline{s}(t), \frac{\partial v_{\lambda,t}}{\partial E}, m_{\lambda,t})}{t},
\end{aligned} \tag{68}$$

and

$$\begin{aligned}
\langle \frac{\partial^2}{\partial m \partial u'} \tilde{H}(\underline{s}(t), \frac{\partial v_{\lambda,t}}{\partial E}, m_{\lambda,t}), m' - m_{\lambda,t} \rangle_{\mathbb{P}(\mathbb{R}^3)} &= \\
\lim_{t \rightarrow 0} \frac{\frac{\partial \tilde{H}}{\partial u'}(\underline{s}(t), \frac{\partial v_{\lambda,t}}{\partial E}, m_{\lambda,t} + t(m' - m_{\lambda,t})) - \frac{\partial \tilde{H}}{\partial u'}(\underline{s}(t), \frac{\partial v_{\lambda,t}}{\partial E}, m_{\lambda,t})}{t}.
\end{aligned} \tag{69}$$

A sufficient condition for the uniqueness of the solution to the mean-field response problem is the monotonicity of the operator associated to

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & a_{22} \end{pmatrix}$$

with

$$\begin{aligned} A_{11} &= -\frac{1}{m} \frac{\partial}{\partial m} \tilde{H}, \\ A_{12} &= A_{21} = \frac{1}{2} \frac{\partial^2}{\partial m \partial u'} \tilde{H}, \\ a_{22} &= \frac{\partial^2}{\partial u'^2} \tilde{H}. \end{aligned} \quad (70)$$

This is true if

$$\begin{aligned} A_{11} &\succ 0, \quad a_{22} > 0, \\ A_{12} - \frac{A_{12}^2}{a_{22}} &\succ 0. \end{aligned} \quad (71)$$

■

References

- Agarwal M, Honig M (2012) Adaptive training for correlated fading channels with feedback. *IEEE Transactions on Information Theory* 58(8):5398–5417
- Basar T, Olsder GJ (1999) Dynamic noncooperative game theory. 2nd edition, Classics in Applied Mathematics, SIAM, Philadelphia
- Belmega EV, Lasaulce S (2011) Energy-efficient precoding for multiple-antenna terminals. *IEEE Trans on Signal Processing* 59(1):329–340
- Belmega EV, Lasaulce S, Debbah M (2009) Power allocation games for MIMO multiple access channels with coordination. *IEEE Trans on Wireless Communications* 8(5):3182–3192
- Bonneau N, Debbah M, Altman E, Hjørungnes A (2008) Non-atomic games for multi-user system. *IEEE Journal on Selected Areas in Communications* 26(7):1047–1058
- Bressan A (2010) Noncooperative differential games. A tutorial
- Buzzi S, Saturnino D (2011) A game-theoretic approach to energy-efficient power control and receiver design in cognitive CDMA wireless networks. *Journal of Selected Topics in Signal Processing* 5(1):137–150
- Cover TM, Thomas JA (1991) Elements of information theory. Wiley-Interscience
- Dumont J, Hachem W, Lasaulce S, Loubaton P, , Najim J (2010) On the capacity achieving covariance matrix of rician mimo channels: an asymptotic approach. *IEEE Trans on Info Theory* 56(3):1048–1069
- Evans L (2010) Partial differential equations. American Mathematical Society
- Fette BA (2006) Cognitive Radio Technology. Elsevier
- Fleming W, Soner H (1993) Controlled Markov processes and viscosity solutions. Applications of mathematics, Springer-Verlag
- Foschini GJ, Miljanic Z (1993) A simple distributed autonomous power control algorithm and its convergence. *IEEE Trans on Vehicular Technology* 42(4):641–646
- Goodman DJ, Mandayam NB (2000) Power control for wireless data. *IEEE Personal Communications* 7:48–54
- Gupta P, Kumar PR (1997) A system and traffic dependent adaptive routing algorithm for ad hoc networks. In: *IEEE Conf. on Decision and Control (CDC)*, San Diego, USA, pp 2375–2380
- Karatzas I, Shreve S (1991) Brownian motion and stochastic calculus. Springer
- Lasaulce S, Tembine H (2011) Game Theory and Learning for Wireless Networks : Fundamentals and Applications. Academic Press
- Lasaulce S, Hayel Y, Azouzi RE, Debbah M (2009) Introducing hierarchy in energy games. *IEEE Trans on Wireless Comm* 8(7):3833–3843
- Lasry JM, Lions PL (2007) Mean field games. *Jpn J Math* 2(1):229–260
- Le Treust M, Lasaulce S (2010) A repeated game formulation of energy-efficient decentralized power control. *IEEE Trans on Wireless Communications*

- Mériaux F, Hayel Y, Lasaulce S, AGarnaev (2011a) Long-term energy constraints and power control in cognitive radio networks. In: IEEE Proc. of the 17th International Conference on Digital Signal Processing (DSP), Corfu, Greece
- Mériaux F, Treust ML, Lasaulce S, Kieffer M (2011b) Energy-efficient power control strategies for stochastic games. In: IEEE Proc. of the 17th International Conference on Digital Signal Processing (DSP), Corfu, Greece
- Meshkati F, Poor HV, Schwartz SC, Narayan BM (2005) An energy-efficient approach to power control and receiver design in wireless data networks. *IEEE Trans on Communications* 53:1885–1894
- Meshkati F, Chiang M, Poor HV, Schwartz SC (2006) A game-theoretic approach to energy-efficient power control in multi-carrier CDMA systems. *IEEE Journal on Selected Areas in Communications* 24(6):1115–1129
- Mitola J, Maguire GQ (1999) Cognitive radio: making software radios more personal. *IEEE Personal Communications* 6(4):13–18
- Olama M, Djouadi S, Charalambous C (2006) Stochastic power control for time-varying long-term fading wireless networks. *EURASIP Journal on Applied Signal Processing (JASP)* Vol. 2006:1–13
- Rodriguez V (2003) An analytical foundation for resource management in wireless communication. In: IEEE Proc. of Globecom, pp 898–902
- Saraydar CU, Mandayam NB, Goodman DJ (2002) Efficient power control via pricing in wireless data networks. *IEEE Trans on Communications* 50(2):291–303
- Tembine H, Huang M (2011) Mean field difference games: McKean-Vlasov dynamics. CDC-ECC, 50th IEEE Conference on Decision and Control and European Control Conference
- Tembine H, Lasaulce S, Jungers M (2010) Joint power control-allocation for green cognitive wireless networks using mean field theory. In: IEEE Proc. of the 5th Intl. Conf. on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM), Cannes, France
- Tse D, Hanly S (1999) Linear multiuser receivers: Effective interference, effective bandwidth and user capacity. In: *IEEE Transactions on Information Theory*, vol 45, pp 641–657
- Tulino A, Verdú S (2004) *Random Matrices and Wireless Communications. Foundations and trends in communication and information theory*, NOW, The Essence of Knowledge
- Yates RD (1995) A framework for uplink power control in cellular radio systems. *IEEE Journal on Selected Areas in Communications* 13(7):1341–1347

B.2 IEEE-ASILOMAR-2012

F. Mériaux, V. Varma et S. Lasaulce. *Mean Field Energy Games in Wireless Networks*. In IEEE Conference on Signals, Systems and Computers (ASILOMAR), pages 671-675, 2012.

Mean Field Energy Games in Wireless Networks

François Mériaux*, Vineeth Varma*,[†] Samson Lasaulce*

*Laboratoire des Signaux et Systèmes – LSS (CNRS-SUPELEC-Paris Sud), Gif-sur-Yvette, France

Email: {meriaux,varma,lasaulce}@lss.supelec.fr

[†]Orange Labs, Issy Les Moulineaux, France

Abstract—This work tackles the problem of energy-efficient distributed power control in wireless networks with a large number of transmitters. The problem is modeled by a dynamic game. Each transmitter-receiver communication is characterized by a state given by the available energy and/or the individual channel state and whose evolution is governed by certain dynamics. Since equilibrium analysis in such a (stochastic) game is generally difficult and even impossible, the problem is approximated by exploiting the large system assumption. Under an appropriate exchangeability assumption, the corresponding mean field game is well defined and studied in detail for special cases. The main contribution of this work is to show how mean field games can be applied to the problem under investigation and provide illustrative numerical results. Our results indicate that this approach can lead to significant gains in terms of energy-efficiency at the resulting equilibrium.

I. INTRODUCTION

We study distributed wireless networks, in which mobile terminals have the liberty to choose their own power policies. This can be due to the absence of a central node to control the terminals in the network or due to complexity issues when there is a large number of terminals. Such a scenario can occur with ad hoc networks [7], unlicensed band communications, and cognitive radio [5], [15]. Modeling the terminals as rational agents, who choose their power control policies to maximize some utilities leads to applying game theory on the problem [8], [10]. In the problem under study, the wireless network model is a multiple access channel (MAC) and the channel access method is code division multiple access (CDMA). In this network, the mobile terminals aim to maximize their expected energy-efficiencies over a given time duration. For this purpose, they must adapt their power control policies to varying channel conditions and decreasing energy in their batteries.

By energy-efficiency, we mean the number of successfully decoded bits at the receiver per Joule consumed at the transmitter, as defined in [6]. In this seminal paper and related works [14], [3], [9], [4], [1], the power control problem is modeled by a sequence of static games independent from one stage to another. But this approach does not capture the interactions that are present among the players when a game is repeated. In [12], it is shown that modeling the problem by a repeated game can lead to more efficient equilibrium power control policies (in the sense of Pareto) than the Nash equilibrium from the static formulation. However, this repeated model uses a normalized stage game which does not depend on the channels realizations. One of the motivations

of our work is to account for the impact of channels realizations by modeling the problem with a stochastic game. Moreover, we also account for long-term energy constraint in the terminal since the remaining energy of the battery decreases when power is consumed. Precisely, the energy-efficient power control problem in MAC under long-term energy constraints is modeled by a stochastic differential game (SDG). But the problem of characterizing the performance of distributed networks modeled by SDG becomes hard and even impossible when the number of players becomes large. The same statement holds for determining individually optimal control strategies. In [13], this problem is overcome by the use of mean field games (MFG). MFG [11] represent a way of approximating a stochastic differential (or difference) game, by a much more tractable model. Under the assumption of individual state information, the idea is precisely to exploit as an opportunity the fact that the number of players is large to simplify the analysis. Typically, instead of depending on the actions and states of all the players, the mean field utility of a player only depends on his own action and state, and depends on the others through an mean field. The main contribution of this paper is to show that the model developed in [13] can be particularized to cases in which, it is possible to derive equilibrium power control policies. These policies are illustrated with some numerical results and they are compared with the equilibrium from the static game formulation and the equilibrium of the repeated discounted game studied in [12].

Our paper is structured as follows. We describe the studied wireless network and the evolution laws for the channels and the energy in Sec. II. In Sec. III, the problem of power control is modeled as an SDG and is shown to converge to an MFG under given conditions. Additionally, the regime of large energy budgets and the regime of quasi-static channels are discussed. Resulting equilibrium power control strategies are compared to the classical power control policies from the static game and the discounted repeated game in Sec. IV.

Notations: In the following, $\nabla_{\underline{x}}f$ and $\Delta_{\underline{x}}f$ respectively represent the gradient and the Laplacian of the function f w.r.t. the vector $\underline{x}$. The divergence operator w.r.t. the vector $\underline{x}$ is denoted by $\text{div}_{\underline{x}}$. The scalar product in the space $\mathbb{A}$ is represented by $\langle \cdot, \cdot \rangle_{\mathbb{A}}$.

II. SYSTEM MODEL

We consider one cell with a single base station and K mobile terminals. Since the uplink power control problem is addressed, mobile terminals are transmitters and the base

station is the receiver. The radio resource is used as a MAC. Consequently the transmitters interfere with each other. Each transmitter sends a signal to a common receiver and has to choose the power level of the transmitted signal. In order to optimize its individual energy-efficiency, i.e., the ratio of its throughput to its transmit power, each transmitter adapts its power level. This choice depends on the quality of the channel between the transmitter and the receiver, the power levels chosen by the other transmitter and the energy remaining. The set of transmitters is denoted by $\mathcal{K} = \{1, \dots, K\}$ and for each transmitter $k \in \mathcal{K}$, the signal-to-interference plus noise ratio (SINR) writes

$$\gamma_k(p_1, \dots, p_K) = \frac{p_k |h_k|^2}{I_k + \sigma^2}, \quad (1)$$

where $p_k \in \mathcal{P}_k$ and $|h_k|^2$ are the power level and the channel gain of transmitter k , respectively. The variance of the noise is represented by σ^2 . The interference term is denoted by $I_k = \frac{1}{N} \sum_{j \in \mathcal{K}, j \neq k} p_j |h_j|^2$ with N a processing gain due to interference management at the receiver. For example, in CDMA systems, N represents the spreading factor.

The instantaneous energy-efficiency, in bit/joule is defined as

$$u_k(p_1, \dots, p_K) = \frac{Rf(\gamma_k(p_1, \dots, p_K))}{p_k} [\text{bit/J}], \quad (2)$$

where R is a constant rate of the transmitter. The function f is the probability of having no outage which takes its values in $[0, 1]$ and depends on the SINR.

Two parameters define the state of each transmitter: the channel coefficients to the receiver and the remaining energy in its battery. The evolution of these parameters is modeled by the two following laws. The channel evolution law is

$$dh_k(t) = \frac{1}{2}(\mu - h_k(t))dt + \eta d\mathbb{W}_k(t), \quad (3)$$

where $\mu \in \mathbb{C}$ and $\eta \in \mathbb{R}$ are constants related to the channel statistics and $\mathbb{W}_k(t)$ are mutually independent Wiener processes. Depending on the value of η , this law can model slow-fading or fast-fading. Additional properties about the asymptotic behavior of the channels are given in [13].

The evolution law of the remaining energy in the battery is given by

$$dE_k(t) = -p_k(t)dt. \quad (4)$$

This means that the energy of the battery decreases with the transmit power consumption.

III. FROM THE STOCHASTIC DIFFERENTIAL GAME FORMULATION TO THE MEAN FIELD GAME

A. Stochastic Differential Game

With the model defined in the previous section, the problem of each transmitter maximizing its expected energy-efficiency over a given time interval can be addressed with an SDG. Time is assumed to be continuous, i.e., $t \in \mathbb{R}$. The time horizon of the game is finite, it is the interval ranging from T to T' .

Definition 1 (SDG model of the power control problem): The stochastic differential power control game is defined by the 5-tuple

$\mathcal{G} = (\mathcal{K}, \{\mathcal{P}_k\}_{k \in \mathcal{K}}, \{\mathcal{X}_k\}_{k \in \mathcal{K}}, \{\mathcal{S}_k\}_{k \in \mathcal{K}}, \{U_k\}_{k \in \mathcal{K}})$ where:

- $\mathcal{K} = \{1, \dots, K\}$ is the set of players. Here, the players correspond to the transmitters.
- $\mathcal{P}_k$ is the set of actions of player $k \in \mathcal{K}$. Here, the action set corresponds to the interval of transmit power values.
- $\mathcal{X}_k$ is the state space of player $k \in \mathcal{K}$. The game state for player $k \in \mathcal{K}$ at time t is defined by $\underline{X}_k(t) = [E_k(t), h_k(t)]^\top$.
- $\mathcal{S}_k$ is the set of feedback control policies for player $k \in \mathcal{K}$. A control policy will be denoted by $p_k(T \rightarrow T')$ which is a function of time between T and T' two reals such that $T' \geq T$;
- the average utility function U_k is defined by:

$$U_k(p(T \rightarrow T')) = \mathbb{E} \left[\int_T^{T'} u_k(p(t), \underline{X}(t)) dt + q(\underline{X}(T')) \right], \quad (5)$$

where $p(T \rightarrow T') = (p_1(T \rightarrow T'), \dots, p_K(T \rightarrow T'))$ is the control strategy profile, $\underline{X}(t) = [\underline{X}_1(t), \dots, \underline{X}_K(t)]$ is the state profile, $q(\underline{X}(T'))$ is the utility at the final state, and u_k is the instantaneous utility.

Even if it can be proven that a Nash equilibrium exists in this game under given conditions [13], obtaining the expression of an equilibrium requires to solve a system of $2K$ coupled equations. Consequently, the complexity of such a system makes its resolution impossible for K large. This makes us consider the MFG associated with the problem to overcome this complexity issue.

B. The Mean field game analysis ($K \rightarrow +\infty$)

1) *Assumptions:* When the number of players goes to infinity, under the assumption of the exchangeability of the players of the game and the convergence of the interference term, the SDG can be proven to converge to an MFG. The exchangeability property (see [13] for more details) is ensured if each player only knows its individual state and implements an homogeneous admissible control: $p_k(t) = \alpha(t, \underline{X}_k(t))$. A sufficient condition for the convergence of the interference term is $\lim_{K, N \rightarrow \infty} \frac{K}{N} = \theta > 0$. In this new game, the set of players is continuous and the generic state of a player is given by $\underline{s}(t) = [E(t), h(t)]^\top$, whose distribution is given by m_t (m_t is the mean field). The SINR can be rewritten as:

$$\hat{\gamma}(\underline{s}(t), m_t) = \frac{p(t)|h(t)|^2}{\sigma^2 + \hat{I}(t, m_t)}, \quad (6)$$

with $\hat{I}(t, m_t)$ the interference resulting from the continuum of other players

$$\hat{I}(t, m_t) = \int_{\underline{s}} |h|^2 \alpha(t, \underline{s}) m_t(d\underline{s}), \quad (7)$$

where $\alpha(t, \underline{s})$ denotes the generic power response at time t and state $\underline{s}$. The instantaneous utility writes

$$\hat{u}(p(t), \underline{s}(t), m_t) = \frac{Rf(\hat{\gamma}(\underline{s}(t), m_t))}{p(t)}. \quad (8)$$

The main advantage of the MFG formulation is that the utility of each player depends only on its own state $\underline{s}(t)$ and a common mean field m_t .

2) *Solution to the mean field best-response problem:* For the mean field optimal trajectory m_t^* , the best-response of the generic player is such that there exists an average utility:

$$\hat{v}_T = \sup_{p(T \rightarrow T')} \mathbb{E} \left[\int_T^{T'} \hat{u}(p(t), \underline{s}(t), m_t^*) dt + q(\underline{s}(T')) \right]. \quad (9)$$

Conversely, the control policy resulting from (9) has to lead to the distribution trajectory m_t^* . Consequently, a solution of the mean field response problem is a solution of the system

$$\begin{cases} \frac{\partial \hat{v}_t}{\partial t} + \tilde{H}(\underline{s}(t), \frac{\partial \hat{v}_t}{\partial E}, m_t) \\ + \frac{1}{2} \langle \mu - h, \nabla_h \hat{v}_t \rangle_{\mathbb{R}^2} + \frac{\eta^2}{2} \Delta_h \hat{v}_t = 0, \\ \frac{\partial m_t}{\partial t} + \frac{\partial}{\partial E} (m_t \frac{\partial}{\partial u'} \tilde{H}(\underline{s}(t), \frac{\partial \hat{v}_t}{\partial E}, m_t)) \\ + \text{div}_h (m_t \frac{1}{2} (\mu - h)) = \frac{\eta^2}{2} \Delta_h m_t, \end{cases} \quad (10)$$

with $\hat{v}_{T'} = q(\underline{s}(T'))$, m_T known and

$$\tilde{H}(\underline{s}, u', m) = \sup_p \{ \hat{u}(p, \underline{s}, m) - p \cdot u' \}. \quad (11)$$

The first equation is a Hamilton-Jacobi-Bellman-Fleming equation which gives the behavior of $\hat{v}_t$ for a given m_t . It is coupled with a Fokker-Planck-Kolmogorov equation which gives the behavior of m_t for a given $\hat{v}_t$. The former one is a backward equation whereas the latter one is a forward equation. It means that an initial condition for m_t and a final condition for $\hat{v}_t$ are needed to solve the system of two equations.

C. Two particular regimes of the mean field game

Solving (10) in the general case is complicated since the state of a transmitter includes both its energy and its channel coefficients. However, it is possible to study the solutions of the mean field response problem for these two parts separately.

1) *Large energy budgets:* If energy budgets are large enough, the variation of energy during the game can be neglected. Consequently, only channel coefficients can be considered as the state of the transmitter. In this case, the mean field problem reduces to

$$\begin{cases} \frac{\partial \hat{v}_t}{\partial t} + \sup_p \hat{u}(p, h, m_t) = 0, \\ \frac{\partial m_t}{\partial t} + \text{div}_h (m_t \frac{1}{2} (\mu - h)) = \frac{\eta^2}{2} \Delta_h m_t. \end{cases} \quad (12)$$

The first equation amounts to choosing the Nash equilibrium from [6] as the power control. The second equation only depends on the channels statistics and give the evolution of the distribution of the channels. The second equation is solved first and the solution m_t is inserted in the first equation to obtain the power control.

2) *Quasi-static channels:* Considering only energy as the state of a transmitter amounts to assume that the channel coefficients are constant during the time interval $[T, T']$. In this case, the mean field problem turns into

$$\begin{cases} \frac{\partial \hat{v}_t}{\partial t} + \sup_p \{ \hat{u}(p, m_t) - p \frac{\partial \hat{v}_t}{\partial E} \} = 0, \\ \frac{\partial m_t}{\partial t} - \frac{\partial}{\partial E} (m_t p^*) = 0, \end{cases} \quad (13)$$

with $p^* = \arg \sup_p \{ \hat{u}(p, m_t) - p \frac{\partial \hat{v}_t}{\partial E} \}$. The first equation gives $\hat{v}_t$ and $p^*(T \rightarrow T')$ given m_t . The second equation gives m_t given $p^*(T \rightarrow T')$. Equilibrium power control policy resulting from this case is illustrated in the following section.

IV. NUMERICAL RESULTS FOR A QUASI-STATIC CHANNELS

In this section, we provide illustrative numerical results for the particular case of quasi-static channels, i.e., when the energy dynamics are faster than the channel dynamics. For the implementation of the proposed scheme, each terminal requires only the knowledge of its own channel state and energy level to choose the transmit power.

A. Comparison with other NE

For the purpose of evaluating our results, we compare the equilibrium of the MFG to the equilibrium of two other well known games.

- 1) The static Nash: This is the classical NE from the work of [6]. The equilibrium point is given by the equation:

$$\forall k \in \mathcal{K}, p_k^* = \frac{\sigma^2}{|h_k|^2} \frac{\beta^*}{1 - \theta \beta^*} \quad (14)$$

where β^* is the unique solution of the equation

$$x f'(x) - f(x) = 0. \quad (15)$$

- 2) The repeated game "operating point": When the power control game is treated as a discounted repeated game, there are several NE. In [12], the authors propose an equilibrium point, defined as the "operating point" which can be very close to the global optimal point. This equilibrium is given by the equation:

$$\forall k \in \mathcal{K}, \tilde{p}_k = \frac{\sigma^2}{|h_k|^2} \frac{\tilde{\beta}}{1 - \theta \tilde{\beta}} \quad (16)$$

where $\tilde{\beta}$ is the unique solution of the equation

$$x(1 - \theta x) f'(x) - f(x) = 0. \quad (17)$$

B. Parameters used

For the purpose of simulations, in order to obtain useful numerical results we take the following parameters:

- 1) the rate $R = 1$ Mbps,
- 2) the noise level with path loss, $\sigma^2 = 0.1$ W,
- 3) the channel gain $\mathbb{E}[h(t)]^2 = 1$,
- 4) the success function $f(\gamma) = \exp(-\frac{0.9}{\gamma})$ [2].

The averaged channel gain for all players at all times is taken to be a constant $\mathbb{E}[h(t)]^2 = 1$. The initial distribution in energy of the terminals m_T is specified for each figure. The final condition $\hat{v}_{T'} = q(\underline{s}(T'))$ is set to 0, in order for the

Fig. No	$T' - T$	$m_t(E, t = 0)$	Plotting
1	20	1	P^* v.s (t, E)
2	20	1	P^* v.s t
3	20	1	m_t v.s (t, E)
4	20	0 if $E \leq 18$, else 1	m_t v.s (t, E)
5	120	1	P^* v.s t
6	120	1	$v_t(E, t = 0)$ v.s E

TABLE I
SIMULATION PARAMETERS FOR THE PRESENTED FIGURES.

comparison to be fair with equilibrium power policies from the static game and the repeated discounted game. Indeed, there is no final reward in these two games. We take the maximum available energy to be $E_{\max} = 20$ J and time duration $T' - T$ is specified for each figure.

C. Discussion

On Fig. 1, the equilibrium power policy plot shows that terminals starting the game with high energy level (20 J in the figure) start transmitting with a high power level. Then this power decreases with time. It can also be noted that terminals starting the game with low energy do not transmit at high power values at the beginning of the game. They first wait for other terminals to empty their battery in order to suffer from less interference. This phenomenon is highlighted on Fig. 2.

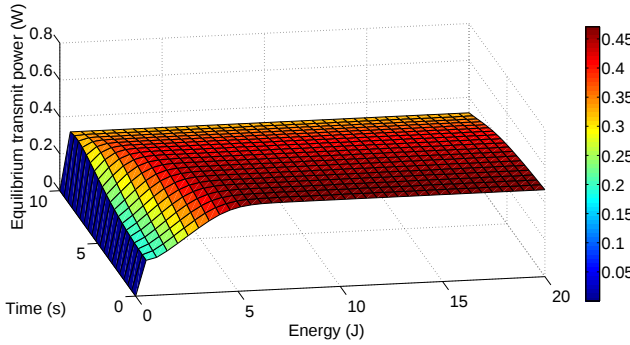


Fig. 1. MFG equilibrium power policy w.r.t. time and energy.

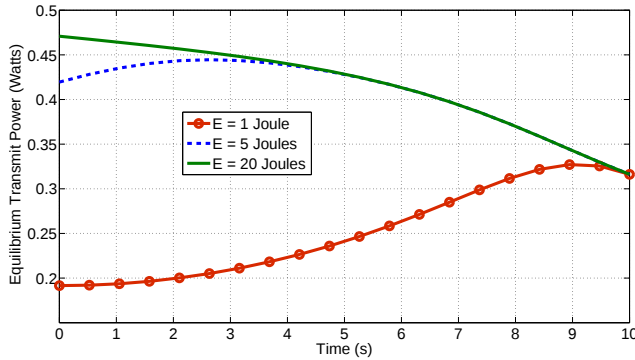


Fig. 2. MFG equilibrium power policies w.r.t. time for 3 initial amounts of energy, for a short game duration.

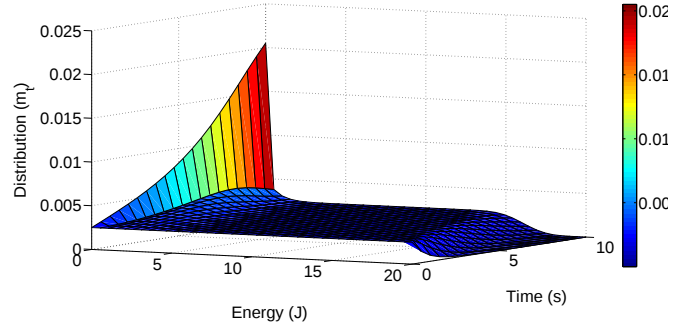


Fig. 3. Evolution of the energy distribution with the equilibrium power policy of the MFG for uniform initial distribution.

On Fig. 3, where terminals start the game with a uniform energy distribution, it can be seen that the amount of terminals with high energy (20 J) decreases with time, whereas an increasing proportion of terminals have empty batteries. A similar behavior is captured on Fig. 4, except that there are only terminals with high energy at the beginning of this case.

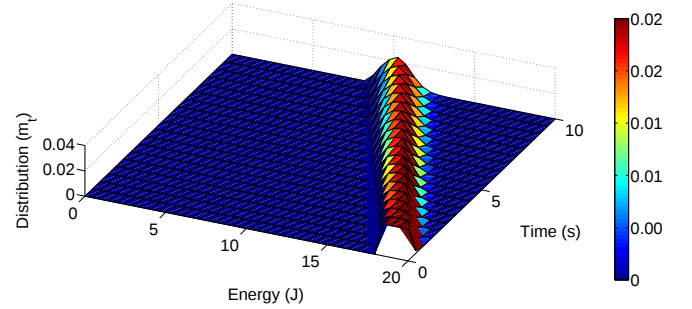


Fig. 4. Evolution of the energy distribution with the equilibrium power policy of the MFG for non-uniform initial distribution.

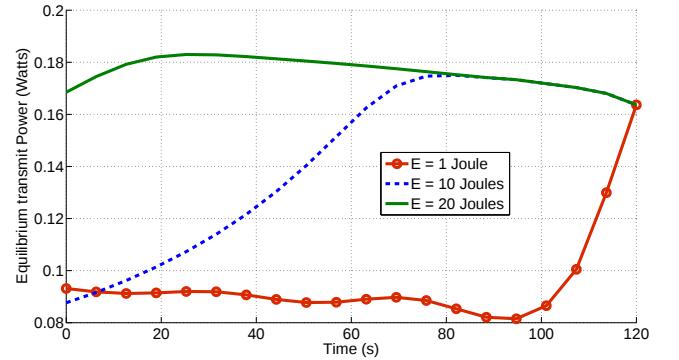


Fig. 5. MFG equilibrium power policies w.r.t. time for 3 initial amounts of energy, for a long game duration.

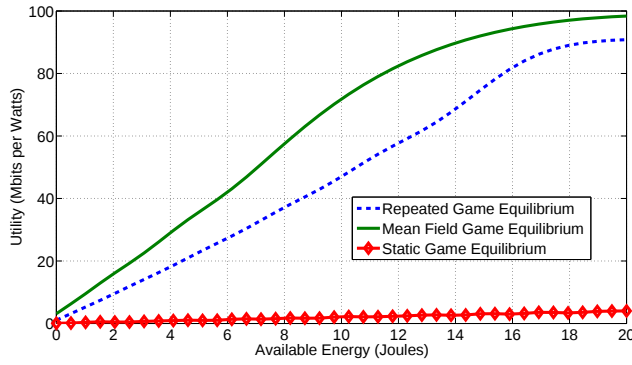


Fig. 6. Comparison of the average energy-efficient utilities with the MFG equilibrium power policy, for a long game duration, against the equilibrium from the static game, and the operating point from the repeated game formulation defined earlier.

Since the solution to the mean field response problem is determined for a given time duration, the equilibrium power control policy naturally depends on this duration. Hence, it is interesting to increase the time duration, as it is done on Fig. 5, to capture the impact on the power policy. Globally, it entices the terminals to consume less power to make their batteries last longer. This low consumption gives the terminals a better energy-efficiency than with the equilibrium from the static game or the equilibrium from the discounted repeated game, as illustrated on Fig. 6.

V. CONCLUSION

In this work, we treat the problem of power control games as a mean field game, taking into account the limited energy available to mobile transmitters, and the effects of channel fading. Provided that convergence conditions are met, the mean field approach offers interesting and useful results. While in the general case, the resulting partial differential equations are hard to solve, we analyze some special cases where numerical results can be obtained. The numerical results are encouraging, but they have only been computed for the case of quasi-static channels.

Some of our key results that can be interesting to engineers and other researchers are presented below:

- 1) Terminals with a low starting energy transmit with lower power at the start and slowly increase their power with time, when the energy distribution is homogeneous and $T' - T$ small.
- 2) In the same case, terminals with a high starting energy start with a high power and slowly decrease their power with time.
- 3) Terminals with a low starting energy transmit with lower power at the start and slowly decrease their power with time and then raise, when the energy distribution is homogeneous and $T' - T$ large.
- 4) The equilibrium of the mean field game when $T' - T$ is large, outperforms other known equilibrium (from [6] and [12]) in terms of energy-efficiency.

A numerical solution to the general case of an evolving channel as well as the energy constraint, is a possibility for future work. Additional possibilities include considering a multi-cellular network, distributed base stations and the multi-carrier case.

REFERENCES

- [1] G. Bacci. Energy-efficient power control for cdma code acquisition over frequency-selective channels. *IEEE Communications Letters*, 16(3):364–367, 2012.
- [2] E. V. Belmega and S. Lasaulce. Energy-efficient precoding for multiple-antenna terminals. *IEEE Trans. on Signal Processing*, 59(1):329–340, January 2011.
- [3] N. Bonneau, M. Debbah, E. Altman, and A. Hjørungnes. Non-atomic games for multi-user system. *IEEE Journal on Selected Areas in Communications*, 26(7):1047–1058, September 2008.
- [4] S. Buzzi and D. Saturnino. A game-theoretic approach to energy-efficient power control and receiver design in cognitive CDMA wireless networks. *Journal of Selected Topics in Signal Processing*, 5(1):137–150, 2011.
- [5] B. A. Fette. *Cognitive Radio Technology*. Elsevier, 2006.
- [6] D. J. Goodman and N. B. Mandayam. Power control for wireless data. *IEEE Personal Communications*, 7:48–54, 2000.
- [7] P. Gupta and P. R. Kumar. A system and traffic dependent adaptive routing algorithm for ad hoc networks. In *IEEE Conf. on Decision and Control (CDC)*, pages 2375–2380, San Diego, USA, December 1997.
- [8] S. Lasaulce, M. Debbah, and E. Altman. Methodologies for analyzing equilibria in wireless games. *IEEE Signal Processing Magazine, Special issue on Game Theory for Signal Processing*, Sep. 2009.
- [9] S. Lasaulce, Y. Hayel, R. E. Azouzi, and M. Debbah. Introducing hierarchy in energy games. *IEEE Trans. on Wireless Comm.*, 8(7):3833–3843, 2009.
- [10] S. Lasaulce and H. Tembine. *Game Theory and Learning for Wireless Networks : Fundamentals and Applications*. Academic Press, August 2011.
- [11] J. M. Lasry and P. L. Lions. Mean field games. *Jpn. J. Math.*, 2(1):229–260, 2007.
- [12] M. Le Treust and S. Lasaulce. A repeated game formulation of energy-efficient decentralized power control. *IEEE Trans. on Wireless Communications*, 2010.
- [13] F. Mériaux, S. Lasaulce, and H. Tembine. Stochastic differential games and energy-efficient power control. *Dynamic Games and Applications*, 2012.
- [14] F. Meshkati, M. Chiang, H. V. Poor, and S. C. Schwartz. A game-theoretic approach to energy-efficient power control in multi-carrier CDMA systems. *IEEE Journal on Selected Areas in Communications*, 24(6):1115–1129, 2006.
- [15] J. Mitola and G. Q. Maguire. Cognitive radio: making software radios more personal. *IEEE Personal Communications*, 6(4):13–18, Aug. 1999.

Publications relatives à l'apprentissage et aux dynamiques de meilleures réponses

C.1 ICST-VALUETOOLS-2011

F. Mériaux, S. Lasaulce et M. Kieffer. *More about Base Station Location Games*. In ICST Conference on Performance Evaluation Methodologies and Tools (VALUE-TOOLS), pages 372-380, 2011.

More about Base Station Location Games

François Mériaux, Samson Lasaulce, Michel Kieffer
L2S - CNRS - SUPELEC - Univ Paris-Sud
3 rue Joliot-Curie F-91192 Gif-sur-Yvette, France
{meriaux, lasaulce, kieffer}@lss.supelec.fr

May 12, 2011

Abstract

This paper addresses the problem of locating base stations in a certain area which is highly populated by mobile stations; each mobile station is assumed to select the closest base station. Base stations are modeled by players who choose their best location for maximizing their uplink throughput. The approach of this paper is to make some simplifying assumptions in order to get interpretable analytical results and insights to the problem under study. Specifically, a relatively complete Nash equilibrium (NE) analysis is conducted (existence, uniqueness, determination, and efficiency). Then, assuming that the base station location can be adjusted dynamically, the best-response dynamics and reinforcement learning algorithm are applied, discussed, and illustrated through numerical results.

1 Introduction

Mobile terminals (MTs) are currently gaining increased autonomy of decision to allow a better use of the available wireless resources. For example, MTs may choose their wireless access technology or the base station (BS) or access point to which they want to connect. We could imagine that this may be done in the future independently of the network operator owner of the BS. A mobile operator deploying BSs for a wireless network will have to deal with these new characteristics. If his goal is to maximize the traffic gathered by his own BSs, he will have to take into account the presence of competitor network operators when deciding on the location of BSs. If every operator involved has the same reasoning, this problem of BSs placement may be cast in the framework of game theory and more precisely in the context of location games.

The history of location games starts with the work of Hotelling [8] in which the notion of spatial competition in a duopoly situation is introduced. More precisely, two firms compete for benefits over a finite segment crowded with customers. This results in the partition of the segment into a convex area of influence for each firm. Plastria [12] gives an overview of optimization approaches to place new facilities in an environment with pre-existing facilities. A large overview on location games is also presented in [7]. Location games are extended to the context of wireless networks with works such as [1] and [2]. The main difference arising

in this new context is the interaction between MTs due to the mutual interference they generate. This point makes the association problem between MTs and BSs complex. As an association between a MT and a BS depends on SINR, the association relies on the respective locations of the MT and the BS, but also on the MTs already connected to the BS. Whereas [2] focuses on the downlink case, in [1] the location of BSs and the association choice of the MTs is treated as a Stackelberg game [15] in the uplink case. The context of our work is similar to the one in [1] but several interesting results are obtained in the present paper. The main *contributions* of this paper can be summarized as follows:

- As in [1] MTs are assumed to operate in the uplink and to be distributed along a one-dimensional region. However, each MT is assumed to select the closest BS (e.g., based on measure given by a GPS -global positioning system- receiver). This leads to a convenient form for the BSs utility functions (Sec. 2.1). As a consequence, the existence of a pure Nash equilibrium can be made rigorously (Sec. 3.1).
- Due to the symmetry of the problem, multiple Nash equilibria generally exist. However, if the locations can be ordered (which is easy for one-dimension regions), the Nash equilibrium can be determined and checked to be unique (Sec. 3.3).
- By making the reasonable assumption that the BS heights are much less than the typical distance between the BSs, the game can be further simplified and shown to be a form of Cournot oligopoly [4].
- In the two-player case, the efficiency of the Nash equilibrium is studied by evaluating the price of anarchy [9]. The influence of deploying its BS in the first place is studied by considering a Stackelberg formulation of the problem (Sec. 4.1).
- Assuming that the BS locations can be adjusted dynamically (which would be relevant in scenario like the one of small cells where only some of the small BS have to be active), the best response dynamics and reinforcement learning algorithm [3][14][11] are performed (Sec. 5).
- The made assumptions lead to several interpretations which could be further analyzed in the light of a more general framework (e.g., in two-dimensional regions).

The remainder of the paper is organized as follows. Section 2 introduces the physical model and the parameters of the K -player game. Section 3 describes the Nash equilibrium of the game in the one-dimensional case. In Section 4, the Nash equilibrium, the Stackelberg equilibrium, and the social optimum are compared. Section 5 presents a way to reach equilibrium using best-response dynamics and reinforcement learning. Finally, Section 6 concludes this work.

2 Model

Consider a plane to which a frame $\mathbb{R}$ is attached. A MT X located in a position $x \in \mathbb{R}^2$ in this plane is linked with a BS X_1 of height ε situated in $x_1 \in \mathbb{R}^2$, see Figure 1. We define the Signal to Noise Ratio (SNR) and the Signal to Interference plus Noise Ratio (SINR) of this MT

$$SNR_X = \frac{P_X \cdot h_{X_1}(x)}{\sigma_{X_1}^2}, \quad (1)$$

$$SINR_X = \frac{P_X \cdot h_{X_1}(x)}{\sigma_{X_1}^2 + I_{X_1}(x)}, \quad (2)$$

where P_X is the transmission power of the MT X , i.e. the level of power chosen by the MT to transmit its signal. $\sigma_{X_1}^2$ is the power of the channel noise, $I_{X_1}(x)$ is some interference term, $h_{X_1}(x)$ is the attenuation introduced by the uplink channel from X to X_1 . Here, it is assumed that

$$h_{X_1}(x) = (|x - x_1|^2 + \varepsilon^2)^{-\frac{\alpha}{2}}, \quad (3)$$

where $|x|$ is the ℓ_2 -norm of x and $\alpha \geq 2$ is the path-loss exponent, $\alpha = 2$ corresponding to the free-space path-loss case. A higher value of α suits to worse channel conditions.

With Single User Decoding (SUD) at the BS X_1 , there is no hierarchy for decoding the incoming signals at the BS. Hence, the signal from MT X is decoded by taking into account the full interference and the uplink capacity between X and X_1 may be written as

$$C_X = \log(1 + SINR_X), \quad (4)$$

Without loss of generality, when several MTs are considered, it is assumed that P_X does not depend on the MT and is normalized, i.e., $P_X = 1$. Moreover, the channel conditions, described by (3), are the same for every MT and the noise power is constant $\sigma_{X_1} = \sigma$. With these assumptions, (4) becomes

$$C_X = \log \left(1 + \frac{(|x - x_1|^2 + \varepsilon^2)^{-\frac{\alpha}{2}}}{\sigma^2 + I_{X_1}(x)} \right). \quad (5)$$

When several BSs are located on the plane, the MTs are assumed to be able to choose the BS they want to be linked with. In this paper, we consider that this association is made based on SNR. Given (3), choosing the BS with the highest SNR is equivalent to choosing the closest BS. Thus, it is assumed that a MT always chooses the closest BS to its location, leading to convex cells for BSs. For example,

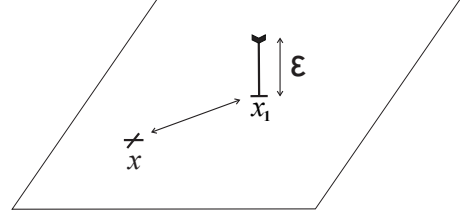


Figure 1: Base station located at x_1 and mobile station located at x

considering two BSs X_1 and X_2 at positions x_1 and x_2 and a MT X at x , one has

$$C_X = \begin{cases} \log \left(1 + \frac{(|x - x_1|^2 + \varepsilon^2)^{-\frac{\alpha}{2}}}{\sigma^2 + I_1(x)} \right) & \text{if } |x - x_1|^2 \leq |x - x_2|^2 \\ \log \left(1 + \frac{(|x - x_2|^2 + \varepsilon^2)^{-\frac{\alpha}{2}}}{\sigma^2 + I_2(x)} \right) & \text{if } |x - x_2|^2 \leq |x - x_1|^2 \end{cases}$$

2.1 Base station utility

The utility of a BS is taken as the sum of uplink capacities it offers to its connected users. We assume that the number of MT is large enough to be represented by a continuous distribution $\rho(x)$. This assumption allows to get the utility for the k -th BS as a continuous sum of the MTs uplink capacities

$$U_k(\underline{x}) = \int_{S_k(\underline{x})} \rho(z) \log \left(1 + \frac{(|z - x_k|^2 + \varepsilon^2)^{-\frac{\alpha}{2}}}{\sigma^2 + I_k(\underline{x})} \right) dz, \quad (6)$$

where $S_k(\underline{x})$ is the subset of the plane where MTs are linked with the k -th BS and $\underline{x} = {}^t(x_1, x_2, \dots, x_K)$ is the vector of locations for the set of BSs. This paper will consider only uniform MT distribution with $\rho(x) = 1$.

When considering interferences, a worst-case scenario is considered: there is no mechanism such as beamforming [10] to lower their effects. We consider that there is no interference between MTs connected with different BSs because of frequency reuse. Then only interference between MTs of a same BS has to be considered. This framework is quite similar to the one of [1], where two competing BSs are assumed to use different frequency bands. In our case, we consider K BSs (with $K \geq 1$) and each of them uses its own frequency band.

Performing SUD at the k -th BS, one gets

$$I_k(\underline{x}) = \int_{S_k(\underline{x})} (|z - x_k|^2 + \varepsilon^2)^{-\frac{\alpha}{2}} dz. \quad (7)$$

Utility of k -th BS (6) then becomes

$$U_k(\underline{x}) = \int_{S_k(\underline{x})} \log \left(1 + \frac{(|z - x_k|^2 + \varepsilon^2)^{-\frac{\alpha}{2}}}{\sigma^2 + \int_{S_k(\underline{x})} (|z' - x_k|^2 + \varepsilon^2)^{-\frac{\alpha}{2}} dz'} \right) dz. \quad (8)$$

2.2 Utility approximation

When considering the low SINR regime, the useful power of each MT is small compared to the interference term in

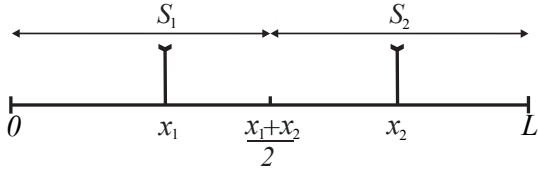


Figure 2: Two base stations competing in a one-dimension space

(8). This is especially true for high MT density. With this assumption, (8) may be approximated as

$$U_k(\underline{x}) \approx \frac{\int_{\mathcal{S}_k(\underline{x})} (|z - x_k|^2 + \varepsilon^2)^{-\frac{\alpha}{2}} dz}{\sigma^2 + \int_{\mathcal{S}_k(\underline{x})} (|z' - x_k|^2 + \varepsilon^2)^{-\frac{\alpha}{2}} dz'} \quad (9)$$

$$= \frac{I_k(\underline{x})}{\sigma^2 + I_k(\underline{x})}.$$

Note that this simplification makes the considered utility based on capacity equivalent to a utility based on SINR such as the ones in [1]. At low-SINR regime, it is equivalent to work with a capacity-based utility or a SINR-based utility.

Also note that one has $f(t) = t/(\sigma^2 + t)$ strictly increasing over $[0, L]$, since its derivative is $f'(t) = \sigma^2/(\sigma^2 + t)^2$. Then maximizing the approximation of $U_k(\underline{x})$ or $I_k(\underline{x})$ is equivalent. Thus the utility we define for the game is $\widehat{U}_k(\underline{x}) = I_k(\underline{x})$, $k \in \mathcal{K}$.

2.3 Definition of the game

In this section, we study the case of K BSs competing on a segment of length L . Each of the BSs uses different carrier frequencies so the MTs of different BSs do not interfere together. As assumed in Section 2, the MT distribution is uniform over the segment, and the set of possible locations for the BS is $[0, L]$. Figure 2 illustrates the context for the two-player case.

Definition 1. *The strategic form of the game is given by*

$$\mathcal{G} = (\mathcal{K}, \{\mathcal{A}_k\}_{k \in \mathcal{K}}, \{\widehat{U}_k\}_{k \in \mathcal{K}})$$

where

- $\mathcal{K} = \{1, \dots, K\}$ is the set of players, which are here BS.
- $\{\mathcal{A}_k\}_{k \in \mathcal{K}}$ is the set of actions players can consider, here

$$\mathcal{A}_k = \{x_k \in [0, L] \mid 0 < x_1 < \dots < x_K < L\}. \quad (10)$$

Denote $\mathcal{A} = \{\underline{x} \in [0, L]^K \mid 0 < x_1 < \dots < x_K < L\}$.

- $\{\widehat{U}_k\}_{k \in \mathcal{K}}$ is the set of utilities players use.

Note that we are interested in location equilibria that do not superimpose several BSs. Thus, if there exists an equilibrium, there exists a spatial order for BSs at this equilibrium, that is why we introduce this order in the action spaces $\{\mathcal{A}_k\}_{k \in \mathcal{K}}$.

3 Nash equilibrium analysis

The aim of this section is to show the existence of a Nash equilibrium in the location game described in Section 2.3 and to characterize this equilibrium.

3.1 Existence

In this section, we focus on the existence of a Nash equilibrium in the defined game. To prove the existence of a Nash equilibrium, the concavity of $\widehat{U}_k(\underline{x})$ with respect to x_k over $\mathcal{A}_k$, $\forall k \in \mathcal{K}$, has to be established.

Lemma 1. *$\widehat{U}_k(\underline{x})$ is concave with respect to x_k over $\mathcal{A}_k$, $\forall k \in \mathcal{K}$.*

The proof of this lemma is in Appendix A.1. Then one has the following theorem.

Theorem 1. *In the game defined by Definition 1, there exists at least one Nash equilibrium.*

Proof. • Using Lemma 1, we know that $\widehat{U}_k(\underline{x})$ is concave with respect to x_k over $\mathcal{A}_k$, $\forall k \in \mathcal{K}$,

- $\widehat{U}_k(\underline{x})$ is continuous with respect to $\underline{x}$ over $\mathcal{A}$, $\forall k \in \mathcal{K}$,
- the set of feasible actions is compact and convex for all players in the game.

The Rosen [13] conditions for the existence of a Nash equilibrium are met and Theorem 1 is thus proved. $\square$

3.2 Multiplicity of NE

Regarding to the uniqueness of the Nash equilibrium, as the characteristics of the BSs (height and noise) are assumed to be identical, it is interesting to note that permuting the order of BSs leads to a symmetric system of equation. Thus, without condition on the order of BSs as in (10), there are $K!$ Nash equilibria for the game and all these equilibria are symmetric, meaning that the set of locations at equilibrium is unique. However, if one imposes the condition order (10), the NE can be shown to be unique by using the Diagonally Strict Concavity (DSC) condition [13].

Theorem 2. *In the game defined by Definition 1, there exists one single Nash equilibrium.*

The proof of this theorem is in Appendix A.2. Having uniqueness under this order condition might seem to be a weak result in comparison with a general uniqueness result. However, in the framework of a dynamic process, the initial locations of the base station might suffice to determine the effectively observed NE (after convergence).

3.3 Determination of the NE

A characterization of the equilibrium is provided in this section with examples for small values of K . A real solution $\underline{x}$

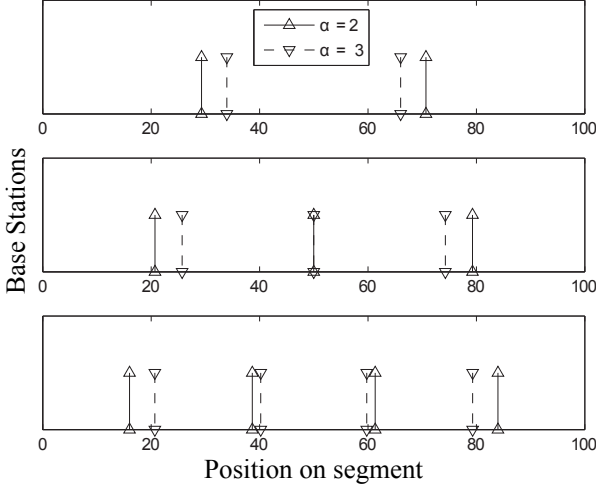


Figure 3: Base Stations at Nash equilibrium for $\alpha = 3$ (dotted line) and $\alpha = 2$ (continuous line). From the top to the bottom, configuration with 2 BSs, 3 BSs and 4 BSs. Parameters are $\varepsilon = 0.1$ and $L = 100$.

for $\alpha > 2$ has to satisfy

$$\begin{cases} x_1 = \frac{2\sqrt{(5.2\frac{2}{\alpha} - 2\frac{4}{\alpha} - 4)\varepsilon^2 + 2\frac{2}{\alpha}x_2^2 - 2\frac{2}{\alpha}x_2}}{4 - 2\frac{2}{\alpha}}, \\ x_k = \frac{x_{k+1} + x_{k-1}}{2}, \forall k \in \{2, \dots, K-1\}, \\ x_K = \frac{-2\sqrt{(5.2\frac{2}{\alpha} - 2\frac{4}{\alpha} - 4)\varepsilon^2 + 2\frac{2}{\alpha}(L - x_{K-1})^2 - 2\frac{2}{\alpha}x_{K-1} + 4L}}{4 - 2\frac{2}{\alpha}}. \end{cases} \quad (11)$$

See Appendix A.3 for detailed derivations. In the case $\alpha = 2$, one obtains

$$\begin{cases} x_1 = \sqrt{2\varepsilon^2 + 2x_2^2} - x_2, \\ x_k = \frac{x_{k+1} + x_{k-1}}{2}, \forall k \in \{2, \dots, K-1\}, \\ x_K = -\sqrt{10\varepsilon^2 + 2(L - x_{K-1})^2} - x_{K-1} + 2L. \end{cases} \quad (12)$$

In the 2-player game with $\alpha = 2$, one gets

$$\begin{cases} x_1^{\text{ne}} = L - \frac{1}{2}\sqrt{2L^2 - 4\varepsilon^2} \\ x_2^{\text{ne}} = \frac{1}{2}\sqrt{2L^2 - 4\varepsilon^2} \end{cases} \quad (13)$$

Figure 3 illustrates equilibria for $\alpha = 2$ and $\alpha = 3$ for respectively two, three, and four BSs.

4 Comparing equilibria

This section compares the Nash equilibrium described in Section 3, the Stackelberg equilibrium, and the social optimum.

4.1 Stackelberg equilibrium

As written in Section 3, the pure Nash equilibrium needs the players to know their spatial order to be reached. If players do not know this order but play in a chronological order, the problem changes. If the first player plays alone at the first stage of the game knowing that other players will

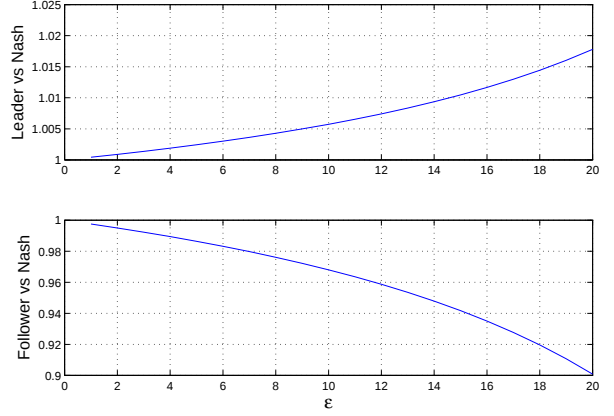


Figure 4: $\frac{\widehat{U}_1^{\text{SE}}}{\widehat{U}_1^{\text{NE}}}$ and $\frac{\widehat{U}_2^{\text{SE}}}{\widehat{U}_2^{\text{NE}}}$ with regard to ε .

place their BSs after, it turns into a Stackelberg game [15] with the first player being the leader of the game. This idea is illustrated with a two-player game with one leader (BS 1) and one follower (BS 2). The leader chooses its position knowing that the follower will place itself after. Both BSs still want to maximize their utilities and BS 1 knows this point. Then, BS 1 knows how BS 2 is going to be placed regarding to its own position, it is simply the best response of BS 2

$$x_2(x_1) = \frac{-2\sqrt{(5.2\frac{2}{\alpha} - 2\frac{4}{\alpha} - 4)\varepsilon^2 + 2\frac{2}{\alpha}(L - x_1)^2 - 2\frac{2}{\alpha}x_1 + 4L}}{4 - 2\frac{2}{\alpha}}.$$

Then BS 1 places itself at a location x_1 solution of

$$\frac{\partial \widehat{U}_1}{\partial x_1}(x_1, x_2(x_1)) = 0. \quad (14)$$

With $\alpha = 2$ and neglecting ε compared to every other lengths of the problem, one gets

$$x_1^{\text{se}} = \left(1 - \sqrt{2} + \sqrt{2 - \sqrt{2}}\right)L, \quad (15)$$

and

$$x_2^{\text{se}} = \left((\sqrt{2} - 1)(1 + \sqrt{2 - \sqrt{2}})\right)L. \quad (16)$$

On Figure 4, we compare the utilities of the leader and the follower of the Stackelberg game versus the utilities of the Nash equilibrium. As we see, the leader of the Stackelberg game has a better utility than what he would have get at the Nash equilibrium. On the contrary, the follower has a worst utility. Hence, in a mobile operator point of view, it is more interesting to deploy its BS first.

4.2 Social optimum

Sections 3 and 4.1 provide equilibria corresponding to situations where the BSs only consider what is best for themselves. But these equilibria are not the best for the MTs.

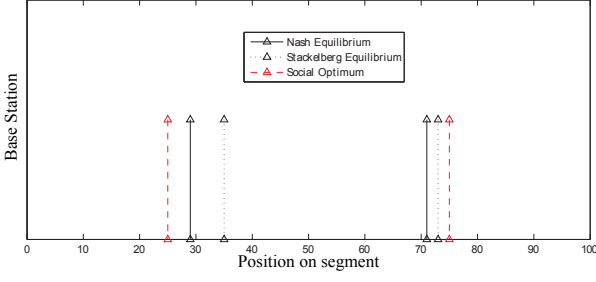


Figure 5: Nash Equilibrium, Stackelberg Equilibrium, and Social Optimum for $\alpha = 2$.

In the MTs point-of-view, the utility to consider is the complete utility sum or *social utility*. In a 2-BS case

$$\widehat{U}^{\text{so}}(x_1, x_2) = \widehat{U}_1(x_1, x_2) + \widehat{U}_2(x_1, x_2). \quad (17)$$

For this utility, we know that there exists an optimum since the strategy set $\mathcal{A}$ is compact and $\widehat{U}_1(x_1, x_2) + \widehat{U}_2(x_1, x_2)$ is continuous with respect to (x_1, x_2) over $\mathcal{A}$. At this optimum, it is proven [1] that

- (i) the BSs place themselves at the middle of their associated subset of segment (respectively $\mathcal{S}_1$ and $\mathcal{S}_2$),
- (ii) the frontier between the two BSs is the middle of the segment.

Thus, we have the optimum

$$x_1^{\text{so}} = \frac{L}{4}, x_2^{\text{so}} = \frac{3L}{4}. \quad (18)$$

Figure 5 shows the Nash equilibrium, the Stackelberg equilibrium, and the social optimum for $\alpha = 2$. The locations of BSs for Nash equilibrium and social optimum are symmetric with respect to the middle of the segment $[0, L]$ whereas this is not the case for the Stackelberg equilibrium.

4.3 Price of anarchy

To compare the Nash and the Stackelberg equilibrium in the two-player case, we look at the utilities at equilibrium. Without any pricing mechanism, the Nash equilibrium and the social optimum are very close in terms of locations. As a result, they are also close in terms of utility sum. The *Price of Anarchy* (PoA), introduced in [9], is an adequate metric to compare these sums

$$\text{PoA}(\text{eq}) = \frac{\max_{\underline{x} \in \mathcal{A}} \sum_{k \in \mathcal{K}} \widehat{U}_k(\underline{x})}{\sum_{k \in \mathcal{K}} \widehat{U}_k(\underline{x}^{\text{eq}})}. \quad (19)$$

Note that the PoA is always stronger than 1 and if the PoA is high, it means that the corresponding equilibrium is not that efficient in terms of utility sum. On the contrary, if the PoA is close to one, the corresponding equilibrium is satisfying. Figure 6 illustrates the behavior of the PoA for the Nash equilibrium and the PoA for the Stackelberg equilibrium as a function of ε for 2-player case. It appears that the Stackelberg equilibrium is less efficient than the Nash equilibrium.

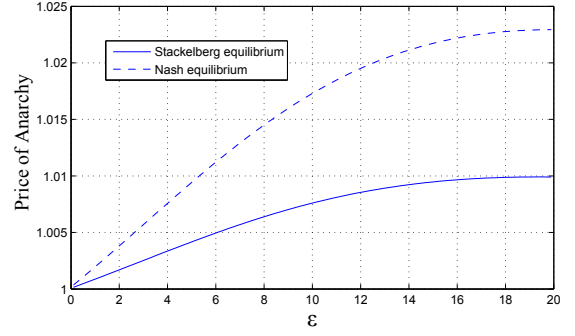


Figure 6: Prices of Anarchy regarding to ε . Parameters are $L = 100$, $\sigma^2 = 10^4$, $\alpha = 2$.

5 Convergence to the Nash equilibrium

Getting to the equilibrium given in Section 3.3 is not simple: it would require that every player knows the utilities of other players. In practice, this is hardly the case. Thus, we present techniques that enable to reach the Nash equilibrium in a decentralized way: best-response dynamics and reinforcement learning. However the BSs need to be movable to perform these two techniques, hence it is more accurate to talk about Mobile Stations (MSs) in the present section. Note that for these two techniques, the only assumption about MSs location is that they cannot superimpose.

5.1 Best response dynamics

The context of this section remains the same as defined in Section 2.3.

The principle of best-response dynamics is that given a realization of actions for its opponents, every player of the game is able to compare its own possible actions and choose which one is best for itself. Precisely the best-response algorithm is the following

1. At every time step t , each MS k chooses its location $x_k^{\text{br}}(t)$ according to

$$x_k^{\text{br}}(t) = \arg \max_{x_k} \widehat{U}_k(x_k, \underline{x}_{-k}(t)).$$

2. Algorithm stops when $|\underline{x}^{\text{br}}(t+1) - \underline{x}^{\text{br}}(t)| < \beta$, with β fixed.

Regarding to the convergence of this algorithm, we study the system of best-responses (11). If we neglect ε regarding to x_1 and $L - x_{K-1}$, one gets a linear system of equations

$$\begin{cases} x_1 = \frac{2 \frac{1+\alpha}{\alpha} - 2 \frac{2}{\alpha}}{4 - 2 \frac{2}{\alpha}} x_2, \\ x_k = \frac{x_{k+1} + x_{k-1}}{2}, \forall k \in \{2, \dots, K-1\}, \\ x_K = \frac{(4 - 2 \frac{1+\alpha}{\alpha})L}{4 - 2 \frac{2}{\alpha}} + \frac{2 \frac{1+\alpha}{\alpha} - 2 \frac{2}{\alpha}}{4 - 2 \frac{2}{\alpha}} x_{K-1}. \end{cases} \quad (20)$$

It is a linear Cournot oligopoly [4].

One distinguishes two ways this algorithm can run.

Simultaneous best response dynamics. At every step, every MS adapt their locations simultaneously. In this case, the evolution of the locations with the algorithm can be expressed by

$$\underline{x}^{\text{br}}(t+1) = \underline{a} + A_K \underline{x}^{\text{br}}(t) \quad (21)$$

with $\underline{a} = {}^t(0, \dots, 0, \frac{(4-2\frac{1+\alpha}{\alpha})L}{4-2\frac{2}{\alpha}})$ and

$$A_K = \begin{pmatrix} 0 & g(\alpha) & 0 & \dots & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \dots & 0 & g(\alpha) & 0 \end{pmatrix}, \quad (22)$$

with $g(\alpha) = \frac{2\frac{1+\alpha}{\alpha} - 2\frac{2}{\alpha}}{4-2\frac{2}{\alpha}}$.

As recalled in Lemma 2.8 of [16], for an irreducible positive square matrix $A = (a_{ij})_{1 \leq i \leq n}$, then either

$$\sum_{j=1}^n a_{ij} = \rho(A) \quad \forall i \in [1, n], \quad (23)$$

or

$$\inf_{1 \leq i \leq n} \sum_{j=1}^n a_{ij} < \rho(A) < \sup_{1 \leq i \leq n} \sum_{j=1}^n a_{ij}, \quad (24)$$

with $\rho(A)$ being the radius of A . In our case, A_K can be verified to be positive and irreducible, then one has

$$g(\alpha) < \rho(A_K) < 1. \quad (25)$$

Hence, the convergence of the algorithm is ensured.

Sequential best response dynamics. The other case is the sequential algorithm: at every step, only one MS adapts its location. Depending on the MS k adapting its location, the algorithm evolves according to

$$\underline{x}^{\text{br}}(t+1) = \underline{a} + A_K^k \underline{x}^{\text{br}}(t), \quad (26)$$

with A_K^k corresponding to the adaptation of the k -th MS. A_K^k may have several forms.

- If the k -th MS, $k \in \{2, \dots, K-1\}$ adapts its location, then

$$A_K^k = \begin{pmatrix} 1 & 0 & \dots & 0 \\ \ddots & \ddots & \ddots & \vdots \\ \vdots & 0 & 1 & 0 \\ \vdots & & \frac{1}{2} & 0 & \frac{1}{2} \\ \vdots & & & 0 & 1 & 0 \\ \vdots & & & & \ddots & \ddots & \ddots \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

In this case, A_K^k is irreducible and positive and

$$\rho(A_K^k) = 1. \quad (27)$$

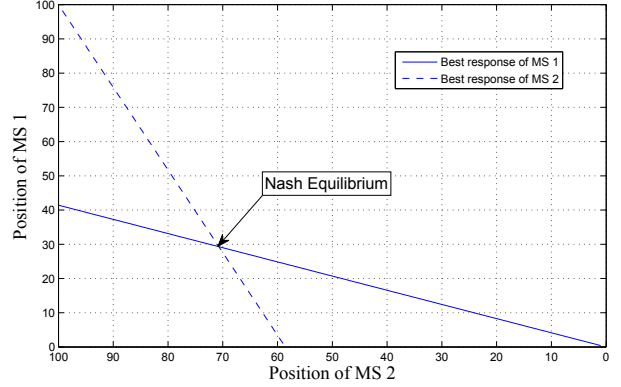


Figure 7: Best-responses of MS 1 and MS 2 for $\alpha = 2$

- If the first MS adapts its location, the matrix has the form

$$A_K^1 = \begin{pmatrix} 0 & g(\alpha) & 0 & \dots \\ 0 & 1 & 0 & \\ \vdots & \ddots & \ddots & \ddots \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

Again, A_K^1 is irreducible and positive, but this time

$$g(\alpha) < \rho(A_K^1) < 1. \quad (28)$$

Note that if the K -th MS adapts its location, the same reasoning can be done.

$$g(\alpha) < \rho(A_K^K) < 1. \quad (29)$$

Hence, $\rho(\prod_{k=1}^K A_K^k) = \prod_{k=1}^K \rho(A_K^k) < 1$ and the algorithm converges.

In Figure 7, we compare MSs 1 and 2 sequential best-responses for $\alpha = 2$.

5.2 Reinforcement Learning

The following notations are used in this section. For a set of MSs $\mathcal{K} = \{1, \dots, K\}$, let $\tilde{\mathcal{A}}_k = \{y_{k1}, \dots, y_{km_k}\}$ be the set of possible locations for MS k (m_k being the cardinality of $\tilde{\mathcal{A}}_k$), which corresponds to a discretization of the set $[0, L]$.

We implement a discrete stochastic learning algorithm in the sense that the action space is discrete and at every step, actions are chosen in a stochastic way. The only information available to a MS is the value of its utility after each iteration (note that the MS does not necessarily knows its utility expression).

We define $\underline{p}_k(t) = (p_{k1}(t), \dots, p_{km_k}(t))$ the probability distribution vector of MS k at time t .

$$P[x_k(t) = y_{ki}] = p_{ki}(t), \quad i \in \{1, \dots, m_k\}. \quad (30)$$

The algorithm used by each MS is then the following.

1. Initialize the distribution probability vector.
 $\forall k \in \mathcal{K}, \forall i \in \{1, \dots, m_k\}, p_{ki}(0) = \frac{1}{m_k}$.
2. At every step t , each MS k chooses a location $x_k(t)$ according to its probability vector $\underline{p}_k(t)$.

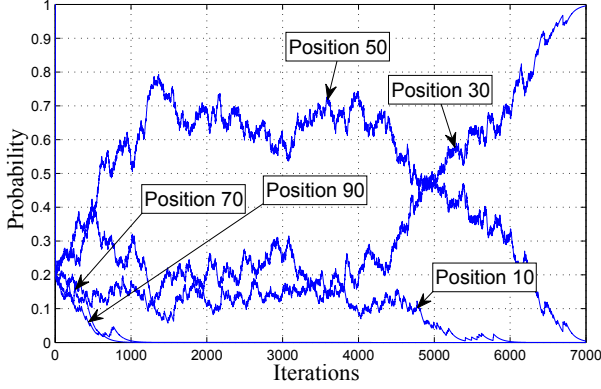


Figure 8: Evolution of action probabilities for BS 1 with discrete stochastic learning.

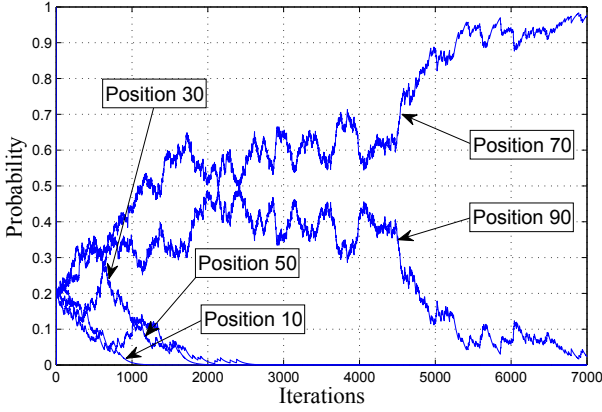


Figure 9: Evolution of action probabilities for BS 2 with discrete stochastic learning.

3. Each MS gets $U_k(t)$.
4. Each MS updates its probability distribution vector $\underline{p}_k(t)$

$$\begin{aligned} p_{ki}(t+1) &= p_{ki}(t) - bU_k(t)p_{ki}(t), \text{ if } x_t(t) \neq y_{ki} \\ p_{ki}(t+1) &= p_{ki}(t) + bU_k(t) \sum_{s \neq i} p_{ks}(t), \text{ if } x_k(t) = y_{ki} \end{aligned} \quad (31)$$

5. Algorithm stops when $\underline{p}_k(t+1) = \underline{p}_k(t)$, else go to step 2.

Figures 8 and 9 illustrate the evolution of the probability distribution vectors of two MSs. The parameters of the simulation are the following: each MS has the same set of possible positions $\{10, 30, 50, 70, 90\}$ and the step of the learning algorithm is $b = 0.01$.

Depending on the choice of b , two phenomena occur.

- The higher the value of b , the lower the convergence time of the algorithm,
- However, if b is chosen too high, the algorithm may converge to locations that do not correspond to an equilibrium of the game defined in Section 2.3.

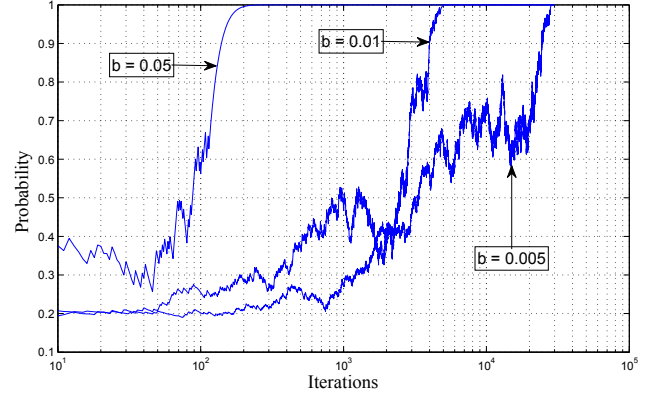


Figure 10: Convergence of probability with respect to step b

Figure 10 illustrates the convergence time as a function of b .

The choice of b is hence a trade-off between convergence time and accuracy of the convergence.

6 Conclusion and perspectives

Obviously the proposed model is relatively simple and should be improved to obtain refined results. However, this led us to several interesting results such as a full characterization of Nash equilibria and interesting behavior in terms of convergence. It would be very relevant to extend this work to two-dimensional scenario, define a suitable order for which uniqueness would be ensured. More connections with the famous multi-source Weber problem [6] should be established to better understand the general problem of *deployment games*. Indeed, if an operator has to locate a set of base stations, the problem becomes more complicated. The problem becomes even more interesting if the set of possible constellations is discrete, which would lead us to make connections with Voronoi games [5]. The authors believe there is large avenue for contributions to the general problem under study, especially in finding relevant assumptions to simplify it without too loss in terms of understanding. Random matrix theory and stochastic geometry might be for great help to achieve this challenging objective.

References

- [1] Eitan Altman, Anurag Kumar, Chandramani Kishore Singh, and Rajesh Sundaresan. Spatial sinr games combining base station placement and mobile association. In *INFOCOM*, pages 1629–1637. IEEE, 2009.
- [2] Eitan Altman, Alonso Silva, Amidou Tembine, and Merouane Debbah. Spatial games and global optimization for the mobile association problem: the downlink case. In *IEEE/SIAM CDC*, 2010.

- [3] Robert R. Bush and Frederick Mosteller. *Stochastic Models for Learning*. Wiley, 1st edition, 1955.
- [4] A. Cournot. *Researches into the principles of wealth*. Irwin Paperback Classics in Economics (original: Recherches sur les principes mathématiques de la théorie des richesses, 1838), 1963 (English Translation).
- [5] Christoph Dürr and Nguyen Kim Thang. Nash equilibria in Voronoi games on graphs. *CoRR*, abs/cs/0702054, 2007.
- [6] Sándor P. Fekete, Joseph S. B. Mitchell, and Karin Beurer. On the continuous fermat-weber problem. *CoRR*, cs.CG/0310027, 2003.
- [7] Jean J. Gabszewicz and Jacques-Francois Thisse. Location. In R.J. Aumann and S. Hart, editors, *Handbook of Game Theory with Economic Applications*, volume 1 of *Handbook of Game Theory with Economic Applications*, chapter 9, pages 281–304. Elsevier, 1992.
- [8] Harold Hotelling. Stability in competition. *The Economic Journal*, 39(153):41–57, 1929.
- [9] Elias Koutsoupias and Christos Papadimitriou. Worst-case equilibria. In *the 16th Annual Symposium on Theoretical Aspects of Computer Science*, pages 404–413, 1999.
- [10] John Litva and Titus K. Lo. *Digital Beamforming in Wireless Communications*. Artech House, Inc., Norwood, MA, USA, 1996.
- [11] Kumpati S. Narendra and Mandayam A. L. Thathachar. *Learning automata: an introduction*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1989.
- [12] Frank Plastria. Static competitive facility location: An overview of optimisation approaches. *European Journal of Operational Research*, 129(3):461–470, March 2001.
- [13] J. B. Rosen. Existence and Uniqueness of Equilibrium Points for Concave N-Person Games. *Econometrica*, 33(3):520–534, 1965.
- [14] P.S. Sastry, V.V. Phansalkar, and M.A.L. Thathachar. Decentralized learning of nash equilibria in multi-person stochastic games with incomplete information. *Systems, Man and Cybernetics, IEEE Transactions on*, 24(5):769–777, May 1994.
- [15] H. Stackelberg. *Marketform und Gleichgewicht*. Oxford, U.K., 1934.
- [16] Richard S. Varga. *Matrix iterative analysis*. Springer series in computational mathematics, 27. Springer, 2nd rev. and expanded ed. edition, 2000.

A Proofs

A.1 Proof of Lemma 1

We need to prove that $\widehat{U}_k(\underline{x})$ is concave with respect to x_k over $\mathcal{A}_k$.

We know that the index of players verify (10). In this case, the regions associated to the BSs are

$$\begin{cases} [0, \frac{x_1+x_2}{2}] \text{ for BS 1,} \\ [\frac{x_{k-1}+x_k}{2}, \frac{x_k+x_{k+1}}{2}] \text{ for BS } k, k \in \{2, \dots, K-1\} \\ [\frac{x_{K-1}+x_K}{2}, L] \text{ for BS } K. \end{cases} \quad (32)$$

Then

$$\begin{cases} \widehat{U}_1(\underline{x}) = \int_0^{\frac{x_1+x_2}{2}} (\epsilon^2 + (x - x_1)^2)^{-\frac{\alpha}{2}} dx \\ \widehat{U}_k(\underline{x}) = \int_{\frac{x_{k-1}+x_k}{2}}^{\frac{x_k+x_{k+1}}{2}} (\epsilon^2 + (x - x_k)^2)^{-\frac{\alpha}{2}} dx, k \in \{2, \dots, K-1\} \\ \widehat{U}_K(\underline{x}) = \int_{\frac{x_{K-1}+x_K}{2}}^L (\epsilon^2 + (x - x_K)^2)^{-\frac{\alpha}{2}} dx \end{cases} \quad (33)$$

which may be rewritten as

$$\begin{cases} \widehat{U}_1(\underline{x}) = \int_{-x_1}^{\frac{x_2-x_1}{2}} (\epsilon^2 + x^2)^{-\frac{\alpha}{2}} dx \\ \widehat{U}_k(\underline{x}) = \int_{\frac{x_{k-1}-x_k}{2}}^{\frac{x_{k+1}-x_k}{2}} (\epsilon^2 + x^2)^{-\frac{\alpha}{2}} dx, k \in \{2, \dots, K-1\} \\ \widehat{U}_K(\underline{x}) = \int_{\frac{x_{K-1}-x_K}{2}}^{\frac{L-x_K}{2}} (\epsilon^2 + x^2)^{-\frac{\alpha}{2}} dx \end{cases} \quad (34)$$

To prove the existence of a Nash equilibrium, the concavity of $\widehat{U}_k(\underline{x})$ with respect to $x_k \forall k \in \mathcal{K}$ has now to be established. One has the first-order partial derivatives

$$\begin{cases} \frac{\partial \widehat{U}_1}{\partial x_1}(\underline{x}) = -\frac{1}{2}(\epsilon^2 + (\frac{x_2-x_1}{2})^2)^{-\frac{\alpha}{2}} + (\epsilon^2 + x_1^2)^{-\frac{\alpha}{2}} \\ \frac{\partial \widehat{U}_k}{\partial x_k}(\underline{x}) = -\frac{1}{2}(\epsilon^2 + (\frac{x_{k+1}-x_k}{2})^2)^{-\frac{\alpha}{2}} + \frac{1}{2}(\epsilon^2 + (\frac{x_{k-1}-x_k}{2})^2)^{-\frac{\alpha}{2}} \\ \frac{\partial \widehat{U}_K}{\partial x_K}(\underline{x}) = -(\epsilon^2 + (L - x_K)^2)^{-\frac{\alpha}{2}} + \frac{1}{2}(\epsilon^2 + (\frac{x_{K-1}-x_K}{2})^2)^{-\frac{\alpha}{2}} \end{cases} \quad (35)$$

and the second-order partial derivatives

$$\begin{cases} \frac{\partial^2 \widehat{U}_1}{\partial x_1^2}(\underline{x}) = -\frac{\alpha}{8} \frac{x_2-x_1}{(\epsilon^2 + (\frac{x_2-x_1}{2})^2)^{\frac{\alpha}{2}+1}} - \alpha \frac{x_1}{(\epsilon^2 + x_1^2)^{\frac{\alpha}{2}+1}} \\ \frac{\partial^2 \widehat{U}_k}{\partial x_k^2}(\underline{x}) = \frac{\alpha}{8} \left(-\frac{x_{k+1}-x_k}{(\epsilon^2 + (\frac{x_{k+1}-x_k}{2})^2)^{\frac{\alpha}{2}+1}} + \frac{x_{k-1}-x_k}{(\epsilon^2 + (\frac{x_{k-1}-x_k}{2})^2)^{\frac{\alpha}{2}+1}} \right) \\ \frac{\partial^2 \widehat{U}_K}{\partial x_K^2}(\underline{x}) = -\alpha \frac{L-x_K}{(\epsilon^2 + (L-x_K)^2)^{\frac{\alpha}{2}+1}} + \frac{\alpha}{8} \frac{x_{K-1}-x_K}{(\epsilon^2 + (\frac{x_{K-1}-x_K}{2})^2)^{\frac{\alpha}{2}+1}} \end{cases} \quad (36)$$

Given (10), we have $\frac{\partial^2 \widehat{U}_k}{\partial x_k^2}(\underline{x}) < 0 \forall k \in \mathcal{K}$. Thus $\widehat{U}_k(\underline{x})$ is concave with respect to x_k over $\mathcal{A}_k$, $\forall k \in \mathcal{K}$.

A.2 Proof of Theorem 2

In the context of K -player game, the DSC [13] condition writes $\forall(\underline{a}, \underline{a}') \in \mathcal{A}^2$ such that $\underline{a} \neq \underline{a}'$

$$\sum_{k=1}^K (a'_k - a_k) \left(\frac{\partial \widehat{U}_k}{\partial x_k}(\underline{a}) - \frac{\partial \widehat{U}_k}{\partial x_k}(\underline{a}') \right) > 0 \quad (37)$$

For clarity reasons, we denote

$$g(a, b) = \left((\epsilon^2 + a^2)^{-\frac{\alpha}{2}} - (\epsilon^2 + b^2)^{-\frac{\alpha}{2}} \right), (a, b) \in \mathbb{R}^2. \quad (38)$$

By (35), it turns

$$\begin{cases} \frac{\partial \widehat{U}_1}{\partial x_1}(\underline{a}) - \frac{\partial \widehat{U}_1}{\partial x_1}(\underline{a}') = g(a_1, a'_1) - \frac{1}{2} \left(g\left(\frac{a_1 - a_2}{2}, \frac{a'_1 - a'_2}{2}\right) \right) \\ \frac{\partial \widehat{U}_k}{\partial x_k}(\underline{a}) - \frac{\partial \widehat{U}_k}{\partial x_k}(\underline{a}') = \frac{1}{2} \left(g\left(\frac{a_{k-1} - a_k}{2}, \frac{a'_{k-1} - a'_k}{2}\right) - g\left(\frac{a_k - a_{k+1}}{2}, \frac{a'_k - a'_{k+1}}{2}\right) \right) \\ \frac{\partial \widehat{U}_K}{\partial x_K}(\underline{a}) - \frac{\partial \widehat{U}_K}{\partial x_K}(\underline{a}') = \frac{1}{2} \left(g\left(\frac{a_{K-1} - a_K}{2}, \frac{a'_{K-1} - a'_K}{2}\right) \right) - g(L - a_K, L - a'_K) \end{cases} \quad (39)$$

Equation (37) becomes

$$\begin{aligned} & (a'_1 - a_1) \left(g(a_1, a'_1) - \frac{1}{2} g\left(\frac{a_1 - a_2}{2}, \frac{a'_1 - a'_2}{2}\right) \right) \\ & + \sum_{k=2}^{K-1} \frac{a'_k - a_k}{2} \left(g\left(\frac{a_{k-1} - a_k}{2}, \frac{a'_{k-1} - a'_k}{2}\right) - g\left(\frac{a_k - a_{k+1}}{2}, \frac{a'_k - a'_{k+1}}{2}\right) \right) \\ & + (a'_K - a_K) \left(\frac{1}{2} g\left(\frac{a_{K-1} - a_K}{2}, \frac{a'_{K-1} - a'_K}{2}\right) - g(L - a_K, L - a'_K) \right) > 0 \end{aligned} \quad (40)$$

which can also be written

$$\begin{aligned} & (a'_1 - a_1) g(a_1, a'_1) \\ & + \sum_{k=2}^K \frac{a'_k - a'_{k-1} - (a_k - a_{k-1})}{2} g\left(\frac{a_{k-1} - a_k}{2}, \frac{a'_{k-1} - a'_k}{2}\right) \\ & + (L - a'_K - (L - a_K)) g(L - a_K, L - a'_K) > 0 \end{aligned} \quad (41)$$

However $\forall (a, b) \in \mathbb{R}^{*2}$,

$$(b - a) g(a, b) > 0 \quad (42)$$

and by the order condition (10)

$$\begin{cases} (a'_1, a_1) \in \mathbb{R}^{*2} \\ \left(\frac{a'_k - a'_{k-1}}{2}, \frac{a_k - a_{k-1}}{2} \right) \in \mathbb{R}^{*2} \quad \forall k \in \{2, \dots, K\} \\ (L - a'_K, L - a_K) \in \mathbb{R}^{*2} \end{cases} \quad (43)$$

Then the DSC condition is verified and the equilibrium is unique.

A.3 Derivation of (11)

To obtain a formal expression of a Nash equilibrium, the intersection of the best-responses has to be considered

$$\frac{\partial \hat{U}_k}{\partial x_k}(\underline{x}) = 0, \quad \forall k \in \mathcal{K}, \quad \forall \alpha \geq 2. \quad (44)$$

(44) leads to

$$\begin{cases} (2^{\frac{2}{\alpha}} (\varepsilon^2 + (\frac{x_2 - x_1}{2})^2))^{\frac{\alpha}{2}} = (\varepsilon^2 + x_1^2)^{\frac{\alpha}{2}} \\ (\varepsilon^2 + (\frac{x_{k+1} - x_k}{2})^2)^{\frac{\alpha}{2}} = (\varepsilon^2 + (\frac{x_{k-1} - x_k}{2})^2)^{\frac{\alpha}{2}}, \quad \forall k \in \{2, \dots, K-1\} \\ (\varepsilon^2 + (L - x_K)^2)^{\frac{\alpha}{2}} = (2^{\frac{2}{\alpha}} (\varepsilon^2 + (\frac{x_{K-1} - x_K}{2})^2))^{\frac{\alpha}{2}}. \end{cases} \quad (45)$$

Since all terms elevated at power $\alpha/2$ are positive, one gets

$$\begin{cases} 2^{\frac{2}{\alpha}} (\varepsilon^2 + (\frac{x_2 - x_1}{2})^2) = \varepsilon^2 + x_1^2 \\ (\frac{x_{i+k} - x_k}{2})^2 = (\frac{x_{k-1} - x_k}{2})^2, \quad \forall k \in \{2, \dots, K-1\} \\ 2^{\frac{2}{\alpha}} (\varepsilon^2 + (\frac{x_{K-1} - x_K}{2})^2) = \varepsilon^2 + (L - x_K)^2 \end{cases} \quad (46)$$

One real solution $\underline{x}$ verifying (10) has thus to satisfy (11) when $\alpha > 2$ and (12) when $\alpha = 2$.

C.2 SPRINGER-GAMENETS-2012

F. Mériaux, S. M. Perlaza, S. Lasaulce, Z. Han et V. Poor. *Achievability of Efficient Satisfaction Equilibria in Self-Configuring Networks*. In Game Theory for Networks, vol. 105, pages 1-15. Springer Berlin Heidelberg, 2012.

Achievability of Efficient Satisfaction Equilibria in Self-Configuring Networks

F. Mériaux^{*}, S. M. Perlaza[†], S. Lasaulce^{*}, Z. Han⁺, and H. V. Poor[†]

^{*} Laboratoire des Signaux et Systèmes - LSS (CNRS-SUPELEC-Paris Sud),
Gif-sur-Yvette, 91192 France.

[†] Department of Electrical Engineering, Princeton University,
Princeton, NJ 08542 USA

⁺ Department of Electrical and Computer Engineering, University of Houston,
Houston, TX 77004 USA

Key words: Supermodular games, Power control, Efficient Satisfaction Equilibrium, Games in Satisfaction Form.

Abstract. In this paper, a behavioral rule that allows radio devices to achieve an efficient satisfaction equilibrium (ESE) in fully decentralized self-configuring networks (DSCNs) is presented. The relevance of ESE in the context of DSCNs is that at such state, radio devices adopt a transmission/receive configuration such that they are able to simultaneously satisfy their individual quality-of-service (QoS) constraints. An ESE is also an efficient network configuration, i.e., individual QoS satisfaction is achieved by investing the lowest possible effort. Here, the notion of effort refers to a preference each radio device independently establishes among its own set of actions. In particular, the proposed behavioral rule requires less information than existing rules, as in the case of the classical best response dynamics and its variants. Sufficient conditions for convergence are presented in a general framework. Numerical results are provided in the context of a particular uplink power control scenario, and convergence from any initial action profile to an ESE is formally proved in this scenario. This property ensures the proposed rule to be robust to the dynamic arrival or departure of radio devices in the network.

1 Introduction

A decentralized self-configuring network (DSCN) is basically an infrastructure-less communication system in which radio devices autonomously choose their own transmit/receive configuration in order to guarantee reliable communication. In particular, a transmit/receive configuration can be described in terms of power allocation policies, coding-modulation schemes, scheduling policies, decoding order, etc. Typical examples of DSCNs are wireless sensor networks, short range networks in the ISM bands (e.g., Wi-Fi, Bluetooth, ZigBee, etc.), femto-cell

networks (e.g., femto cells in LTE-A) and adhoc networks in general. The underlying feature of DSCNs is that transmitters directly communicate with their respective receivers without the intervention of a central controller. Thus, the main limitation of these networks to actually provide QoS is the mutual interference arising from the uncoordinated interaction of radio devices subject to mutual interference. Within this context, the notion of QoS provisioning translates into the need for designing behavioral rules such that radio devices autonomously adapt their transmission configurations in order to meet the minimum requirements for their communications to take place satisfactorily. In particular, similar reconfigurable capabilities have been already mentioned in [6] in the context of cognitive radios.

In general, the decentralized nature of the QoS provisioning task in DSCN has fostered the use of tools from game theory (GT) [7, 5], strategic learning theory [9], distributed optimization and variational inequality theory [11] to the analysis of QoS provisioning in this scenario. In this paper, we focus on a particular formulation of the QoS provisioning problem, namely games in satisfaction form [8]. More specifically, we provide behavioral rules that allow radio devices to achieve an efficient satisfaction equilibrium (ESE) in DSCNs. The notion of ESE, as introduced in [8], refers to a network state in which all the network devices are able to satisfy their individual QoS constraints by investing the minimum effort. Often, we associate the notion of high effort with transmit/receive configurations that might represent an additional waste of energy to satisfy the individual QoS constraints. In this context, one of the main contributions of this paper is the introduction of a behavioral rule that allows the network to achieve an ESE using only local information. Another important contribution is a set of sufficient conditions to observe the convergence to an ESE of the proposed rule.

In order to show the potential of our contributions in the context of DSCNs, we consider a particular scenario of power control in the uplink of a single-cell system in which devices must guarantee a minimum signal to interference plus noise ratio (SINR). Interestingly, we highlight that in this particular scenario, the proposed behavioral rule converges to an ESE independently of the initial state of the network. This result contrasts with the existing literature. For instance, in [1], Altman *et al.* studied the problem in the general framework of compact sublattices as action sets. Therein, under the assumption that a solution to the problem exists, they established that a simple behavioral rule known in game theory as the best response dynamics (BRD) [4] only converge to the solution from particular starting points. When the transmit power sets are continuous, Yates *et al.* proved that the BRD converge from any initial point in [14]. In the case of discrete actions sets, an algorithm close to the BRD is proposed in [12]. However, there are still conditions on the starting point to ensure the convergence of the algorithm.

The remainder of this paper unfolds as follows. In Sec. 2, we revisit the notion of satisfaction equilibrium (SE) and ESE and we formulate the QoS provisioning problem in the most general terms. In Sec. 3, we describe our main contribution: a behavioral rule that allows DSCNs to converge to an ESE, when action

sets correspond to compact sublattices. In Sec. 4, we present numerical results in a particular scenario as described above in order to verify our theoretical contributions. Finally, we conclude our work in Sec. 5.

2 QoS Provisioning and Games in Satisfaction Form

2.1 QoS Problem Formulation

Consider a DSCN comprising a set $\mathcal{K} = \{1, \dots, K\}$ of K transmitter/receiver pairs to which we refer as links. Each radio device autonomously chooses its optimal transmit/receive configuration in order to satisfy its own QoS requirements. Here, we denote by $k \in \mathcal{K}$ the k -th link, independently of whether it is the transmitter or the receiver that is the device performing the self-adaptation of the link configuration. We denote by a_k the transmit/receive configuration adopted by the link k , and we denote by $\mathcal{A}_k$ the set of all its possible choices. For all $k \in \mathcal{K}$, $\mathcal{A}_k$ is assumed to be a compact sublattice, as in [1, 13]. $\mathcal{A} = \mathcal{A}_1 \times \dots \times \mathcal{A}_K$ represents the set of all configuration profiles. This structure has the advantage of comprising both compact continuous sets and discrete sets¹. We denote by $\mathbf{a}_{-k} = (a_1, \dots, a_{k-1}, a_{k+1}, \dots, a_K)$ the vector obtained by dropping the k -th component of the vector $\mathbf{a}$. We denote the space in which the vector $\mathbf{a}_{-k}$ exists by $\mathcal{A}_{-k}$. With a slight abuse of notation, we write the vector $\mathbf{a}$ as $(a_k, \mathbf{a}_{-k})$, in order to emphasize its k -th component. A transmit/receive configuration can be described by parameters such as the power allocation policy, modulation scheme, constellation size, decoding order, scheduling policy, etc. The instantaneous performance of radio device k is determined by a set of Q_k functions

$$\begin{cases} u_k^{(1)} : \mathcal{A} \rightarrow \mathbb{R}, \\ \vdots \\ u_k^{(Q_k)} : \mathcal{A} \rightarrow \mathbb{R}. \end{cases}$$

Typical performance metrics are transmission rate, transmission delay, bit error rate, energy efficiency, or any combination of those. We denote the minimum and maximum acceptable values of the performance metric $u_k^{(q_k)}$ by $\Gamma_k^{(q_k, \min)}$ and $\Gamma_k^{(q_k, \max)}$, respectively. Thus, we say that the configuration profile $\mathbf{a} \in \mathcal{A}$ satisfies the QoS constraints of the DSCN if for all link k the following set of inequalities are satisfied :

$$\begin{cases} \Gamma_k^{(1, \min)} < u_k^{(1)}(a_k, \mathbf{a}_{-k}) < \Gamma_k^{(1, \max)}, \\ \vdots \\ \Gamma_k^{(Q_k, \min)} < u_k^{(Q_k)}(a_k, \mathbf{a}_{-k}) < \Gamma_k^{(Q_k, \max)}. \end{cases} \quad (1)$$

¹ The results of Sec. 3.2 and Sec. 3.4 apply to the general framework of compact sublattices whereas the results of Sec. 3.5 apply only to discrete configuration sets.

Note that the performance metrics of link k depend not only on its own configuration a_k but also on the configurations $\mathbf{a}_{-k}$ adopted by all the other links. Thus, in order to ease our notation, we define the correspondence $f_k: \mathcal{A}_{-k} \rightarrow 2^{\mathcal{A}_k}$ that determines all the possible configurations of player k that satisfies its QoS constraints. That is $\forall a_k \in \mathcal{A}_k$

$$\begin{aligned} a_k \in f_k(\mathbf{a}_{-k}) &\Leftrightarrow \\ \forall q \in \{1, \dots, Q_k\}, \Gamma_k^{(q, \min)} &< u_k^{(q)}(a_k, \mathbf{a}_{-k}) < \Gamma_k^{(q, \max)}. \end{aligned} \quad (2)$$

The problem of all the links wanting to satisfy their QoS constraints at the same time can naturally be described as a game.

2.2 Game formulation

As defined in [8], a game in satisfaction form is fully described by the following triplet

$$\widehat{\mathcal{G}} = (\mathcal{K}, \{\mathcal{A}_k\}_{k \in \mathcal{K}}, \{f_k\}_{k \in \mathcal{K}}). \quad (3)$$

In this triplet, $\mathcal{K}$ represents the set of players, $\mathcal{A}_k$ is the strategy set of player $k \in \mathcal{K}$, and the correspondence f_k determines the set of actions of player k that allows its satisfaction given the actions played by all the other players. A strategy profile is denoted by vector $\mathbf{a} = (a_1, \dots, a_K) \in \mathcal{A}$. In general, an important outcome of a game in satisfaction form is the one where all players are satisfied, that is, an SE. The notion of SE was formulated as a fixed point in [8] as follows:

Definition 1 (Satisfaction Equilibrium). *An action profile $\mathbf{a}^+$ is an equilibrium for the game $\widehat{\mathcal{G}} = (\mathcal{K}, \{\mathcal{A}_k\}_{k \in \mathcal{K}}, \{f_k\}_{k \in \mathcal{K}})$ if*

$$\forall k \in \mathcal{K}, \quad a_k^+ \in f_k(\mathbf{a}_{-k}^+). \quad (4)$$

As we shall see, the SE is often not unique and thus, there might exist some SEs that are of particular interest. In the following, we introduce the notion of an efficient SE (ESE). For this intent, we consider a cost function for each player of the game, in order to model the notion of effort or cost associated with a given action choice. For all $k \in \mathcal{K}$, the cost function $c_k: \mathcal{A}_k \rightarrow [0, 1]$ satisfies the following condition: $\forall (a_k, a'_k) \in \mathcal{A}_k^2$, it holds that

$$c_k(a_k) < c_k(a'_k), \quad (5)$$

if and only if, a_k requires a lower effort than action a'_k when it is played by player k . Under the notion of effort, the set of SEs that are of particular interest are those that require the lowest individual efforts. We formalize this notion of equilibrium using the following definition.

Definition 2 (Efficient Satisfaction Equilibrium). *An action profile $\mathbf{a}^*$ is an ESE for the game $\widehat{\mathcal{G}} = (\mathcal{K}, \{\mathcal{A}_k\}_{k \in \mathcal{K}}, \{f_k\}_{k \in \mathcal{K}})$, with cost functions $\{c_k\}_{k \in \mathcal{K}}$, if*

$$\forall k \in \mathcal{K}, \quad a_k^* \in f_k(a_{-k}^*) \quad (6)$$

and

$$\forall k \in \mathcal{K}, \quad \forall a_k \in f_k(a_{-k}^*), \quad c_k(a_k) \geq c_k(a_k^*). \quad (7)$$

Note that the effort associated by each player with each of its actions does not depend on the choice of effort made by other players. Here, we have left players to individually choose their cost functions, which adds another degree of freedom to the modeling of the QoS problem in DSCNs.

Note also that a game in satisfaction form is not a game with a constrained set of actions, as is the case in the formulation presented in [3]. Here, a player can use any of its actions independently of all the other players. The dependency on the other players' actions enters through whether the player under study is satisfied or not.

2.3 Power control game

In the rest of this paper, we use the context of uplink power control in a single-cell as a case study. Although most of our results apply in a general context, we concentrate in the uplink power control problem as presented in [12, 14], to illustrate our results.

Consider K transmitter/receiver pairs denoted by index $k \in \mathcal{K}$. For all $k \in \mathcal{K}$, transmitter k uses power level $p_k \in \mathcal{A}_k$, with $\mathcal{A}_k$ generally defined as a compact sublattice. For each player $k \in \mathcal{K}$, we denote $p_k^{\min}$ and $p_k^{\max}$ the minimum and the maximum power levels in $\mathcal{A}_k$, respectively. For every couple $(i, j) \in \mathcal{K}^2$, we denote by g_{ij} the channel gain coefficient between transmitter i and receiver j . The considered metric for each pair k is the Shannon rate given by

$$u_k(p_k, \mathbf{p}_{-k}) = \log_2 \left(1 + \frac{p_k g_{kk}}{\sigma_k^2 + \sum_{j \neq k} p_j g_{jk}} \right) \quad [\text{bps/Hz}], \quad (8)$$

where σ_k^2 is the noise variance at receiver k .

The QoS requirement for each pair k is to have a channel capacity $u_k(p_k, \mathbf{p}_{-k})$ higher than a given threshold Γ_k bps/Hz. The satisfaction correspondence of link k is then

$$\begin{aligned} f_k(\mathbf{p}_{-k}) &= \{p_k \in \mathcal{A}_k \mid u_k(p_k, \mathbf{p}_{-k}) \geq \Gamma_k\} \\ &= \left\{ p_k \in \mathcal{A}_k \mid p_k \geq (2^{\Gamma_k} - 1) \frac{\sigma_k^2 + \sum_{j \neq k} p_j g_{jk}}{g_{kk}} \right\}. \end{aligned} \quad (9)$$

3 Convergence to an Efficient Satisfaction Equilibrium

In this section, we provide sufficient conditions for convergence of the BRD and the robust blind response dynamics (RBRD) to an ESE of the game $\hat{\mathcal{G}} = (\mathcal{K}, \{\mathcal{A}_k\}_{k \in \mathcal{K}}, \{f_k\}_{k \in \mathcal{K}})$, with cost functions $\{c_k\}_{k \in \mathcal{K}}$.

3.1 Best response dynamics

In the context of a game in satisfaction form $\widehat{\mathcal{G}} = (\mathcal{K}, \{\mathcal{A}_k\}_{k \in \mathcal{K}}, \{f_k\}_{k \in \mathcal{K}})$, with cost functions $\{c_k\}_{k \in \mathcal{K}}$, we define the best response (BR) correspondence of player k , given that the other players adopt the reduced action profile $\mathbf{a}_{-k}$, as follows:

$$BR_k(\mathbf{a}_{-k}) = \arg \min_{a_k \in f_k(\mathbf{a}_{-k})} c_k(a_k). \quad (10)$$

We consider a BRD defined as the behavioral rule in which players sequentially update their action following the Gauss-Seidel method [2]. At step $n + 1$ of the algorithm, all the players sequentially update their actions with the following rule:

$$a_k^{(n+1)} = BR_k(a_1^{(n+1)}, \dots, a_{k-1}^{(n+1)}, a_{k+1}^{(n)}, \dots, a_K^{(n)}). \quad (11)$$

For a discrete set of actions, the BRD can be compared to the asynchronous version of the Minimum Feasible Value Assignment (MFVA) algorithm presented in [12]. The difference is that in [12], players only move to their optimal satisfying action if they are not satisfied with actions played at the previous step. In the BRD, players move to their optimal satisfying action independently of their satisfaction at the previous step.

3.2 Convergence of the BRD

To study the convergence of the BRD, we first define some notation of interest. Let $\mathbf{a} = (a_1, \dots, a_N)$ and $\mathbf{b} = (b_1, \dots, b_N)$ be two action profiles and let $\mathbf{c} = \mathbf{a} \vee \mathbf{b}$ denote the maximum of $(\mathbf{a}, \mathbf{b})$ component wise, i.e., $\mathbf{c} = (c_1, \dots, c_N)$ with $c_n = \max(a_n, b_n) \ \forall n \in \{1, \dots, N\}$. In a similar way, $\mathbf{a} \wedge \mathbf{b}$ denotes $\min(\mathbf{a}, \mathbf{b})$ component wise.

Definition 3 (S-modularity). *The function $g: \mathcal{A} \rightarrow \mathbb{R}$ is said to be supermodular if for any $\mathbf{a}, \mathbf{b} \in \mathcal{A}$*

$$g(\mathbf{a} \wedge \mathbf{b}) + g(\mathbf{a} \vee \mathbf{b}) \geq g(\mathbf{a}) + g(\mathbf{b}), \quad (12)$$

and said to be submodular if

$$g(\mathbf{a} \wedge \mathbf{b}) + g(\mathbf{a} \vee \mathbf{b}) \leq g(\mathbf{a}) + g(\mathbf{b}). \quad (13)$$

In the case of the cost function defined in (5), c_k depends only on the actions of player k . Hence, c_k is both supermodular and submodular. As a result, (12) and (13) are equalities.

Definition 4 (Ascending and descending properties). *The correspondence f_k is said to possess the ascending property (respectively the descending property) if for any two elements $\mathbf{a}_{-k}$ and $\mathbf{a}'_{-k}$ of the set $\mathcal{A}_{-k}$, with $\mathbf{a}_{-k} = \mathbf{a}_{-k} \wedge \mathbf{a}'_{-k}$ implies that $\forall a_k \in f_k(\mathbf{a}_{-k})$ and $\forall a'_k \in f_k(\mathbf{a}'_{-k})$,*

$$\begin{cases} \min(a_k, a'_k) \in f_k(\mathbf{a}_{-k}), \\ \max(a_k, a'_k) \in f_k(\mathbf{a}'_{-k}), \end{cases} \quad (14)$$

or for the descending property

$$\begin{cases} \max(a_k, a'_k) \in f_k(\mathbf{a}_{-k}), \\ \min(a_k, a'_k) \in f_k(\mathbf{a}'_{-k}). \end{cases} \quad (15)$$

An important consequence of the ascending (or descending) property is that

$$\forall \mathbf{a}_{-k} \in \mathcal{A}_{-k}, f_k(\mathbf{a}_{-k}) \neq \emptyset. \quad (16)$$

The definition of an ascending set can easily be understood in the context of distributed power control. In such a context, the ascending property means that if all the other players increase their powers, player k also has to increase its own power if it wants to remain satisfied. Also note that if the ascending property is ensured, then there is always at least one satisfying power level for player k , whatever the other players are playing. In particular, when all the players are at maximum power levels, there exists a satisfying power for player k , which is a strong assumption.

Proposition 1. *Assume that for all $k \in \mathcal{K}$, $f_k(\cdot)$ is nonempty and compact for all the values of their arguments, $f_k(\cdot)$ has either the ascending or the descending property and $f_k(\cdot)$ is continuous. Then the following holds:*

- (i) *An ESE exists.*
- (ii) *If the dynamics start with the action profile associated with the highest or lowest effort in $c_k(\cdot)$, for all $k \in \mathcal{K}$, the BRD converge monotonically to an ESE.*
- (iii) *If the dynamics start from an SE, the trajectory of the best response converges to an ESE. It monotonically evolves in all components.*
- (iv) *In a two-player game, the BRD converge to an ESE from any starting point.*

The proof of Prop. 1 comes from Th. 1 in [1] and Th. 2.3 in [13]. We simply have to verify that the right assumptions hold for the ascending case and the descending case:

- Let $f_k(\cdot)$ be ascending for all $k \in \mathcal{K}$. c_k is a cost function player k wants to minimize, in particular c_k is a submodular function, and thus $-c_k$ is a supermodular function player k wants to maximize and Th. 1 from [1] holds, i.e., (i, ii, iii) in Prop. 1 are ensured when the sets are ascending.
- Let $f_k(\cdot)$ be descending for all $k \in \mathcal{K}$. A similar reasoning can be made: c_k is a submodular function player k wants to minimize and the same theorem holds as well, i.e., (i, ii, iii) in Prop. 1 are ensured when the sets are descending.

In both ascending and descending cases, (iv) in Prop. 1 is obtained from Th. 2.3 in [13].

3.3 BRD in the uplink power control game

In the general framework of compact sublattices as strategy sets (including continuous and discrete action sets), the BRD converge only from given starting points (see [1, 13]). However, in the uplink power control problem, it has been shown in [10, 14] that when strategy sets are continuous, the BRD converge from any initial point. When strategy sets are discrete, the convergence of the BRD from any initial point to an equilibrium is not guaranteed. In [12], it is shown that the MFVA converges only when all the transmitters start at their lowest power levels. In the following, we consider a 3-player uplink power control game to illustrate the non-convergence of the BRD from a particular initial action profile.

Example 1. In this example, we refer to the notation introduced in Sec. 2.3. Let us consider $K = 3$ pairs of transmitters/receivers. For all $k \in \mathcal{K}$, transmitter k uses power level $a_k \in \{p^{\min}, p^{\max}\}$. Given the constraints from Sec. 2.3, let consider channel gains such that

$$\begin{aligned} f_1(p^{\min}, p^{\min}) &= f_3(p^{\min}, p^{\min}) = \{p^{\min}, p^{\max}\}, \\ f_1(p^{\min}, p^{\max}) &= f_3(p^{\min}, p^{\max}) = \{p^{\min}, p^{\max}\}, \\ f_1(p^{\max}, p^{\min}) &= f_3(p^{\max}, p^{\min}) = \{p^{\max}\}, \\ f_1(p^{\max}, p^{\max}) &= f_3(p^{\max}, p^{\max}) = \{p^{\max}\}, \end{aligned} \quad (17)$$

and

$$\begin{aligned} f_2(p^{\min}, p^{\min}) &= \{p^{\min}, p^{\max}\}, \\ f_2(p^{\min}, p^{\max}) &= \{p^{\max}\}, \\ f_2(p^{\max}, p^{\min}) &= \{p^{\min}, p^{\max}\}, \\ f_2(p^{\max}, p^{\max}) &= \{p^{\max}\}. \end{aligned} \quad (18)$$

We can check that f_k has the ascending property for all $k \in \mathcal{K}$. For each pair k , the cost of the power level is given by the identity cost function $c_k(a_k) = a_k$. This game has two ESEs:

- $(p^{\min}, p^{\min}, p^{\min})$ where all the players transmit at their lowest power level. No player has interest in deviating from its action since any other action has a higher cost (even though the player would remain satisfied).
- $(p^{\max}, p^{\max}, p^{\max})$ where all the players have to transmit at maximum power to be satisfied. If one deviates from its action, it will not be satisfied anymore.

But depending on the initial action profile of the BRD, the BRD may not converge to an ESE. For instance, assume that the BRD start at $\mathbf{p}^{(0)} = (p^{\max}, p^{\min}, p^{\max})$. At step 1, player 1 chooses the action that minimizes $c_1(\cdot)$ given the previous actions of the other players $\mathbf{p}_{-1}^{(0)} = (p^{\min}, p^{\max})$, i.e.,

$$p_1^{(1)} = BR_1(p^{\min}, p^{\max}) = p^{\min}.$$

Player 2 chooses the action that minimizes $c_2(\cdot)$ given the most recent actions of the other players $(p_1^{(1)}, \mathbf{p}_{-(1,2)}^{(0)}) = (p^{\min}, p^{\max})$, i.e.,

$$p_2^{(1)} = BR_2(p^{\min}, p^{\max}) = p^{\max}.$$

Player 3 chooses the action that minimizes $c_3(\cdot)$ given $(p_1^{(1)}, p_2^{(1)}) = (p^{\min}, p^{\max})$, i.e.,

$$p_3^{(1)} = BR_3(p^{\min}, p^{\max}) = p^{\min}.$$

At step 2, player 1 chooses the action that minimizes $c_1(\cdot)$ given the previous actions of the other players $\mathbf{p}_{-1}^{(1)} = (p^{\max}, p^{\min})$, i.e.,

$$p_1^{(2)} = BR_1(p^{\max}, p^{\min}) = p^{\max}.$$

Player 2 chooses the action that minimizes $c_2(\cdot)$ given the most recent actions of the other players $(p_1^{(2)}, \mathbf{p}_{-(1,2)}^{(1)}) = (p^{\max}, p^{\min})$, i.e.,

$$p_2^{(2)} = BR_2(p^{\max}, p^{\min}) = p^{\min}.$$

Player 3 chooses the action that minimizes $c_3(\cdot)$ given $(p_1^{(2)}, p_2^{(2)}) = (p^{\max}, p^{\min})$, i.e.,

$$p_3^{(2)} = BR_3(p^{\max}, p^{\min}) = p^{\max}.$$

The algorithm is back at the starting point, and it is clear that it will continue in this infinite loop.

3.4 Robust blind response dynamics

The BRD have significant drawbacks. First, it was just shown that in a K -player game with $K > 2$, the dynamics may not converge to an ESE depending on the initial action profile. Second, to determine the BR, each player has to know the set $f_k(\mathbf{a}_{-k}) \forall \mathbf{a}_{-k} \in \mathcal{A}_{-k}$. To overcome these drawbacks, we propose a new algorithm that requires less information about the game for each player and can still be proven to converge to an ESE. Let us start by defining the robust blind response (RBR) by $RBR_k: \mathcal{A} \rightarrow \mathcal{A}_k$, such that :

$$(a_k, \mathbf{a}_{-k}) \rightarrow \begin{cases} a'_k, & \text{if } a'_k \in f_k(\mathbf{a}_{-k}), a_k \in f_k(\mathbf{a}_{-k}) \text{ and } c_k(a'_k) \leq c_k(a_k), \\ a'_k, & \text{if } a'_k \in f_k(\mathbf{a}_{-k}) \text{ and } a_k \notin f_k(\mathbf{a}_{-k}), \\ a_k, & \text{otherwise,} \end{cases} \quad (19)$$

with action a'_k being randomly chosen in $\mathcal{A}_k$, such that $\forall a_k \in \mathcal{A}_k, \Pr(a'_k = a_k) > 0$. Each time the RBR is used, a player $k \in \mathcal{K}$ randomly chooses an action in its strategy set $\mathcal{A}_k$ without taking into account the constraints of other players. Player k only has to know if the new action and the previous one allow the satisfaction of its individual constraints and to compare their respective costs. If both actions allow the satisfaction of the constraints, it chooses the one with the lowest cost. If the new action allows the satisfaction of the individual constraints whereas the previous one does not, it moves to the new action. Otherwise, it keeps the same action. When all the players sequentially use the RBR such that $\forall k \in \mathcal{K}$

$$a_k^{(n+1)} = RBR_k(a_1^{(n+1)}, \dots, a_{k-1}^{(n+1)}, a_{k+1}^{(n)}, \dots, a_K^{(n)}), \quad (20)$$

we refer to these dynamics as the RBR dynamics (RBRD). Our main result in this section is stated in the following theorem.

Theorem 1. *Assume that for all $k \in \mathcal{K}$, $f_k(\cdot)$ is nonempty and compact for all the values of their arguments, $f_k(\cdot)$ has the ascending property and it is continuous, and $c_k(\cdot)$ is strictly increasing. Then, the following holds:*

- (i) *If the dynamics start from an SE, the sequence of RBRs converges to an ESE. It monotonically decreases in all components.*
- (ii) *If the dynamics start with the actions associated with the highest effort in $c_k(\cdot)$, $\forall k \in \mathcal{K}$, the sequence of RBRs converges monotonically to an ESE.*
- (iii) *In a two-player game, the sequence of RBRs converge to an ESE from any starting point.*

Proof. Applying Prop. 1, we know that there exists an ESE for the game $\widehat{\mathcal{G}} = (\mathcal{K}, \{\mathcal{A}_k\}_{k \in \mathcal{K}}, \{f_k\}_{k \in \mathcal{K}})$. The convergence of the RBRD to an ESE is proven in two steps. First, we show for (i, ii, iii) that the RBRD converge to a fixed point. Second, we explain why this fixed point has to be an ESE.

- (i) Assume that the dynamics start from an SE: $\mathbf{a}^{SE}$ and this SE is not an ESE (otherwise, the convergence is trivial). Let player $k \in \mathcal{K}$ be the first player to actually change its action at step n to $a_k^{(n)}$; necessarily this action has a lower cost than a_k^{SE} because a satisfied player can only move to another satisfying action with a lower cost. Let the next player to move be denoted by j . From its point of view $(a_k^{(n)}, \mathbf{a}_{-\{k,j\}}^{SE}) = (a_k^{(n)}, \mathbf{a}_{-\{k,j\}}^{SE}) \wedge \mathbf{a}_{-j}^{SE}$. Hence, due to the ascending property of f_j and the strict monotony of c_j , necessarily its new action $a_j^{(n')} \leq a_j^{SE}$, and so forth. For each $k \in \mathcal{K}$ the sequence $\{a_k^{(n)}\}_{n \in \mathbb{N}}$ is decreasing in a compact set. Thus, the algorithm converges to a limit.
- (ii) Assume that the dynamics start from action profile $\mathbf{a}^{\max} = (a_1^{\max}, \dots, a_K^{\max})$ and this point is not an SE (otherwise refer to (i)). Let player k update its action first, at step n . Necessarily, its updated action $a_k^{(n)}$ is lower than $a_k^{\max}$. Then $\forall j \neq k, j \in \mathcal{K}$

$$(\mathbf{a}_{-\{j,k\}}^{\max}, a_k^{(n)}) = (\mathbf{a}_{-\{j,k\}}^{\max}, a_k^{(n)}) \wedge \mathbf{a}_{-j}^{\max}. \quad (21)$$

Due to the ascending property of f_j and the strict monotony of c_j , the update of player j is hence lower than $a_j^{\max}$, and so forth. Again, for each player $k \in \mathcal{K}$, the sequence of action $\{a_k^{(n)}\}_{n \in \mathbb{N}}$ is decreasing in a compact set and the algorithm converges to a limit.

- (iii) In a two-player game, assume the dynamics start from a random action profile $(a_1^{(0)}, a_2^{(0)})$. Assume player 1 is the first player that updates its action to get satisfied, at step n . The action profile is then $(a_1^{(n)}, a_2^{(0)})$. In the next move, either the same player 1 decreases its action, remaining satisfied, or player 2 moves to an action that satisfies it, leading to an action profile $(a_1^{(n)}, a_2^{(n')})$. If this profile is an SE, the dynamics converge according to (i). Otherwise

player 1 is no longer satisfied and has to update its action. If $a_2^{(n')} < a_2^{(0)}$, then due to the ascending property and the strict monotonicity of c_1 , player 1 will only move to a lower action than $a_1^{(n)}$. Then player 2 will also have to move to a lower action than $a_2^{(n')}$ for analogous reasons, and so forth. The sequences $\{a_1^{(n)}\}_{n \in \mathbb{N}}$ and $\{a_2^{(n)}\}_{n \in \mathbb{N}}$ are hence decreasing in a compact set, so they converge to a limit. If $a_2^{(n')} > a_2^{(0)}$, the sequences are increasing in a compact set and converge as well.

We now have to prove that a fixed point is an ESE. Consider that $\mathbf{a}^*$ is a fixed point for RBR_k , $\forall k \in \mathcal{K}$. By the definition of RBR_k this means that there exists no $a_k \in \mathcal{A}_k$ such that $a_k \in f_k(\mathbf{a}_{-k}^*)$ and $c_k(a_k) \leq c_k(a_k^*)$, which is exactly the definition of the ESE. This completes the proof. ■

The main advantage of these dynamics over BRD in a general framework is that the former require only local information and the knowledge of an explicit expression for f_k is no longer relevant. Only the knowledge of whether the corresponding player is satisfied or not is sufficient to implement the RBR.

3.5 RBRD in the uplink power control game

A very interesting property occurs for the RBR in the uplink power control game with discrete action sets.

Theorem 2. *In the power allocation game defined above in Sec. 2.3, with discrete action sets, i.e., $\forall k \in \mathcal{K}$, $\mathcal{A}_k = \{p_k^{(1)}, \dots, p_k^{(N_k)}\}$ with N_k the number of power levels in action set $\mathcal{A}_k$, the RBRD converge to an ESE from any starting point.*

Proof. We show in this proof that from any starting point of the dynamics, there is a non-null probability that the dynamics move to a particular SE with a given way. Note that the particular sequence of events we describe here is not always the way the dynamics run. It is simply a sequence that can occur with a non-null probability, but there are many other possible sequences that lead to an SE.

Assume $\mathbf{p}^{(0)} = (p_1^{(0)}, \dots, p_k^{(0)})$ is the starting power profile of the dynamics. Consider all the unsatisfied players at this point and assume that they all move to their maximum possible power levels (this may happen with a non-null probability). These levels satisfy them since the ascending property gives us

$$\forall k \in \mathcal{K}, \forall \mathbf{p}_{-k} \in \mathcal{A}_{-k}, p_k^{\max} \in f_k(\mathbf{p}_{-k}). \quad (22)$$

This increase of power levels may cause some of the satisfied players at the starting point not to be satisfied anymore. We also assume that these players move to their maximum power levels. And the same is done until no unsatisfied player remains. So we get a power profile made of the highest power levels for some of the players and the initial power levels for the others, and every player is satisfied at this point: it is an SE.

Finally, from (i) of Th. 1, the dynamics converge to an ESE, which completes the proof. ■

Th. 2 highlights a very interesting property of the RBRD when players enter or quit the game (or when the channel coefficients vary). Indeed, if K transmitters are in any given ESE $\mathbf{p}^*$ and a new transmitter enters the game, a new game starts with $K + 1$ players. Thus, from Th. 2, it can be stated that convergence to a new ESE, if it exists, is ensured from the new starting point $(\mathbf{p}^*, p_{k+1})$.

4 Numerical results

In this section, we provide numerical results for the uplink power control game with discrete action sets as defined in Sec. 2.3.

In Fig. 1, we show the sequences of actions converging to an ESE for the RBRD in a 2-player power control game. The colored region is the satisfaction region, i.e., the region allowing both players to be satisfied. The coloring of this region follows the sum of the costs for each player. The RBR first converges to the satisfaction region, then converges to an ESE while remaining in the satisfaction region.

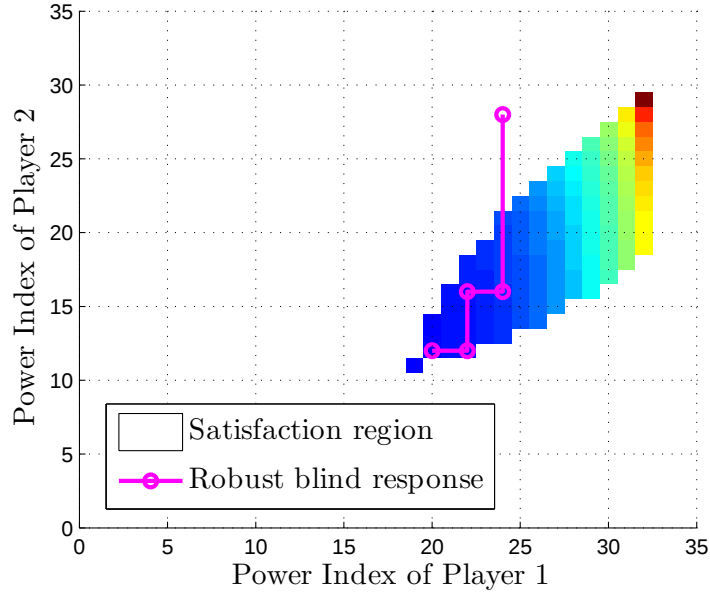


Fig. 1. Sequence of power indices for the RBRD in the uplink 2-player power control game. The colored region is the satisfaction region, i.e., the region where the two players mutually satisfy their constraints.

The scenario we consider in Fig. 2 and Fig. 3 highlights the advantages of RBRD over the BRD in a 3-player game: during the first 200 steps, only

transmitters 1 and 3 are in the game, then transmitter 2 joins them for the 200 next steps, and finally transmitter 3 leaves for the last 200 steps. On each of the two figures, we show the sequence of power indices for the three players, knowing that each action set is made of $N_k = 32$ possible power levels from 10^{-6}W to 10^{-2}W . We also show the satisfaction states of the three players: for each step of the dynamics, if all the player are satisfied, the satisfaction state is 1, otherwise it is 0. Fig. 2 and Fig. 3 correspond to the behavior of the BRD and the RBRD, respectively. The channel parameters and the starting points of the two simulations are exactly the same. Channel gains are $g_{22} = 10^{-5}$, $g_{11} = g_{33} = g_{13} = g_{21} = g_{32} = 10^{-6}$, $g_{12} = g_{23} = g_{31} = 10^{-7}$, and transmitters 1, 2, and 3 start at power levels 10^{-3}W , $10^{-5/2}\text{W}$, and $10^{-9/4}\text{W}$, respectively. The utility constraints I_1 , I_2 , and I_3 are taken as 1.2 bps/Hz, 1.5 bps/Hz, and 1.2 bps/Hz, respectively. The variance of the noise is fixed at 10^{-10}W for all the transmitters. It is interesting to notice that the BRD converge to ESE during the first and third phase but when transmitter 2 enters the game in the second phase, the BRD do not converge to an ESE. Instead, they enter a loop and we can see that the transmitters are not satisfied. Concerning the RBRD, although their convergence time is longer, they converge in the three phases and another interesting fact is that transmitters are satisfied during a longer amount of time compared to the BRD.

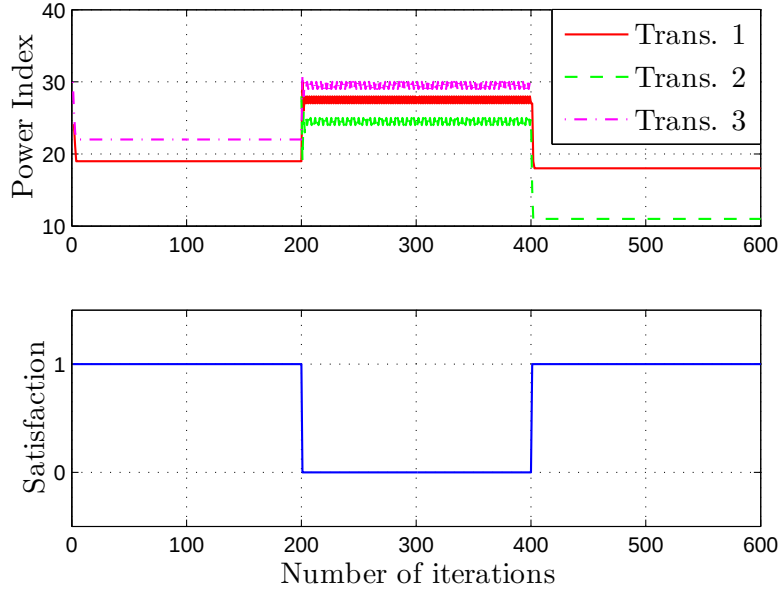


Fig. 2. Sequences of power indices and satisfaction states for the BRD in the 3-player uplink power control game.

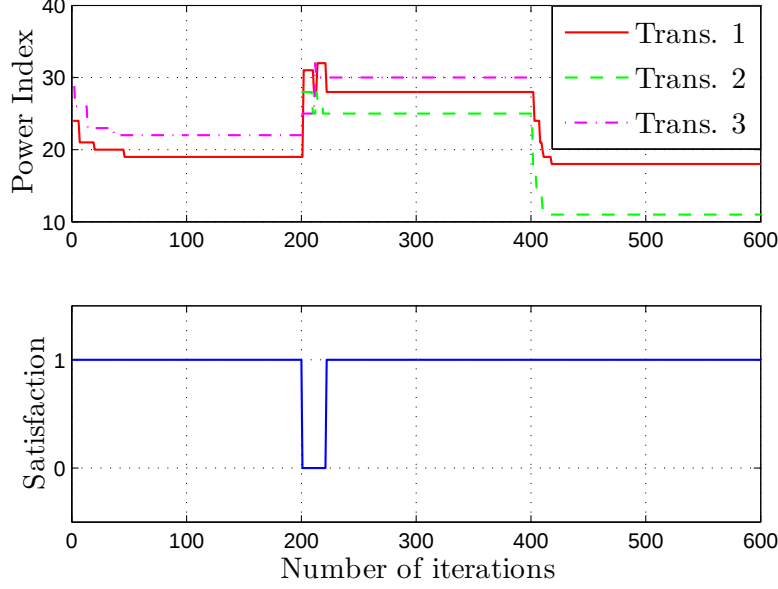


Fig. 3. Sequences of power indices and satisfaction states for the RBRD in the 3-player uplink power control game.

5 Conclusion and future work

In this work, we have proposed a behavior rule that converges to an ESE in the general framework of compact sublattices as actions sets. Compared to the BRD, the proposed rule requires far less information although its convergence time is longer. Applying this rule to the uplink power control game with discrete actions sets has been shown to be of great interest since the dynamics are proven to converge to an ESE from any starting action profile. This particular feature allows the proposed rule to be robust to the entrance or the exit of players in the power control game.

However, a strong assumption of this work is to assume that for every player, for any action profile of the other players, there exists an action satisfying the considered player. In the power control game, it would be more relevant to take into account scenarios in which the power levels of the other players are too high and a given player cannot be satisfied for any action it can play. Hence, a natural perspective of this work is to relax this assumption and study the convergence of the dynamics in this context.

References

1. E. Altman and Z. Altman. S-modular games and power control in wireless networks. *IEEE Transactions on Automatic Control*, 48(5):839–842, May 2003.
2. D. P. Bertsekas and J. N. Tsitsiklis. *Parallel and Distributed Computation: Numerical Methods*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1989.
3. G. Debreu. A social equilibrium existence theorem. *Proceedings of the National Academy of Sciences of the United States of America*, 38(10):886–893, October 1952.
4. D. Fudenberg and J. Tirole. *Game Theory*. MIT Press, Cambridge, MA, 1991.
5. Z. Han, D. Niyato, W. Saad, T. Basar, and A. Hjørungnes. *Game Theory in Wireless and Communication Networks: Theory, Models and Applications*. Cambridge University Press, Cambridge, UK, 2011.
6. S. Haykin. Cognitive radio: Brain-empowered wireless communications. *IEEE Journal on Selected Areas in Communications*, 23(2):201–220, 2005.
7. S. Lasaulce and H. Tembine. *Game Theory and Learning in Wireless Networks: Fundamentals and Applications*. Elsevier Academic Press, Waltham, MA, 2011.
8. S. M. Perlaza, H. Tembine, S. Lasaulce, and M. Debbah. Quality-of-service provisioning in decentralized networks: A satisfaction equilibrium approach. *IEEE Journal of Selected Topics in Signal Processing*, 6(2):104–116, Apr. 2012.
9. L. Rose, S. Lasaulce, S.M. Perlaza, and M. Debbah. Learning equilibria with partial information in decentralized wireless networks. *IEEE Communications Magazine*, 49(8):136–142, Aug. 2011.
10. G. Scutari, S. Barbarossa, and D.P. Palomar. Potential games: A framework for vector power control problems with coupled constraints. In *the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toulouse, France, May 2006.
11. G. Scutari, D. P. Palomar, F. Facchinei, and J.-S. Pang. Convex optimization, game theory, and variational inequality theory in multiuser communication systems. *IEEE Signal Processing Magazine*, 27(3):35–49, May 2010.
12. C. C. Wu and D. P. Bertsekas. Distributed power control algorithms for wireless networks. *IEEE Transactions on Vehicular Technology*, 50(2):504–514, Mar. 2001.
13. D. D. Yao. S-modular games, with queueing applications. *Queueing Systems*, 21(3-4):449–475, 1995.
14. R. D. Yates. A framework for uplink power control in cellular radio systems. *IEEE Journal on Selected Areas in Communications*, 13(7):1341–1347, Sep. 1995.

Bibliographie

- [3GPP 2009] 3GPP. *E-UTRA Physical layer procedures (Release 9)*. Technical Specification TS 36.213 V9.0.1, 3GPP, 2009.
- [Achdou 2010] Y. Achdou et I. Capuzzo-Dolcetta. *Mean Field Games : Numerical Methods*. SIAM Journal on Numerical Analysis, vol. 48, no. 3, pages 1136–1162, 2010.
- [Agarwal 2012] M. Agarwal et M. Honig. *Adaptive Training for Correlated Fading Channels With Feedback*. IEEE Transactions on Information Theory, vol. 58, no. 8, pages 5398–5417, 2012.
- [Altman 2003] E. Altman et Z. Altman. *S-Modular Games and Power Control in Wireless Networks*. IEEE Transactions on Automatic Control, vol. 48, no. 5, pages 839–842, 2003.
- [Altman 2009] E. Altman, A. Kumar, C. Singh et R. Sundaresan. *Spatial SINR Games Combining Base Station Placement and Mobile Association*. In IEEE INFOCOM 2009, pages 1629–1637, 2009.
- [Basar 1999] T. Basar et G. J. Olsder. *Dynamic Noncooperative Game Theory*. Classics in Applied Mathematics, SIAM, Philadelphia, 1999.
- [Belmega 2009] E. V. Belmega, S. Lasaulce et M. Debbah. *Power Allocation Games for MIMO Multiple Access Channels with Coordination*. IEEE Transactions on Wireless Communications, vol. 8, no. 5, pages 3182–3192, 2009.
- [Belmega 2011] E. V. Belmega et S. Lasaulce. *Energy-Efficient Precoding for Multiple-Antenna Terminals*. IEEE Transactions on Signal Processing, vol. 59, no. 1, pages 329–340, 2011.
- [Bertsekas 1989] D. P. Bertsekas et J. N. Tsitsiklis. *Parallel and Distributed Computation : Numerical Methods*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1989.
- [Bonneau 2008] N. Bonneau, M. Debbah, E. Altman et A. Hjørungnes. *Non-Atomic Games for Multi-User System*. IEEE Journal on Selected Areas in Communications, vol. 26, no. 7, pages 1047–1058, 2008.
- [Bressan 2011] A. Bressan. *Noncooperative Differential Games*. Milan Journal of Mathematics, vol. 79, no. 2, pages 357–427, 2011.
- [Brown 1951] G. W. Brown. *Iterative Solutions of Games by Fictitious Play*. Activity analysis of production and allocation, vol. 13, no. 1, pages 374–376, 1951.
- [Bush 1955] R. Bush et F. Mosteller. *Stochastic Models for Learning*. Wiley Sons, New York, 1ère édition, 1955.
- [Buzzi 2011] S. Buzzi et D. Saturnino. *A Game-Theoretic Approach to Energy-Efficient Power Control and Receiver Design in Cognitive CDMA Wireless*

- Networks*. IEEE Journal of Selected Topics in Signal Processing, vol. 5, no. 1, pages 137–150, 2011.
- [Charalambous 2005] C. D. Charalambous, S. M. Djouadi et S. Z. Denic. *Stochastic Power Control for Wireless Networks via SDEs : Probabilistic QoS Measures*. IEEE Transactions on Information Theory, vol. 51, no. 12, pages 4396–4401, 2005.
- [Cournot 1963] A. Cournot. *Researches into the principles of wealth* (original : *Recherches sur les principes mathématiques de la théorie des richesses*, 1838). Irwin Paperback Classics in Economics, 1963.
- [Cover 2006] T. M. Cover et J. A. Thomas. *Elements of Information Theory*. Wiley-Interscience, New York, 2nde édition, 2006.
- [Debreu 1952] G. Debreu. *A Social Equilibrium Existence Theorem*. Proceedings of the National Academy of Sciences of the United States of America, vol. 38, no. 10, pages 886–893, 1952.
- [Dürr 2007] C. Dürr et N. K. Thang. *Nash Equilibria in Voronoi Games on Graphs*. Algorithms–ESA, pages 17–28, 2007.
- [Dutta 1995] P. K. Dutta. *A folk Theorem for Stochastic games*. Journal of Economic Theory, vol. 66, pages 1–32, 1995.
- [Fekete 2005] S. P. Fekete, J. S. B. Mitchell et K. Beurer. *On the Continuous Fermat-Weber Problem*. Operations Research, vol. 53, no. 1, pages 61–76, 2005.
- [Feng 2007] T. Feng, T. R. Field et S. Haykin. *Stochastic Differential Equation Theory Applied to Wireless Channels*. IEEE Transactions on Communications, vol. 55, no. 8, pages 1478–1483, 2007.
- [Fette 2009] B. A. Fette. *Cognitive Radio Technology*. Elsevier, 2009.
- [Fleming 2006] W. H. Fleming et H. M. Soner. *Controlled Markov Processes and Viscosity Solutions*, volume 25. Springer New York, 2006.
- [Foschini 1993] G. J. Foschini et Z. Miljanic. *A Simple Distributed Autonomous Power Control Algorithm and its Convergence*. IEEE Transactions on Vehicular Technology, vol. 42, no. 4, pages 641–646, 1993.
- [Fudenberg 1991] D. Fudenberg et J. Tirole. *Game Theory*. MIT Press, Cambridge, MA, 1991.
- [Fudenberg 2011] D. Fudenberg et Y. Yamamoto. *The Folk Theorem for Irreducible Stochastic Games with Imperfect Public Monitoring*. Journal of Economic Theory, vol. 146, no. 4, pages 1664–1683, 2011.
- [Gabszewicz 1992] J. J. Gabszewicz et J. F. Thisse. *Location*. Handbook of game theory with economic applications, vol. 1, pages 281–304, 1992.
- [Goodman 2000] D. J. Goodman et N. B. Mandayam. *Power Control for Wireless Data*. IEEE Personal Communications, vol. 7, no. 2, pages 48–54, 2000.

- [Gupta 1997] P. Gupta et P. R. Kumar. *A System and Traffic Dependent Adaptive Routing Algorithm for Ad Hoc Networks*. In IEEE Conference on Decision and Control (CDC), volume 3, pages 2375–2380, 1997.
- [Han 2011] Z. Han, D. Niyato, W. Saad, T. Basar et A. Hjørungnes. *Game Theory in Wireless and Communication Networks : Theory, Models and Applications*. Cambridge University Press, Cambridge, UK, 2011.
- [Haykin 2005] S. Haykin. *Cognitive Radio : Brain-Empowered Wireless Communications*. IEEE Journal on Selected Areas in Communications, vol. 23, no. 2, pages 201–220, 2005.
- [Hörner 2011] J. Hörner, T. Sugaya, S. Takahashi et N. Vieille. *Recursive Methods in Discounted Stochastic Games : An Algorithm for $\delta \hat{=} 1$ and a Folk Theorem*. Econometrica, vol. 79, no. 4, pages 1277–1318, 2011.
- [Hotelling 1929] H. Hotelling. *Stability in Competition*. The Economic Journal, vol. 39, no. 153, pages 41–57, 1929.
- [Huang 2003] M. Huang, P. E. Caines et R. P. Malhamé. *Individual and Mass Behaviour in Large Population Stochastic Wireless Power Control Problems : Centralized and Nash Equilibrium Solutions*. In IEEE Conference on Decision and Control (CDC), volume 1, pages 98–103, 2003.
- [Karatzas 1991] I. Karatzas et S. E. Shreve. *Brownian Motion and Stochastic Calculus*, volume 113. Springer, 1991.
- [Lasaulce 2009] S. Lasaulce, Y. Hayel, R. El-Azouzi et M. Debbah. *Introducing Hierarchy in Energy Games*. IEEE Transactions on Wireless Communications, vol. 8, no. 7, pages 3833–3843, 2009.
- [Lasaulce 2011] S. Lasaulce et H. Tembine. *Game Theory and Learning for Wireless Networks : Fundamentals and Applications*. Academic Press, 2011.
- [Lasry 2007] J. M. Lasry et P. L. Lions. *Mean Field Games*. Japanese Journal of Mathematics, vol. 2, no. 1, pages 229–260, 2007.
- [Le Treust 2010] M. Le Treust et S. Lasaulce. *A Repeated Game Formulation of Energy-Efficient Decentralized Power Control*. IEEE Transactions on Wireless Communications, vol. 9, no. 9, pages 2860–2869, 2010.
- [Litva 1996] J. Litva et T. K. Lo. *Digital Beamforming in Wireless Communications*. Artech House, Inc., Norwood, MA, USA, 1996.
- [Mériaux 2011a] F. Mériaux, Y. Hayel, S. Lasaulce et A. Garnaev. *Long-Term Energy Constraints and Power Control in Cognitive Radio Networks*. In IEEE International Conference on Digital Signal Processing (DSP), pages 842–848, 2011.
- [Mériaux 2011b] F. Mériaux, S. Lasaulce et M. Kieffer. *More about Base Station Location Games*. In ICST Conference on Performance Evaluation Methodologies and Tools (VALUETOOLS), pages 372–380, 2011.
- [Mériaux 2011c] F. Mériaux, M. Le Treust, S. Lasaulce et M. Kieffer. *A Stochastic Game Formulation of Energy-Efficient Power Control : Equilibrium Utilities*

- and Practical Strategies*. In IEEE International Conference on Digital Signal Processing (DSP), pages 665–671, 2011.
- [Mériaux 2012a] F. Mériaux, S.M. Perlaza, S. Lasaulce, Z. Han et V. Poor. *Achievability of Efficient Satisfaction Equilibria in Self-Configuring Networks*. In Game Theory for Networks, volume 105 of *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, pages 1–15. Springer Berlin Heidelberg, 2012.
- [Mériaux 2012b] F. Mériaux, S. Valentin et S. Lasaulce. *An Energy-Efficient Power Allocation Game with Selfish Channel State Feedback in Cellular Networks*. In IEEE Conference on Performance Evaluation Methodologies and Tools (VALUETOOLS), pages 290–297, 2012.
- [Mériaux 2012c] F. Mériaux, V. Varma et S. Lasaulce. *Mean Field Energy Games in Wireless Networks*. In IEEE Conference on Signals, Systems and Computers (ASIOMAR), pages 671–675, 2012.
- [Mériaux 2013] F. Mériaux, S. Lasaulce et H. Tembine. *Stochastic Differential Games and Energy-Efficient Power Control*. Dynamic Games and Applications, vol. 3, no. 1, pages 3–23, 2013.
- [Meshkati 2005] F. Meshkati, H. V. Poor, S. C. Schwartz et B. M. Narayan. *An Energy-Efficient Approach to Power Control and Receiver Design in Wireless Data Networks*. IEEE Transactions on Communications, vol. 53, pages 1885–1894, 2005.
- [Meshkati 2006] F. Meshkati, M. Chiang, H. V. Poor et S. C. Schwartz. *A Game-Theoretic Approach to Energy-Efficient Power Control in Multi-Carrier CDMA Systems*. IEEE Journal on Selected Areas in Communications, vol. 24, no. 6, pages 1115–1129, 2006.
- [Mitola 1999] J. Mitola et G. Q. Maguire. *Cognitive Radio : Making Software Radios more Personal*. IEEE Personal Communications, vol. 6, no. 4, pages 13–18, 1999.
- [Narendra 1989] K. S. Narendra et M. A. Thathachar. Learning Automata : an Introduction. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1989.
- [Nash 1950] J. F. Nash. *Equilibrium Points in n -Person Games*. Proceedings of the national academy of sciences, vol. 36, no. 1, pages 48–49, 1950.
- [Olama 2006] M. Olama, S. Djouadi et C. Charalambous. *Stochastic Power Control for Time-Varying Long-Term Fading Wireless Networks*. EURASIP Journal on Applied Signal Processing (JASP), vol. Vol. 2006, pages 1–13, 2006.
- [Perlaza 2012] S. M. Perlaza, H. Tembine, S. Lasaulce et M. Debbah. *Quality-Of-Service Provisioning in Decentralized Networks : A Satisfaction Equilibrium Approach*. IEEE Journal of Selected Topics in Signal Processing, vol. 6, no. 2, pages 104–116, 2012.
- [Plastria 2001] F. Plastria. *Static Competitive Facility Location : An Overview of Optimisation approaches*. European Journal of Operational Research, vol. 129, no. 3, pages 461–470, 2001.

- [Rodriguez 2003] V. Rodriguez. *An Analytical Foundation for Resource Management in Wireless Communication*. In IEEE Global Telecommunications Conference (Globecom), pages 898–902, 2003.
- [Rosen 1965] J. B. Rosen. *Existence and Uniqueness of Equilibrium Points for Concave N-Person Games*. *Econometrica*, vol. 33, no. 3, pages 520–534, 1965.
- [Saraydar 2002] C. U. Saraydar, N. B. Mandayam et D. J. Goodman. *Efficient Power Control via Pricing in Wireless Data Networks*. *IEEE Transactions on Communications*, vol. 50, no. 2, pages 291–303, 2002.
- [Sastry 1994] P. S. Sastry, V. V. Phansalkar et M. A. L. Thathachar. *Decentralized Learning of Nash Equilibria in Multi-Person Stochastic Games with Incomplete Information*. *IEEE Transactions on Systems, Man and Cybernetics*, vol. 24, no. 5, pages 769–777, 1994.
- [Scutari 2006] G. Scutari, S. Barbarossa et D.P. Palomar. *Potential Games : A Framework for Vector Power Control Problems With Coupled Constraints*. In the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), volume 4, pages IV–IV, 2006.
- [Shapley 1953] L. Shapley. *Stochastic Games*. *Proceedings of the National Academy of Sciences of the United States of America*, vol. 39, no. 10, pages 1095–1100, 1953.
- [Silva 2010] A. Silva, H. Tembine, E. Altman et M. Debbah. *Spatial Games and Global Optimization for the Mobile Association Problem : the Downlink Case*. In IEEE Conference on Decision and Control (CDC), pages 966–972, 2010.
- [Sorin 1992] S. Sorin. *Repeated Games with Complete Information*. *Handbook of Game Theory with Economic Applications*, vol. 1, pages 71–107, 1992.
- [Stackelberg 1934] H. Stackelberg. *Marktform und Gleichgewicht*. Springer, Wien, 1934.
- [Sutton 1998] R. S. Sutton et A. G. Barto. *Reinforcement Learning : An Introduction*. Cambridge Univ Press, 1998.
- [Tembine 2010] H. Tembine, S. Lasaulce et M. Jungers. *Joint Power Control-allocation for Green Cognitive Wireless Networks using Mean Field Theory*. In IEEE International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM), pages 1–5, 2010.
- [Tembine 2011] H. Tembine et M. Huang. *Mean Field Difference games : McKean-Vlasov Dynamics*. In IEEE Conference on Decision and Control and European Control Conference (CDC-ECC), pages 1006–1011, 2011.
- [Tulino 2004] A. Tulino et S. Verdú. *Random Matrix Theory and Wireless Communications*, volume 1. Now Publishers Inc, 2004.
- [Van Veen 1988] B. D. Van Veen et K. M. Buckley. *Beamforming : A Versatile Approach to Spatial Filtering*. *IEEE Signal Processing Magazine*, vol. 5, no. 2, pages 4–24, 1988.

- [Wu 2001] C. C. Wu et D. P. Bertsekas. *Distributed Power Control Algorithms for Wireless Networks*. IEEE Transactions on Vehicular Technology, vol. 50, no. 2, pages 504–514, 2001.
- [Wyner 1974] A. D. Wyner. *Recent Results in Shannon Theory*. IEEE Transactions on Information Theory, vol. 20, no. 1, pages 2–10, 1974.
- [Yates 1995] R. D. Yates. *A Framework for Uplink Power Control in Cellular Radio Systems*. IEEE Journal on Selected Areas in Communications, vol. 13, no. 7, pages 1341–1347, 1995.