



**HAL**  
open science

# Méthodes décentralisées d'allocation des ressources dans le canal d'interférence acoustique sous-marin

Antony Pottier

► **To cite this version:**

Antony Pottier. Méthodes décentralisées d'allocation des ressources dans le canal d'interférence acoustique sous-marin. Réseaux et télécommunications [cs.NI]. Ecole nationale supérieure Mines-Télécom Atlantique, 2018. Français. NNT : 2018IMTA0095 . tel-02095995

**HAL Id: tel-02095995**

**<https://theses.hal.science/tel-02095995>**

Submitted on 11 Apr 2019

**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.

# THÈSE DE DOCTORAT DE

L'ÉCOLE NATIONALE SUPÉRIEURE MINES-TÉLÉCOM  
ATLANTIQUE BRETAGNE PAYS DE LA LOIRE - IMT ATLANTIQUE

COMUE UNIVERSITÉ BRETAGNE LOIRE

ÉCOLE DOCTORALE N° 601

*Mathématiques et Sciences et Technologies  
de l'Information et de la Communication*

Spécialité : Télécommunications

Par

**Antony POTTIER**

**Méthodes décentralisées d'allocations des ressources  
dans le canal d'interférence acoustique sous-marin**

Thèse présentée et soutenue à IMT Atlantique, le 15 Novembre 2018

Unité de recherche : UMR CNRS 6285 LabSTICC

Thèse N° : 2018IMTA095

## Rapporteurs avant soutenance :

Jean-Marie Gorce  
Benoit Geller

Professeur, INSA Lyon  
Professeur, ENSTA PartiTech

## Composition du Jury :

Président : Emmanuel Radoi  
Rapporteurs : Jean-Marie Gorce  
Benoit Geller  
Examineurs : Veronica Belmega  
François-Xavier Socheleau  
Dir. de thèse : Christophe Laot

Professeur, Université de Bretagne Occidentale  
Professeur, INSA Lyon  
Professeur, ENSTA PartiTech  
Maître de conférences, ENSEA  
Maître de conférences, IMT Atlantique  
Professeur, IMT Atlantique



Sous le sceau de l'Université Bretagne Loire

# IMT Atlantique Bretagne-Pays de la Loire

En accréditation conjointe avec l'Ecole Doctorale MathSTIC

---

## Méthodes d'allocation des ressources dans le canal d'interférence acoustique sous-marin

---

### Thèse de Doctorat

Spécialité : Télécommunications

Présentée par **Antony Pottier**

Département : Signal et Communications

Laboratoire : UMR CNRS 6285 Lab-STICC

Pôle : CACS

Directeur de thèse : Christophe LAOT

Soutenue le 15 novembre 2018

#### Jury :

M. Benoit Geller, Professeur, ENSTA Paristech (Rapporteur)  
M. Jean-Marie Gorce, Professeur, INSA Lyon (Rapporteur)  
Mme Elena Veronica Belmega, Maître de conférences, ENSEA (Examineur)  
M. Emanuel Radoï, Professeur, Université de Bretagne Occidentale (Examineur)  
M. Christophe Laot, Professeur, IMT Atlantique (Examineur)  
M. François-Xavier Socheleau, Maître de conférences, IMT Atlantique (Examineur)



Cette thèse a été préparée à IMT Atlantique au sein du département Signal et Communications (UMR CNRS 6285 Lab-STICC). Cette thèse a été co-financée par la région Bretagne sur une bourse ARED. Une partie des travaux présentés dans ce manuscrit est issue d'une collaboration avec le Dr P. Mitchell de Université de York au Royaume-Uni, effectuée lors d'un séjour d'études de 3 mois financé par l'Université de Bretagne-Loire, l'IMT Atlantique et le LabSTICC. Ce travail a également été soutenu par DGA-TN qui a fourni une partie des données réelles exploitées dans ce manuscrit.



# Remerciements

J'aimerais tout d'abord témoigner ma profonde gratitude et ma sympathie envers Christophe Laot et François-Xavier Socheleau avec qui j'ai eu le plaisir de travailler durant mon passage à IMT Atlantique, et qui m'ont proposé ce sujet de thèse. Je les remercie pour la qualité de leur encadrement, leur rigueur scientifique et leurs conseils toujours pertinents.

Je remercie vivement les membres du jury pour l'intérêt qu'ils ont porté à mes recherches. Je remercie Emmanuel Radoï, qui m'a fait l'honneur de présider le jury. Merci à Jean-Marie Gorce et Benoit Geller d'avoir accepté de rapporter cette thèse et pour l'attention et le soin qu'ils ont apporté à la lecture du manuscrit. Merci également à Veronica Belmega pour sa participation au jury et l'intérêt qu'elle a porté à ces travaux.

Merci à Paul D. Mitchell de m'avoir accueilli dans son équipe du département Electronic Engineering à l'Université de York.

Merci enfin aux membres du département Signal et Communications avec qui j'ai eu le plaisir de passer ces quelques années. Je leur témoigne ici ma plus sincère amitié.



# Résumé

L'onde acoustique est utilisée par de nombreux systèmes et organismes sous-marins pour communiquer, naviguer ou inférer des informations à propos de l'environnement. Le développement des activités humaines liées au monde de la mer (océanographie, lutte sous-marine, pêche, prospection pétrolière, etc...) induit une augmentation du nombre de sources acoustiques en activité simultanée dans l'océan. L'environnement acoustique sous-marin (ASM) est alors partagé par de nombreuses sources hétérogènes (sonars, mammifères marins, systèmes de communications et de télémétrie, etc...) entrant involontairement en compétition pour l'utilisation de la ressource offerte par le canal de transmission. Néanmoins, et contrairement aux communications radiofréquences, l'activité temporelle et fréquentielle des systèmes n'est pas normalisée par des standards internationaux. Ils sont en revanche contraints par la physique du milieu sous-marin, qui leur impose des bandes passantes souvent très faibles, dépendante de la distance et de la fréquence de transmission, et des délais de propagation longs. En l'absence de régulation, l'accès anarchique au canal par plusieurs sources peut induire un masquage mutuel des signaux acoustiques de sorte qu'il soit impossible d'établir un lien de communication fiable.

L'objectif de ce travail de thèse est de proposer des solutions permettant aux systèmes de communications ASM d'adapter leurs paramètres de transmission de manière intelligente, autonome et décentralisée selon l'environnement acoustique dans lequel ils évoluent. À plusieurs égards, les problématiques de ce sujet de thèse sont proches de celles ayant motivé le développement des recherches sur la radio cognitive dans le monde des communications terrestres. Les systèmes de communications doivent en effet être capables, étant donné l'état du canal de transmission, d'identifier et de mettre en œuvre la meilleure stratégie de transmission au sens d'un certain critère : maximisation du débit, du rapport signal à bruit ou de la capacité du lien, minimisation de la probabilité d'erreur ou de la puissance transmise sous contrainte de débit, etc. Cependant, les spécificités du milieu acoustique sous-marin, les sources d'interférences de différentes natures et l'absence de standards de communications posent de nouvelles difficultés.

## Mots clés

Communications acoustiques sous-marines, théorie des jeux, apprentissage multi-agents.



# Abstract

Underwater acoustic waves are used by many systems and biologic organisms to communicate, navigate or infer information about the environment. Future developments of human activities underwater for oceanography, warfare, fishing or other applications imply an increase of the number of active acoustic sources in the oceans. The underwater environment is therefore shared by many heterogeneous sources (sonars, marine mammals, telemetry and communication systems, ...) which involuntarily compete for using the physical resource offered by the channel. Nevertheless, and unlike radiofrequency communications, the temporal and frequential activity of underwater systems is not normalized by international standards. They are however constrained by physical properties of the underwater environment, which imposes low bandwidths - which also depend on distance and transmission frequency - and long propagation delays. Without regulation, several sources accessing anarchically to the channel may provoke mutual interferences which are detrimental to the robustness of the communication links.

The goal of this thesis is to provide solutions allowing autonomous and decentralized and adaptation for the transmission strategies of underwater acoustic communication systems, according to the environment in which they operate. To some extent, this thesis deals with topics that are close to what has motivated the development of research on cognitive radio. Indeed, given the state of the channel, communication systems should be able to identify and implement the best transmission strategy regarding some performance criterion : capacity, signal to noise ratio, error rate, transmission power under rate constraints, ... However, the specific properties of the underwater environment, the heterogeneity of interfering acoustic sources and the absence of standardized communication protocols rise new difficulties.

## **Keywords**

Underwater acoustic communications, game theory, multi-agent learning



# Table des matières

<b>Remerciements</b>	<b>3</b>
<b>Résumé</b>	<b>5</b>
<b>Abstract</b>	<b>7</b>
<b>Table des matières</b>	<b>11</b>
<b>Acronymes</b>	<b>13</b>
<b>Notations</b>	<b>15</b>
<b>Introduction générale</b>	<b>17</b>
<b>1 Communications dans le milieu de propagation sous-marin</b>	<b>23</b>
1.1 Propriétés physiques du milieu de propagation . . . . .	24
1.1.1 Célérité de l’onde acoustique . . . . .	24
1.1.2 Trajets multiples . . . . .	25
1.1.3 Pertes de propagation . . . . .	26
1.1.4 Sources et modèles de bruit . . . . .	27
1.1.5 Hétérogénéités et variabilité de l’environnement . . . . .	29
1.2 Impact de la physique du milieu sur les signaux de communication . . .	30
1.2.1 Fréquence optimale et bande passante . . . . .	30
1.2.2 Effets Doppler . . . . .	31
1.2.3 Réponse impulsionnelle de canal doublement dispersif . . . . .	32
1.2.4 Temps et bande de cohérence . . . . .	34
1.3 Le canal acoustique sous-marin comme canal d’interférence . . . . .	35
1.4 Adaptation au canal et gestion de l’interférence . . . . .	37
1.4.1 Communications ASM point-à-point et adaptatives . . . . .	37
1.4.2 Communications multi-utilisateurs et interférence . . . . .	39
1.4.3 Réseaux sous marins . . . . .	40
1.5 Conclusion . . . . .	41
<b>2 Notions de théorie des jeux</b>	<b>43</b>
2.1 Jeux sous forme stratégique . . . . .	44
2.1.1 Définition et exemples introductifs . . . . .	44
2.1.2 Stratégies pures et randomisées, utilité espérée . . . . .	47

2.1.3	Jouer de manière rationnelle . . . . .	48
2.1.4	Équilibre de Nash . . . . .	49
2.1.5	Équilibre corrélé . . . . .	53
2.1.6	Pareto-optimalité . . . . .	54
2.1.7	Apprentissage distribué des équilibres de Nash . . . . .	55
2.2	Jeux de satisfaction . . . . .	56
2.2.1	Jeu et équilibre de satisfaction . . . . .	57
2.2.2	Apprentissage décentralisé des équilibres de satisfaction . . . . .	60
2.3	Conclusion . . . . .	61
<b>3</b>	<b>Partage de spectre robuste et décentralisé</b>	<b>63</b>
3.1	Formulation du problème . . . . .	64
3.1.1	Modèle de la transmission et hypothèses . . . . .	64
3.1.2	Jeu de maximisation du débit moyen . . . . .	66
3.1.3	Connaissance incomplète de la fonction d'utilité . . . . .	67
3.2	Solution de water-filling itératif . . . . .	69
3.2.1	Meilleure réponse et équilibre de Nash . . . . .	69
3.2.2	Apprentissage décentralisé d'un équilibre . . . . .	70
3.2.3	Unicité de l'équilibre et convergence de l'algorithme d'apprentissage	71
3.3	Incertitude quant aux statistiques du canal . . . . .	73
3.4	Études et résultats sur canaux synthétiques et sondés en mer . . . . .	75
3.4.1	Méthode d'évaluation . . . . .	75
3.4.2	Étude d'un jeu à 2 joueurs . . . . .	77
3.4.3	Étude générale . . . . .	81
3.4.4	Résultats sur canaux sondés en mer . . . . .	84
3.4.5	Aspects pratiques . . . . .	87
3.5	Conclusion . . . . .	88
<b>4</b>	<b>Partage décentralisé du canal sous contraintes de qualité de service</b>	<b>91</b>
4.1	Allocation de puissance sous contrainte d'utilité . . . . .	92
4.1.1	Formulation du problème . . . . .	92
4.1.2	Équilibre de Nash et water-filling itératif . . . . .	93
4.1.3	Résultats sur des canaux synthétiques . . . . .	95
4.2	Approche par les jeux de satisfaction . . . . .	98
4.2.1	Formulation du problème . . . . .	98
4.2.2	Algorithme d'apprentissage aveugle . . . . .	99
4.2.3	Modèle des transmissions . . . . .	102
4.2.4	Étude de cas : Communications DSSS . . . . .	103
4.2.5	Étude de cas : Communications OFDM . . . . .	106
4.2.6	Résultats sur des canaux synthétiques . . . . .	109
4.3	Conclusion . . . . .	114
<b>5</b>	<b>Réseaux de capteurs sous marins décentralisés et adaptatifs</b>	<b>117</b>
5.1	Modèle du réseau . . . . .	118
5.2	Sélection d'un point d'accès avec $Q$ -learning . . . . .	121
5.2.1	Apprentissage dans un processus de décision de Markov . . . . .	121

5.2.2	Algorithme $Q$ -learning . . . . .	123
5.2.3	Application à la sélection de point d'accès . . . . .	125
5.3	Résultats numériques . . . . .	126
5.4	Conclusion . . . . .	130
<b>Conclusion générale et perspectives</b>		<b>131</b>
<b>Liste des publications</b>		<b>135</b>
<b>A Calcul des fonctions d'utilité pour le partage de spectre robuste</b>		<b>137</b>
A.1	Minimisation contre l'interférence la plus préjudiciable . . . . .	137
A.2	Canaux à évanouissements de Rice . . . . .	138
<b>B Unicité de l'équilibre de Nash du partage de spectre robuste</b>		<b>141</b>
<b>C Convergence de l'algorithme de water-filling itératif</b>		<b>143</b>
C.1	Convergence de la dynamique de meilleures réponse asynchrone . . . . .	143
C.2	Convergence du water-filling itératif . . . . .	145
<b>D Cas particuliers de convergence des algorithmes 1 et 3</b>		<b>147</b>
D.1	Convergence pour les jeux de satisfaction en DSSS . . . . .	147
D.2	Convergence pour les jeux de satisfaction en OFDM . . . . .	148
D.2.1	Water-filling et retour de l'état du canal . . . . .	148
D.2.2	Allocation ON/OFF et retour sur 1 bit . . . . .	149
<b>Bibliographie</b>		<b>160</b>



# Acronymes

<b>ASM</b>	Acoustique Sous-Marin(e)
<b>AUV</b>	Autonomous Underwater Vehicle
<b>CSI</b>	Channel State Information
<b>CDMA</b>	Code Division Multiple Acces
<b>CSMA</b>	Carrier Sense Multiple Access
<b>DSSS</b>	Direct Sequence Spread Spectrum
<b>d.d.p</b>	densité de probabilité
<b>FDMA</b>	Frequency Division Multiple Acces
<b>FHSS</b>	Frequency Hopping Spread Spectrum
<b>FSK</b>	Frequency Shift Keying
<b>IES</b>	Interférence entre symboles
<b>IESP</b>	Interférence entre sous-porteuses
<b>i.i.d</b>	indépendant et identiquement distribué
<b>MIMO</b>	Multiple Input Multiple Output
<b>MMSE</b>	Minimum Mean Square Error
<b>NATO</b>	North Atlantic Treaty Organization
<b>OFDM</b>	Orthogonal Frequency Division Multiplexing
<b>OTAN</b>	Organisation du Traité de l'Atlantique Nord
<b>PAPR</b>	Peak-to-Average Power Ratio
<b>PSK</b>	Phase Shift Keying
<b>QAM</b>	Quadrature Amplitude Modulation
<b>RF</b>	Radio-fréquences
<b>RMS</b>	Root Mean Square
<b>RSB</b>	Rapport Signal sur Bruit
<b>RSIB</b>	Rapport Signal sur Interférence plus Bruit
<b>RX</b>	Récepteur
<b>SINR</b>	Signal-to-Interference plus Noise Ratio
<b>SNR</b>	Signal-to-Noise Ratio
<b>TDMA</b>	Time Division Multiple Access
<b>TEB</b>	Taux d'Erreur Binaire
<b>TX</b>	Emetteur
<b>WSSUS</b>	Wide-Sense Stationary and Uncorrelated Scattering



# Notations

$x$	scalaire, élément d'un ensemble
$\mathbf{x}$	vecteur
$\mathbf{X}$	matrice
$\mathcal{X}$	ensemble

## Ensembles

$\mathbb{C}$	ensemble des nombres complexes
$\mathbb{R}$	ensemble des nombres réels
$\mathbb{Z}$	ensemble des entiers relatifs
$\mathbb{N}$	ensemble des entiers naturels
$\Delta\mathcal{X}$	ensemble des distributions de probabilités sur $\mathcal{X}$
$\mathcal{M}_{n \times p}(\mathbb{K})$	ensemble des matrices à $n$ lignes et $p$ colonnes dont les coefficients sont à valeurs dans $\mathbb{K}$
$\mathcal{L}^2(\mathbb{R})$	espace de Hilbert des fonctions de carré intégrable

## Symboles et opérateurs

$\mathbf{I}_n$	matrice identité $n \times n$
$\mathbf{1}_n$	vecteur colonne dont les $n$ éléments sont égaux à 1
$\mathbb{1}_{x \in \mathcal{X}}$	fonction indicatrice valant 1 $\forall x \in \mathcal{X}$ et 0 sinon.
$\mathbf{D}(\mathbf{x})$	matrice diagonale contenant les éléments du vecteur $\mathbf{x}$
$ \mathcal{X} $	cardinal de l'ensemble $\mathcal{X}$
$(\cdot)^*$	conjugué
$\mathbf{X}^T$	transposée
$\mathbf{X}^H$	transposée et conjuguée
$\odot$	produit de Hadamard
$[\cdot]_{k,l}$	élément de la $k$ -ième ligne et $l$ -ième colonne d'une matrice
$\ \mathbf{X}\ _F$	norme de Frobenius, <i>i.e.</i> , $\sqrt{\text{Tr}[\mathbf{X}^H \mathbf{X}]}$
$\ \mathbf{x}\ _2$	norme euclidienne du vecteur $\mathbf{x}$ , soit, $\sqrt{\sum_i x_i^2}$
$\ \mathbf{x}\ _1$	norme $\ell_1$ du vecteur $\mathbf{x}$ , soit, $\sum_i  x_i $
$\ \cdot\ _{\ell_0}$	norme $\ell_0$ , <i>i.e.</i> , nombre d'éléments non-nuls de $\mathbf{x}$

$\text{Tr}(\cdot)$	trace d'une matrice carrée
$\Re$	partie réelle
$\Im$	partie imaginaire
$ \cdot $	module
$\langle \cdot, \cdot \rangle$	produit scalaire
$\delta(\cdot)$	distribution de Dirac
$\delta_{ij}$	symbole de Kroenecker
$\nabla$	opérateur nabla

### Probabilités, variables aléatoires

$(\Omega, \mathcal{F}, \mathbb{P})$	espace probabilisé où $\Omega$ est un ensemble non vide, $\mathcal{F}$ est une tribu sur $\Omega$ et $\mathbb{P}$ est la mesure de probabilité sur $\mathcal{F}$ .
$X$	variable aléatoire dont les réalisations sont notées $X(\omega)$ pour $\omega \in \Omega$
$\mathbb{E}[X]$	espérance mathématique <i>i.e.</i> $\int X d\mathbb{P}$
$\mathbb{E}_X[\phi(X, Y)]$	espérance de $\phi(X, Y)$ par rapport à la loi de $X$ , <i>i.e.</i> , $\int \phi(X, Y) d\mathbb{P}_X$
$\mathbb{E}[X   Y]$	espérance de $X$ conditionnellement à $Y$
$\mathcal{N}(\mu, \sigma^2)$	loi normale (ou gaussienne) de moyenne $\mu$ et de variance $\sigma^2$
$\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	loi normale (ou gaussienne) complexe multivariée
$\mathcal{U}(\mathcal{I})$	loi uniforme sur l'ensemble $\mathcal{I}$
$\mathcal{E}(\lambda)$	loi exponentielle de paramètre $\lambda$
$\mathcal{P}(\lambda)$	loi de Poisson de paramètre $\lambda$
$\mathcal{B}(p)$	loi de Bernouilli de paramètre $p$
$X \sim \mathcal{L}$	variable aléatoire $X$ distribuée selon la loi (ou mesure) $\mathcal{L}$

### Théorie des Jeux

$\mathcal{I}$	ensemble de joueurs numérotés de 1 à $I$
$\mathcal{A}_i$	espace des actions disponibles au joueur $i$
$\mathcal{A}$	ensemble produit des espaces d'actions des joueurs
$\mathcal{A}_{-i}$	ensemble produit des espaces d'actions de tous les joueurs excepté $i$ , <i>i.e.</i> , $\mathcal{A} \setminus \mathcal{A}_i$
$a_i$	une action $a_i \in \mathcal{A}_i$ du joueur $i$
$a_{-i}$	un profil d'action de tous les joueurs autres que $i$ , $a_{-i} \in \mathcal{A}_{-i}$
$a$	un profil d'action, $a \in \mathcal{A}$
$\pi_i$	une stratégie mixte du joueur $i$ , $\pi_i \in \Delta \mathcal{A}_i$
$u_i$	fonction d'utilité du joueur $i$ dans le jeu sous forme normale
$u_i(a)$	utilité du joueur $i$ pour le profil d'actions $a$
$u_i(\pi)$	utilité espérée du joueur $i$ pour le profil de stratégies mixtes $\pi$

# Introduction générale

## Motivations et contexte de l'étude

L'onde acoustique se révèle être un support privilégié pour la transmission et l'inférence d'informations dans le milieu marin, là où les ondes électromagnétiques souffrent de fortes atténuations. De nombreuses questions économiques et environnementales motivent le développement des techniques de l'acoustique sous-marine (ASM) pour l'analyse, l'exploration et l'exploitation de l'environnement marin. Leur utilisation est indispensable à de nombreuses applications liées au monde de la mer où sonars, modems et écho-sondeurs sont des exemples de systèmes acoustiques employés à des fins militaires, scientifiques, ou commerciales (océanologie, guerre sous-marine, prospection pétrolière, etc...) [Lurton, 1998]. Les observatoires de fond de mer, les drones et robots sous-marins télé-opérés (AUVs et ROVs) ou les submersibles habitables sont d'autres exemples où la communication d'informations à des distances plus ou moins longues est nécessaire. La capacité à établir des liens de communication ASM fiables et efficaces apparaît alors essentielle.

L'environnement sous-marin est intrinsèquement bruyant ; pluie, vent, vagues et vocalises des mammifères marins étant des exemples de contributions naturelles au bruit ambiant. Les développements actuels et futurs des activités humaines sous-marines présagent une augmentation du nombre de sources acoustiques en activité simultanée dans les mers et les océans. Cela implique inévitablement une augmentation du bruit d'origine anthropique, dont les sources diverses incluent sonars, modems et sondeurs, ainsi que les bruits de bateaux, de forages, etc. Cependant, les questions de la pollution sonore des océans par l'activité humaine et de sa régulation n'ont été que peu abordées par les organisations politiques [McCarthy, 2004]. L'augmentation des niveaux de bruits a des conséquences néfastes à plusieurs niveaux. Son effet sur la faune marine est notable, la survie des mammifères marins, par exemple, reposant sur les ondes acoustiques en tant que moyen de leurs interactions sociales et d'écholocation [Tyack, 1986]. Les performances des systèmes sous-marins sont également drastiquement réduites par le bruit et les sources d'interférences acoustiques externes. Ainsi, le développement des activités humaines sous l'océan se heurte à une contradiction opposant l'augmentation du nombre de systèmes ASM à la difficulté de faire cohabiter les sources acoustiques hétérogènes peuplant l'océan en l'absence de réglementation.

En particulier, et au contraire des télécommunications terrestres, l'activité temporelle et fréquentielle des modems ASM n'est que peu réglementée par des normes et

des standards. Il n'existe pas non plus de structure de réseaux pré-établie sur laquelle la cohabitation et le partage du canal de communication entre des systèmes hétérogènes pourrait reposer. Ceux-ci sont toutefois contraints par la physique du milieu. L'atténuation croissante avec la fréquence et la distance impose des fréquences porteuses et des bandes passantes souvent faibles, de l'ordre du kilohertz à la dizaine de kilohertz, ce qui limite fortement les débits atteignables. De plus, la faible célérité de l'onde acoustique sous l'eau (1500  $m/s$ ) induit des délais de propagation longs de plusieurs secondes, voire dizaines de secondes. La ressource offerte par le milieu marin en tant que canal de communication est donc limitée et sa disponibilité dépend de plusieurs facteurs, tels que

- le nombre et la nature des sources actives simultanément,
- leur situation géographique, ou leur position relative,
- la bande de fréquence des transmissions,
- l'activité temporelle des sources (temps, durée, répétition).

Dès lors, l'accès anarchique au canal ASM peut induire un masquage mutuel de signaux de sorte qu'il est impossible d'établir des liens de communication fiables [McGee et al., 2014; Wang et al., 2012]. Cette compétition pour l'accès aux ressources du canal peut se formaliser, en théorie de l'information, par le modèle de canal d'interférence.

Le problème de l'interférence et les questions de « l'intérêt et la faisabilité de standardiser, allouer ou contrôler les fréquences acoustiques pour les télécommunications sous-marines » [NAS, 1970] se sont posées au début des années 1970. Le spectre utilisable pour les communications ASM étant relativement étroit, il est clairement irréaliste d'envisager une régulation fixe du spectre, où des bandes de fréquences prédéfinies seraient allouées en fonction des applications. Jusqu'à très récemment il n'existait pas de standard de communication en dehors du STANAG 1074 de l'OTAN concernant les téléphones sous-marins analogiques, et dont la première version fut publiée en 1985. La standardisation par l'OTAN du protocole de communications ASM numériques JANUS [NATO, 2017] en Avril 2017 est une avancée notable en ce qu'elle offre un cadre permettant l'interopérabilité et la découverte de systèmes hétérogènes dans le but de former des réseaux ad-hoc [Potter et al., 2014]. Néanmoins, l'adoption d'un tel protocole ne résout pas le problème du partage efficace des ressources physiques limitées offertes par le canal de communication ASM.

L'ambition de ce travail de thèse est de proposer des solutions pragmatiques d'allocation des ressources entre systèmes de communications ASM en conflit pour l'accès au canal. L'hétérogénéité des systèmes et leur inter-opérabilité limitée, voire inexistante, imposent un cadre strictement non-coopératif pour résoudre le problème d'allocation des ressources. Par « *non-coopératifs* », on exprime l'idée que les différents systèmes simultanément en activité dans le canal ASM n'ont pas la possibilité d'échanger des messages afin de se mettre d'accord sur un schéma équitable de partage de la ressource. Cela peut même s'étendre à l'idée qu'ils ne sont pas conscients du fait qu'ils entrent en compétition pour l'utilisation de cette ressource. En l'absence d'infrastructures de réseau standardisées il est également important que l'allocation se fasse de manière décentralisée. Du fait de la décentralisation et de la non-coopération imposées par le contexte, les utilisateurs du canal ne se voient pas « allouer » des ressources au sens strict, dans la mesure où cette allocation ne résulte pas d'une autorité extérieure, mais font plutôt face à un problème de *décision* à prendre quant à l'utilisation

d'une ressource particulière. Cette décision correspond à une *stratégie de transmission*. Concrètement, cela peut se traduire par le choix de la puissance transmise, de la (ou des) bande(s) de fréquences occupée(s), de l'ordre de modulation, etc... Dans ce cadre de travail, les outils de la théorie des jeux et de l'apprentissage automatique vont se révéler utiles à formaliser les problèmes d'accès conflictuel à une ressource, ainsi qu'à trouver des solutions d'allocation équilibrée et des algorithmes permettant d'apprendre ces solutions aux systèmes de communications de manière autonome et décentralisée.

D'une manière générale, l'objectif sera de permettre aux émetteurs d'adapter leurs paramètres de transmission de manière intelligente, autonome et décentralisée selon l'environnement acoustique dans lequel ils évoluent, l'environnement incluant ici le canal et les utilisateurs qui y accèdent. Aussi, étant donné le contexte non-coopératif et décentralisé qui nous est imposé, le choix d'une stratégie de transmission doit être une décision prise localement, à partir d'informations sur l'environnement connues ou apprises localement. La capacité d'adapter sa stratégie est importante d'une part parce que les systèmes de communication ASM ne sont pas en mesure de prévoir à l'avance la nature des sources d'interférences rencontrées, mais aussi parce que l'environnement marin est dynamique, et les conditions du canal de communication varient à différentes échelles de temps et d'espace. À plusieurs égards, les problématiques du sujet de cette thèse sont proches de celles ayant motivé le développement des recherches sur la radio cognitive dans le monde des télécommunications terrestres [Baldo et al., 2008; Mitola, 2000]. Les émetteurs doivent en effet être capables, étant donné l'état du canal ASM, d'identifier et de mettre en œuvre la meilleure stratégie de transmission au sens d'un certain critère : maximisation du débit, du rapport signal à bruit ou de la capacité du lien, minimisation de la probabilité d'erreur ou de la puissance transmise sous contrainte de débit, etc. Cependant, les spécificités de l'environnement marin, les sources d'interférences de différentes natures et l'absence de protocole de communications et d'infrastructures réseau standardisées posent de nouvelles difficultés.

## Plan du manuscrit et contributions

Ce manuscrit s'organise autour de cinq chapitres. Chaque chapitre débute par sa propre introduction, permettant de situer le contexte et l'état de l'art, et se conclut par une ouverture sur les chapitres suivants et, le cas échéant, une synthèse des contributions.

Le premier chapitre présente des généralités sur les communications et le canal ASM, afin de mieux préciser le contexte qui n'a été présenté que succinctement dans cette introduction générale. Plusieurs détails importants sont donnés sur la physique du milieu et les distorsions subies par les signaux de communications. Le canal ASM est présenté en toute généralité comme canal d'interférence, avec des justifications plus précises quant à l'utilisation de ce modèle. Le chapitre se termine par un bref état de l'art sur les travaux menés sur les thèmes de l'adaptation au canal et de la gestion de l'interférence en communications ASM.

Le deuxième chapitre présente les éléments de théorie des jeux qui seront utilisés dans le manuscrit. La théorie des jeux a été l'outil de modélisation central employé

dans cette thèse. De nombreux modèles de jeux différents existent au sein de cette théorie : jeux bayésiens, stochastiques, répétés, etc... Seuls les modèle de jeux sous forme stratégique et de satisfaction ont été utilisés, car ils sont suffisants pour déduire des stratégies d'équilibre dans les problèmes d'allocation de ressources en communications ASM qui seront traités dans les chapitres suivants. Le chapitre 2 donne donc tous les éléments permettant de décrire et résoudre ces jeux. Les algorithmes d'apprentissage des équilibres sont également présentés.

Le troisième chapitre présente un problème de partage de spectre entre systèmes de communications ASM non-coopératifs utilisant la modulation OFDM. L'objectif est, pour chaque émetteur, de trouver l'allocation de puissance maximisant son débit (dépendant de l'interférence générée par les autres émetteurs) sous contrainte de puissance. Ce problème peut se formaliser comme un jeu non-coopératif où les joueurs sont les émetteurs OFDM. La formalisation du problème et ses solutions sont similaires à celles présentées par [Scutari et al., 2008a,b] en théorie des communications. Une méthode de partage de spectre robuste face aux erreurs d'estimation du canal est également proposée. Ce chapitre couvre les articles

- A. Pottier, F.-X. Socheleau, and C. Laot. Robust Noncooperative Spectrum Sharing Games in Underwater Acoustic Interference Channels. *IEEE J. of Oceanic Eng.*, 42(4) :1019 – 1034, October 2017
- A. Pottier, F.-X. Socheleau, and C. Laot. Distributed Power Allocation Strategy in Shallow Water Acoustic Interference Channels. In *IEEE Int. workshop on Signal Processing Advances in Wireless Communications (SPAWC) 2016*, Jul. 2016b
- A. Pottier, F.-X. Socheleau, and C. Laot. Adaptive Power Allocation for Noncooperative OFDM Systems in UWA Interference Channels. In *Proc. 3rd Underwater Acoustic Communications and Networking Conf. (UComms)*, Sept. 2016c

Le quatrième chapitre présente le problème de partage de ressource dans le canal ASM sous contrainte de qualité de service. Le problème de partage de spectre minimisant la puissance transmise sous contrainte de débit d'information, complémentaire du chapitre 3, est présenté en première partie. On aboutit à des solutions similaires, cependant la faisabilité du problème n'est pas toujours garantie. La difficulté est contournée grâce aux jeux de satisfaction [Goonewardena et al., 2017; Ross and Chaib-draa, 2006]. Un algorithme d'apprentissage des équilibres de satisfaction ne nécessitant qu'une connaissance de l'environnement encodée sur 1 bit d'information est proposé et plusieurs exemple de jeux faisant intervenir des émetteurs OFDM et DSSS non coopératifs dans le canal ASM sont étudiés. Ce chapitre reprend les articles

- A. Pottier, F.-X. Socheleau, and C. Laot. Quality-of-Service Satisfaction Games for Noncooperative Underwater Acoustic Communications. *IEEE Access*, 6, April 2018b
- A. Pottier, F.-X. Socheleau, and C. Laot. Power-Efficient Spectrum Sharing for Noncooperative Underwater Acoustic Communication Systems. In *MTS/IEEE OCEANS'16 Monterey*, Sept. 2016a

Le cinquième et dernier chapitre propose une ouverture sur l'apprentissage par renforcement dans les réseaux sous marins, à travers l'étude d'un exemple d'application à un réseau de capteurs. Le réseau est constitué de capteurs déployés au fond de la mer, dédiés à la collecte de mesures physiques (pression, température). Ces capteurs sont

supposés devoir remonter leur données régulièrement à la surface en communiquant, par voie acoustique, avec des bouées de surface considérées comme récepteurs. L'asynchronisme, une surcharge minimale du réseau et la capacité d'adaptation des capteurs aux fluctuations temporelles de l'environnement sont des propriétés désirables dans ce genre de réseaux afin d'en faciliter le déploiement, la maintenance et la configuration. On propose un schéma d'accès aux bouées basée sur l'apprentissage par renforcement qui permet d'offrir toute ces propriétés aux capteurs communicants. Ce travail est issu d'une collaboration avec l'Université de York, qui a donné lieu à un séjour de trois mois au sein du département Electronic Engineering. Il reprend l'article

- A. Pottier, P. D. Mitchell, F.-X. Socheleau, and C. Laot. Q-Learning based Adaptive Channel Selection for Underwater Sensor Networks. In *Proc. 4th Underwater Acoustic Communications and Networking Conf. (UComms)*, August. 2018a

Le manuscrit se termine par une conclusion générale, synthétisant les différentes contributions et les perspectives pour de futurs travaux sur le sujet.



# Chapitre 1

## Communications dans le milieu de propagation sous-marin

L'objectif d'un système de communications numériques est de transmettre de l'information d'un émetteur à un récepteur, sous forme de symboles ou d'éléments binaires, à un débit mesuré habituellement en symboles par secondes ou en bits par secondes. Cette transmission s'opère dans un milieu dictant les lois de propagation des ondes sous la forme desquelles se matérialisent physiquement les signaux de communications. Ces signaux transportent les symboles contenant l'information communiquée et occupent une largeur de bande déterminée dans le spectre fréquentiel, autour d'une ou plusieurs fréquences porteuses. Le milieu physique constitue ce que l'on appelle un canal de communication. Les canaux de communication possèdent une *capacité* en terme de nombre de symboles par seconde pouvant être transmis avec un taux d'erreur arbitrairement faible. Cette notion de capacité représente mathématiquement le compromis entre performance (débit) et fiabilité (faible taux d'erreur) d'un système de communications numériques. Lorsque qu'une communication s'établit à travers un canal, une certaine ressource physique de ce canal est utilisée : typiquement, une certaine bande de fréquence est utilisée par un système pendant une certaine durée. L'un des principaux problèmes qui se pose lors de la conception de réseaux de communications est le partage de la ressource disponible pour établir des transmissions fiables entre chaque élément du réseau. Lorsque que plusieurs systèmes non-coopératifs entrent en compétition pour l'usage d'une ressource, il interfèrent au détriment à la fois de leur robustesse et de leur fiabilité.

Les propriétés physiques de l'environnement sous-marin permettent aux ondes acoustiques de se propager sur de longues distances. Les ondes acoustiques sous-marines (ASM) sont ainsi le support privilégié des signaux de communication dans ce milieu. Cependant, la physique se présente aussi comme un facteur limitant des performances des modems de communications acoustiques sous-marins qui, en comparaison des modems radio-fréquence (RF), fonctionnent à des débits très bas. Ces différences de performances sont à l'image des différences d'ordres de grandeur entre certains paramètres physiques fondamentaux de chacun des deux milieux de propagation. À titre d'exemple, le rapport entre la célérité de l'ondes électromagnétique dans l'air et celle de l'onde acoustique sous-marine est de  $10^5$  ( $3 \cdot 10^8$  m/s contre environ 1500 m/s,

respectivement). Les fréquences porteuses des communications RF peuvent atteindre quelques gigahertz, tandis qu'elles n'excèdent pas quelques dizaines de kilohertz pour les communications ASM. Les bandes passantes dans ce milieu sont sévèrement limitées, et ce, d'autant plus que la distance de transmission augmente. Par conséquent, le canal de communication ASM est généralement considéré comme large bande, les fréquences porteuses et les largeurs des bandes de transmission ayant le même ordre de grandeur. La limite actuelle de performance des modems ASM se traduit par un produit débit-distance de transmission habituellement inférieur à  $40 \text{ kbits/s} \times \text{km}$  [Kilfoyle and Baggeroer, 2000]. La cohabitation dans le même canal de sources acoustiques de différentes origines est également difficile, en raison, d'une part, de l'absence d'interopérabilité, et d'autre part à cause du manque de ressources disponibles pour communiquer sans interférences entre utilisateurs. Ceci, couplé aux longs délais de propagation, pose des difficultés à concevoir des protocoles fiables et efficaces adaptés aux réseaux de communications. Enfin, le milieu marin présente des variabilités et des hétérogénéités à plusieurs échelles spatiales et temporelles, ce qui nécessite une certaine flexibilité et une capacité d'adaptation des modems ASM à la dynamique de l'environnement dans lequel ils évoluent.

Ce premier chapitre a pour objectif de présenter en détail les particularités du milieu ASM considéré en tant que canal de communication afin de donner au lecteur une meilleure compréhension des différentes hypothèses de travail qui seront posées dans les chapitres suivants. Les propriétés physiques importantes du canal de communication ASM sont tout d'abord présentées avant de donner plus de précisions sur les distorsions subies par les signaux de communications se propageant sous l'eau. Les communications ASM sont ensuite présentées comme un problème de communications dans un canal d'interférence du point de vue de la théorie de l'information, et le chapitre se conclut sur un bref état de l'art de la recherche dans ce domaine.

## 1.1 Propriétés physiques du milieu de propagation

### 1.1.1 Célérité de l'onde acoustique

La célérité d'une onde dans un milieu est un paramètre physique fondamental et déterminant de sa propagation. Dans l'eau, la célérité de l'onde acoustique sous l'eau est de l'ordre de  $1500 \text{ m.s}^{-1}$ . Cependant, le milieu marin n'est pas iso-célère. Cette grandeur varie en fonction de la température, de la salinité et de la profondeur (à cause de la pression) selon la relation [Clay and Medwin, 1998]

$$c = 1449.2 + 4.6T - 0.055T^2 + 0.00029T^3 + (1.34 - 0.01T)(S - 35) + 0.016z, \quad (1.1)$$

où  $T$  est la température en degré en centigrades,  $S$  est la salinité en ‰ et  $z$  est la profondeur en mètres. Dans une colonne d'eau, le profil de célérité donne la vitesse du son en fonction de la profondeur. En général, la célérité de l'onde acoustique dans le milieu marin est comprise entre 1440 et 1540  $\text{m/s}$ .

Des différences significatives existent entre les profils de célérité des environnements grands fonds et ceux des environnements petits fonds. En petits fonds, où les

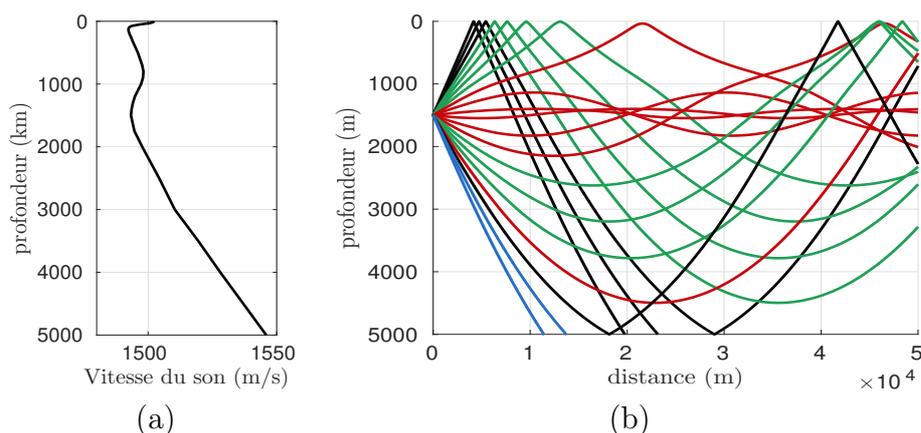


FIGURE 1.1 – (a) Exemple de profil de célérité typique d'un canal grands fonds. (b) Tracé de rayons avec une source à 1500 m de profondeur. En noir : réflexions surface et fond, en bleu : réflexion fonds, en vert : réflexions surface, en rouge : trajets directs.

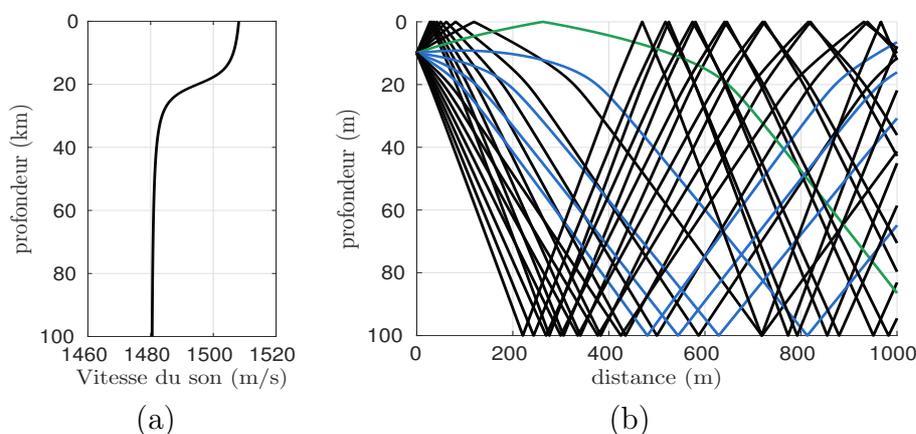


FIGURE 1.2 – (a) Exemple de profil de célérité typique d'un canal petits fonds. (b) Tracé de rayons avec une source à 10 m de profondeur. En noir : réflexions surface et fond, en bleu : réflexion fonds, en vert : réflexions surface.

profondeurs de colonne d'eau excèdent rarement la centaine de mètres, la vitesse du son est principalement déterminée par la température. Elle diminue avec la température, et donc, généralement, avec la profondeur. En grands fonds, les températures devenant faibles et homogènes à partir d'une certaine profondeur, la célérité augmente du fait d'une augmentation de la pression. Ces différences conditionnent fortement la propagation des ondes acoustiques, comme il le sera expliqué à la section suivante.

### 1.1.2 Trajets multiples

L'onde ASM suit de multiples trajets de propagation lors son cheminement dans le milieu. Ce phénomène trouve son explication dans deux propriétés physiques : la présence d'interfaces de réflexion formées par la surface de l'eau et les fonds marins, et la célérité variable avec la profondeur.

Les réflexions aux interfaces sont majoritairement responsables du phénomène

de trajets multiples dans les canaux petits fonds. La variation de célérité est analogue à une variation d'indice de réfraction et à partir des lois de Snell-Descartes il est possible de montrer que les rayons, définis comme les directions perpendiculaires au front d'onde, ont tendance à se courber vers la région de la colonne d'eau où la célérité est minimale. La réfraction est la cause principale du phénomène de trajets multiples dans le cas de transmissions longues distances dans un canal grand fond. De plus, la courbure des rayons peut créer des zones d'ombres où aucune énergie acoustique en rapport avec l'onde considérée ne peut être perçue. Les effets des divers phénomènes responsables de la propagation par trajets multiples sont illustrés aux figures 1.1 et 1.2.

### 1.1.3 Pertes de propagation

Les pertes de propagation subies par un signal se propageant dans le canal ASM dépendent à la fois de la fréquence et de la distance de transmission. La dépendance à la distance s'explique par les pertes par divergence géométrique résultantes de l'étalement spatial de l'énergie acoustique à mesure que l'onde se propage, tandis que la dépendance à la fréquence vient des pertes d'amortissement. Ces dernières dépendent du milieu de propagation. Dans l'eau de mer, l'amortissement des ondes acoustiques résulte d'une part de la viscosité de l'eau, conséquence d'une dissipation d'énergie thermique, et d'autre part de la relaxation de composés chimiques tels que l'acide borique [Lurton, 2011]. Ces pertes augmentent avec la fréquence.

Soit un signal de fréquence  $f$  (en kHz) se propageant dans le canal acoustique selon un trajet de longueur  $d$  (en mètres). Les pertes de propagation en dB subies par ce signal, négligence faite des pertes de réflexion, sont données par [Urlick, 1983] :

$$A_{dB}(d, f) = k \times 10 \log_{10}(d) + 10^{-3} \times d \times \alpha_{dB}(f) \quad (1.2)$$

où  $k$  est un facteur de dispersion spatiale compris entre 1 et 2 (1 : diffusion cylindrique, 2 : diffusion sphérique) et  $\alpha_{dB}(f)$  est le coefficient d'amortissement exprimé en dB perdus par kilomètre en fonction de la fréquence. Plusieurs modèles existent pour ce coefficient, où des paramètres tels que la salinité, la température de l'eau et la profondeur peuvent être pris en compte. En communications ASM, le modèle couramment employé est celui donné par la formule de Thorp [Lurton, 1998; Urlick, 1983], soit

$$\alpha_{dB}(f) \approx 0.11 \times \frac{f^2}{f^2 + 1} + 44 \times \frac{f^2}{f^2 + 4100} + 2.75 \times 10^{-4} \times f^2 + 0.003. \quad (1.3)$$

Les figures 1.3-(a) et 1.3-(b) montrent l'influence de la distance et de la fréquence de transmission d'un signal acoustique sous-marin sur les pertes par diffusion et par amortissement qu'il subit.

Les réflexions fonds et surfaces induisent également des pertes. Ces deux interfaces ne sont pas parfaitement planes et présentent des irrégularités aléatoires dues au relief et aux vagues provoquées par le vent et les marées. Une onde, au contact d'une surface irrégulière, se voit réfléchiée en une composante spéculaire cohérente et une composante diffusée aléatoirement dans toutes les directions. Une partie de l'énergie est également transmise au second milieu. Lors d'une réflexion, l'onde incidente subit

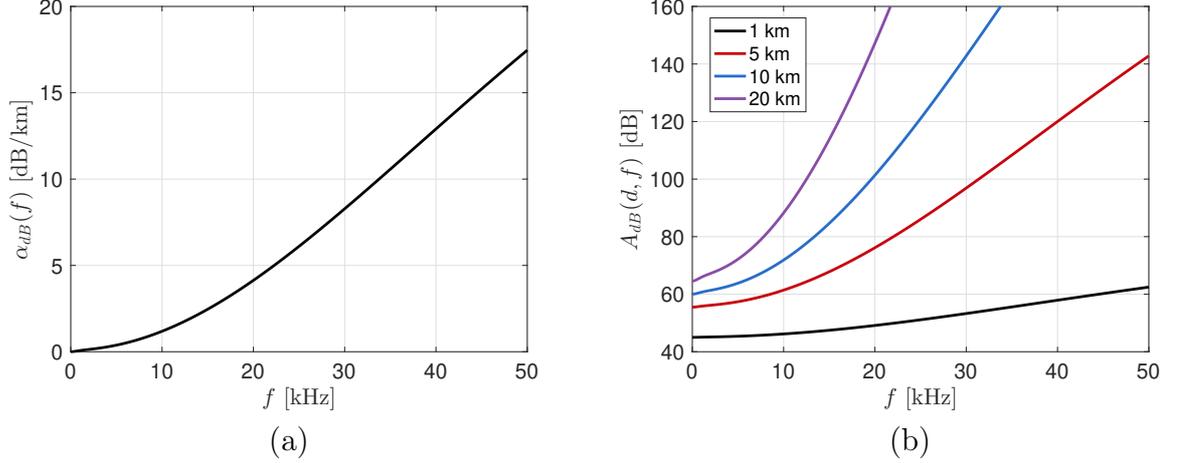


FIGURE 1.3 – Pertes de propagation dans le canal ASM. Gauche : coefficient d’amortissement. Droite : pertes de propagation selon (1.2).

des pertes dans la partie diffuse et la partie transmise, dont l’importance dépend de plusieurs paramètres tels que l’état de la mer dans le cas d’une réflexion à la surface, le relief du fond marin et le type de sédiments lorsque la réflexion se produit sur le fond marin.

La perte d’énergie de la composante cohérente réfléchié par rapport à l’énergie initiale de l’onde incidente peut être modélisée par un coefficient de réflexion exprimé en fonction de la fréquence  $f$  (en kHz), de la célérité  $c$  et de l’angle d’incidence  $\theta$  à la surface du dioptre idéal par [Lurton, 1998]

$$\tilde{R}(\theta) = \bar{R}(\theta) e^{-2\left(\frac{2\pi f}{c}\sigma\right)^2 \cos^2(\theta)}, \quad (1.4)$$

où  $\sigma$  est l’écart-type des irrégularités aléatoires de la surface autour du dioptre idéal. Le coefficient de réflexion au dioptre idéal,  $\bar{R}(\theta)$  dépend des caractéristiques acoustiques de l’eau ( $\rho, c$ ) et du second milieu ( $\rho', c'$ ) en termes de densités et de célérités ainsi que de l’angle de réfraction  $\theta'$  de l’onde transmise au second milieu<sup>1</sup>. Il est donné par

$$\bar{R}(\theta) = \frac{\rho c \cos \theta' - \rho' c' \cos \theta}{\rho c \cos \theta' + \rho' c' \cos \theta}. \quad (1.5)$$

Ainsi voit-on grâce l’équation (1.4) que la fréquence du signal participe aussi aux pertes par réflexions. Son augmentation participe également à la réduction du rapport de l’énergie de la composante spéculaire cohérente sur l’énergie de la composante diffusée après réflexion.

#### 1.1.4 Sources et modèles de bruit

L’environnement sous-marin est bruyant car peuplé de nombreuses sources de sons, même à l’état naturel, abstraction faite des bruits provoqués par les activités humaines. La densité spectrale de puissance du bruit ambiant peut être décrite par des mo-

1. En cas de réflexion totale aucune énergie n’est transmise au second milieu et  $\bar{R}(\theta) = 1$ .

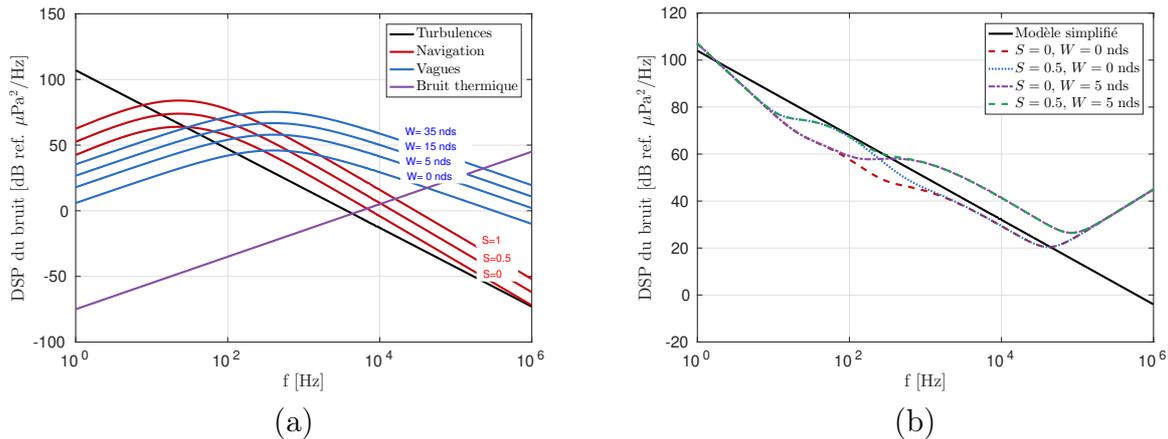


FIGURE 1.4 – Gauche : Densités spectrales de puissance des contributions au bruit selon le modèle de l'équation (1.6). Droite : Densité spectrale de puissance de bruit en sommant les différentes contributions. Modèle simplifié selon l'équation 1.7 avec  $N_0 = 50$  dB.

dèles empiriques [Coates, 1990; Knudsen et al., 1948; Urick, 1984; Wenz, 1972]. En communication ASM, le modèle de Coates est couramment utilisé [Stojanovic and Preisig, 2009]. Ce modèle propose comme sources principales de bruit ambiant les turbulences, le trafic maritime, les vagues et le bruit thermique. Ces différentes contributions sont chacune prépondérante dans une bande de fréquence. Elles sont modélisées par des variables aléatoires dont les densités spectrales de puissance respectives sont exprimées, par Hertz et en dB par rapport à une pression de  $1\mu\text{Pa}$  à 1 mètre de la source, selon les formules suivantes :

$$\begin{aligned}
 \text{Turbulences} & : N_T(f) = 17 - 30 \log_{10}(f) \\
 \text{Trafic maritime} & : N_S(f) = 40 + 20(S - 0.5) + 26 \log_{10}(f) - 60 \log_{10}(f + 0.03) \\
 \text{Vagues} & : N_W(f) = 50 + 7.5\sqrt{W} + 20 \log_{10}(f) - 40 \log_{10}(f + 0.4) \\
 \text{Bruit thermique} & : N_C(f) = -15 + 20 \log_{10}(f),
 \end{aligned} \tag{1.6}$$

où  $f$  est la fréquence donnée en kHz,  $S$  est un facteur de trafic maritime compris entre 0 et 1, et  $W$  est la vitesse du vent en m/s.

Ce genre de modèle ne prend évidemment pas en compte toutes les sources de bruits pouvant être rencontrées dans l'océan. Les sources biologiques (claquement des crevettes, mammifères marins, ...), par exemple, sont ici négligées. Des exemples de densités spectrales de puissance de bruit sont illustrés à la figure 1.4-(b), avec le détail des différentes contributions de l'équation (1.6) à la figure 1.4-(a) pour plusieurs valeurs des paramètres de vitesse du vent et d'activité du trafic maritime. Dans la bande de fréquence allant de 100 Hz à 50 kHz, dans laquelle opère la vaste majorité des systèmes de communication ASM, on peut approximer la densité spectrale de puissance du bruit par une fonction linéaire décroissante sur l'échelle logarithmique des fréquences [Knudsen et al., 1948] :

$$N_K(f) = N_0 + 18 \log_{10}(f) \tag{1.7}$$

où la fréquence  $f$  est donnée en kHz et  $N_0$  est une constante choisie selon l'état de la mer. La table 1.1 donne la correspondance entre l'état de la mer, la vitesse du vent et la constante  $N_0$  du modèle de bruit simplifié. Ce dernier modèle est utile pour évaluer les

$W$ (nds)	0	1-3	4-6	7-10	11-16	17-21	22-27
État de la mer (Beaufort)	0	1	2	3	4	5	6
$N_0$ (dB)	44.5	50	55	61.5	64.5	66.5	68.5

TABLE 1.1 – Correspondances entre vitesse du vent en noeuds, l'état de la mer en Beaufort et la constante  $N_0$  en dB.

performances de systèmes de communication ASM par simulations lorsqu'on ne dispose pas d'information précise quant à l'intensité et la nature des sources de bruit ambiant. On peut également paramétrer la constante  $N_0$  selon une puissance de bruit spécifiée dans la bande de transmission considérée.

Le bruit ambiant ASM est souvent considéré comme étant distribué selon une loi normale avec une densité spectrale de puissance continue [Lurton, 1998; Stojanovic and Preisig, 2009]. Il faut noter cependant que, selon le contexte, ces hypothèses peuvent être mises en défaut. Par exemple, le claquement des crevettes est plutôt modélisé par une distribution de Lévy que par une gaussienne [Mahmood et al., 2012], et les bruits de bateaux peuvent présenter des DSP constituées de raies spectrales aux fréquences de rotations des hélices et leurs harmoniques.

### 1.1.5 Hétérogénéités et variabilité de l'environnement

Le milieu sous-marin est hétérogène et dynamique. Les propriétés physiques dictant la propagation des ondes acoustiques fluctuent spatialement et temporellement, et à différentes échelles.

Comme mentionné à la section 1.1.1, le profil de célérité varie selon les zones géographiques en conséquence de changements de température, de salinité et de pression. Les cycles saisonniers et les marées conditionnent également les profils de célérité, notamment en petits fonds du fait des changements de températures. Les conditions de propagation peuvent donc être très différentes selon les endroits du globe et les époques de l'année.

À des échelles plus locales, des changements dans les conditions de propagation depuis une source vers un récepteur peuvent se produire du fait de mouvements de la source et/ou du récepteur. Ces mouvements peuvent être volontaires ou le résultat des vagues ou d'une dérive par rapport au courant marin. Les effets de distorsions produits par la propagation de l'onde acoustique sur le signal transmis varient alors au gré des petites variations des distances et des profondeurs de transmission. La rugosité variable des interfaces fond et surface est également source de fluctuations des amplitudes, des phases et des retards des échos réfléchis [Loubet et al., 1990]. Certains phénomènes plus sporadiques provoquent des variations imprévisibles des niveaux de bruits. C'est le cas par exemple de l'activité humaine, des pépiements des mammifères marins, ou de la formation de bulles d'air à la surface due aux vagues.

Du point de vue de la théorie des communications, les phénomènes présentant des dynamiques temporelles de l'ordre de la journée à la saison peuvent être considérés

responsables de variations à longs termes du rapport signal à bruit (RSB), dont l'impact sur la performance théorique d'un récepteur est connu [Gallager, 2008; Proakis, 2001]. Les phénomènes dont les variabilités se présentent à des échelles de temps de l'ordre d'un temps symbole, d'un paquet, ou d'une trame impactent, en revanche, les performances instantanées de la réception.

## 1.2 Impact de la physique du milieu sur les signaux de communication

### 1.2.1 Fréquence optimale et bande passante

Sur la base des modèles de bruit et de pertes de propagation données à la section précédente, il est possible d'établir des bilans de liaison pour une sinusoïde de fréquence  $f$  se propageant sur un chemin de longueur  $l$  [Stojanovic and Preisig, 2009]. En considérant un trajet direct, la grandeur  $P_{dB}(l, f) = -A_{dB}(l, f) - N_K(f)$  exprime, en décibels, la perte en terme de rapport signal sur bruit subie par la sinusoïde. La fréquence  $f$  est donnée en kHz et la distance  $l$  en km afin d'être cohérent avec les équations (1.2) et (1.7).

La figure 1.5 illustre plusieurs exemples de ce rapport signal à bruit (SNR) en fonction de la fréquence pour différentes distances de transmission. Plusieurs observations peuvent être faites. La première est qu'un trajet de propagation est d'autant plus sélectif en fréquence que la distance parcourue augmente. Cela implique que la bande passante disponible pour une communication et donc la capacité du canal associé diminuent avec la distance. Il apparaît ensuite qu'il existe une fréquence optimale au sens où les pertes de bilan de liaison sont minimales pour une distance et une bande transmission fixées. Enfin, le canal ASM est large bande car les bandes de fréquences utilisables ont des largeurs du même ordre de grandeur que leurs fréquences centrales.

Ce modèle de bilan de liaison est trop optimiste pour réaliser des études réalistes de performances de modems ASM ou de capacité du canal. Les trajets multiples parti-

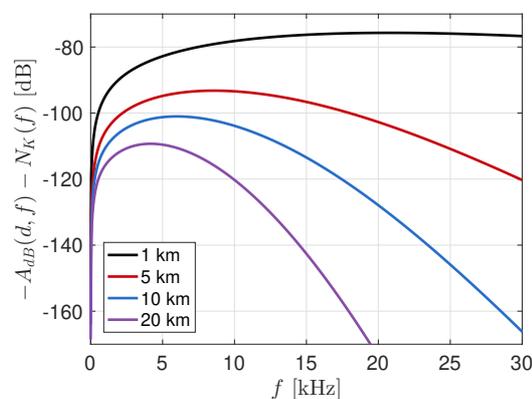


FIGURE 1.5 – Bilan de liaison pour une sinusoïde de fréquence  $f$  se propageant à différentes distances. Modèle de bruit simplifié selon l'équation 1.7 avec  $N_0 = 50$  dB.

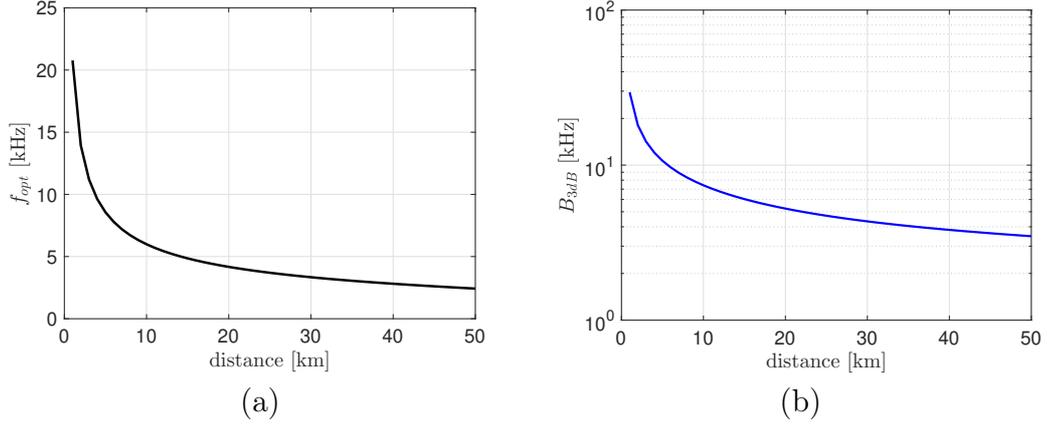


FIGURE 1.6 – Gauche : Fréquence optimale en fonction de la distance. Droite : Bande à 3 dB autour de la fréquence optimale.

cipent également à la sélectivité fréquentielle du canal et il faudrait prendre également en compte sa variabilité temporelle. Cependant il est utile pour donner des bornes en termes de bande de fréquence raisonnablement utilisable pour une distance de transmission fixée. Sur les figures 1.6-(a) et 1.6-(a) sont illustrées la fréquence et la bande à 3 dB optimales, respectivement, en fonction de la distance de transmission.

### 1.2.2 Effets Doppler

De la faible célérité des ondes ASM résultent des effets Doppler prononcés du point de vue récepteur lorsque lui et/ou la source sont en mouvement. Dans un canal bande étroite, où la fréquence porteuse est très grande devant la largeur de bande des signaux, cet effet est approximé par un simple décalage de la fréquence centrale du signal reçu. Dans le canal ASM large bande, cet effet nécessite une modélisation plus précise, faisant explicitement intervenir la compression/dilatation temporelle du signal via une remise à l'échelle de la variable temporelle [Sharif et al., 2000]. En d'autres termes, considérant  $s(t)$  le signal issu de la source acoustique et  $r(t)$  le signal reçu après la propagation à travers un canal idéal, seulement affecté par l'effet Doppler, nous avons

$$r(t) = s((1 - \Delta)t) \quad (1.8)$$

où  $\Delta = v/c$  est le facteur Doppler, ratio de la vitesse relative  $v$  des mouvements de la source et du récepteur et de la célérité dans le milieu.

Comme il l'a été mentionné dans la section précédente, les déplacements de la source et du récepteur peuvent être volontaires ou le fruit d'un mouvement qui leur est imprimé par les vagues ou les courants. Or, la transformée de Fourier du signal affecté par l'effet de contraction/dilatation Doppler  $s((1 - \Delta)t)$  est donnée par  $S\left(\frac{f}{\Delta}\right) / |\Delta|$ . Considérant que le signal  $s(t)$  occupe une bande  $B$  autour de sa fréquence centrale  $f_c$ , l'effet Doppler lui fait alors subir une dilation/compression fréquentielle et décalage de sorte qu'il occupe une bande  $\left[-\frac{B}{2}(1 - \Delta) ; \frac{B}{2}(1 - \Delta)\right]$  autour de  $(1 - \Delta)f_c$ .

### 1.2.3 Réponse impulsionnelle de canal doublement dispersif

Un signal transmis par un émetteur est reçu, après la propagation dans le canal ASM, comme une somme d'échos retardés dont l'amplitude et la phase fluctuent dans le temps. Chaque écho est la résultante d'un trajet de propagation particulier, sélectif en fréquence et dont les fluctuations temporelles reflètent celles du milieu marin. De plus, le retard de chaque écho est lui même variable dans le temps du fait des éventuels effets Doppler. Résumé mathématiquement, on a le modèle général suivant pour la réponse impulsionnelle du canal ASM, exprimé en fonction des retards  $\tau$  et du temps  $t$  :

$$h(\tau, t) = \sum_{k=1}^K c_k(t) g_k(\tau - \tau_k(t)) \quad (1.9)$$

où  $K$  est le nombre de trajets,  $c_k(t) \in \mathbb{C}$  est un coefficient complexe modélisant les fluctuations temporelles du trajet  $k$ ,  $g_k(\tau)$  exprime sa sélectivité fréquentielle due à l'absorption, et la dépendance au temps du retard  $\tau_k(t)$  capture le caractère large bande du canal et les effets Doppler. Pour un signal  $s(t)$  transmis à travers le canal  $h(\tau, t)$ , le signal reçu s'exprime par

$$r(t) = \int h(\tau, t) s(t - \tau) d\tau. \quad (1.10)$$

Le canal est dit dispersif en temps (ou sélectif en fréquences) si pour  $s(t) = \delta(t)$ , la durée de  $r(t)$  est strictement positive. Réciproquement, un canal est dispersif en fréquence si pour  $s(t) = e^{j2\pi f_0 t}$ , la bande occupée par  $r(t)$  est strictement positive. De la dispersion temporelle se produit dès lors que le nombre de trajets  $K$  est strictement supérieur à 1, tandis que la dispersion fréquentielle se traduit par la dépendance au temps de la phase, de l'amplitude et des retards de ces trajets, à travers les coefficients  $c_k(t)$  et le retard  $\tau_k(t)$ . Du fait de la propagation de l'onde acoustique par trajets multiples et des fluctuations temporelles de l'environnement marin, le canal ASM est doublement dispersif.

Les fluctuations temporelles de l'environnement marin sont dépendantes d'un grand nombre de paramètres physiques, qui varient eux-mêmes selon différentes échelles de temps. Il est alors difficile d'adopter un modèle de canal déterministe. Par conséquent, la réponse impulsionnelle  $h(\tau, t)$  est plutôt modélisée comme un processus stochastique. Lorsque le canal est stationnaire au sens large et à diffusion décorrélée (WS-SUS), il est possible de décrire comment le canal disperse l'énergie dans le domaine du retard et des fréquences Doppler à travers la fonction de diffusion [Bello, 1963; Kilfoyle and Baggeroer, 2000]

$$S_h(\tau, f) = \int R_h(\tau, u) e^{-j2\pi f u} du \quad (1.11)$$

avec

$$R_h(\tau, u) = \mathbb{E} [(h(\tau, t) - \mathbb{E}[h(\tau, t)]) (h(\tau, t + u) - \mathbb{E}[h(\tau, t + u)])^*] \quad (1.12)$$

l'auto-covariance du canal considéré stationnaire au sens large par rapport au temps  $t$ . Si le canal est WSSUS, la fonction de diffusion suffit à le caractériser. De cette fonction

peuvent être dérivés le spectre Doppler

$$S_{\Delta}(f) = \int |S_h(\tau, f)|^2 d\tau, \quad (1.13)$$

qui caractérise la dispersion fréquentielle du canal, et le profil puissance-délai

$$S_d(\tau) = \int |S_h(\tau, f)|^2 df, \quad (1.14)$$

qui caractérise sa dispersion temporelle.

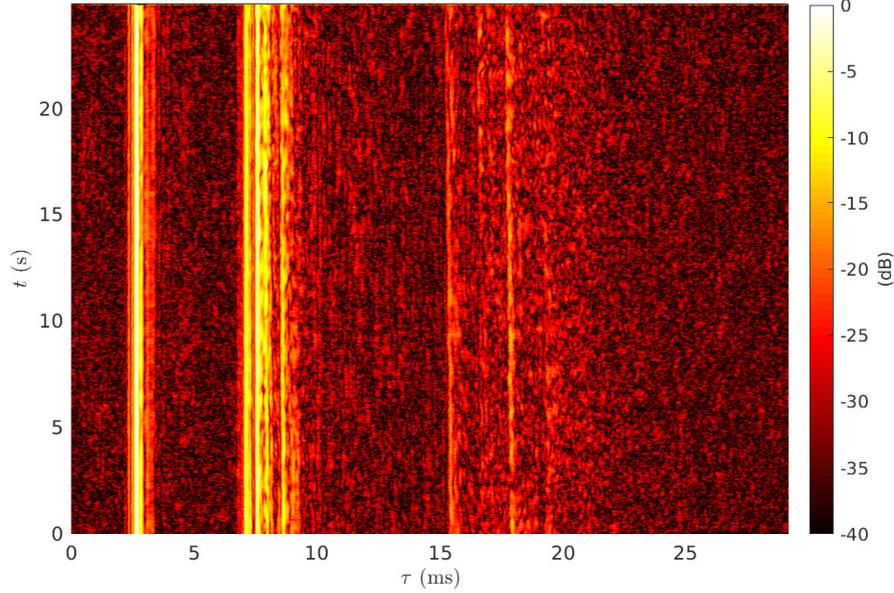


FIGURE 1.7 – Réponse impulsionnelle  $h(t, \tau)$  d'un canal ASM sondé en Mer Méditerranée.

L'hypothèse WSSUS peut être mise en défaut par de nombreux canaux ASM [Socheleau et al., 2011; van Walree, 2011] et sa validité dépend à la fois de l'échelle de temps d'observation et des propriétés intrinsèques du canal. En effet, les statistiques du canal ASM peuvent fluctuer plus ou moins lentement, au gré des phénomènes de différentes échelles temporelles décrits à la section 1.1.5. Enfin, il n'existe pas de consensus dans la communauté des communications ASM quant à la modélisation statistique des évanouissements du canal. Parmi plusieurs exemple possibles, on peut citer [Socheleau et al., 2011] où une distribution de Rice est choisie pour modèle des évanouissements rapides supposés stationnaires conditionnellement aux fluctuations long-terme du canal considérées déterministes. Ces fluctuations sont modélisées par une loi log-normale dans [Qarabaqi and Stojanovic, 2013]. Une distribution de Rayleigh est choisie dans [van Walree and Otnes, 2013] pour les évanouissements rapides de canaux sondés en mer, tandis que d'autres canaux semblent être mieux modélisés par une distribution K [Yang and Yang, 2006]. Ainsi, le choix d'un modèle statistique de canal ASM est toujours sujet à discussions, chaque canal réel pouvant être plus ou moins proche de ce modèle selon l'endroit et le moment où il a été sondé. La figure 1.7 montre un canal sondé en rade de Toulon (la compression/dilatation temporelle due à l'effet Doppler a été compensée), à une fréquence de 10,5 kHz sur une bande de 8 kHz. Cette figure illustre bien la dispersion et la variabilité temporelle typique des canaux ASM.

## 1.2.4 Temps et bande de cohérence

L'impact de la dispersion en temps du canal prend la forme d'interférences entre symboles (IES), qui sont en général d'autant plus importantes que la rapidité de modulation est grande. L'étalement temporel d'un canal ASM peut être de l'ordre de plusieurs dizaines de millisecondes jusqu'à la centaine de millisecondes et augmente avec le ratio distance/profondeur dans les canaux petits fonds [Kilfoyle and Baggeroer, 2000; Stojanovic and Preisig, 2009; van Walree, 2011]. La dispersion temporelle du canal étant le phénomène dual de sa sélectivité fréquentielle, la connaissance de l'étalement temporel nous informe également quant à la bande de fréquence utilisable sans subir l'IES.

L'étalement temporel du canal peut se définir à travers le profil délai-puissance par [van Walree, 2011]

$$\tau_{rms} = \sqrt{\frac{\int (\tau - \tau_g)^2 S_d(\tau) d\tau}{\int S_d(\tau) d\tau}} \quad (1.15)$$

où  $S_d(\tau)$  est défini à l'équation (1.14) et  $\tau_g = \frac{\int \tau S_d(\tau) d\tau}{\int S_d(\tau) d\tau}$ . La bande de cohérence du canal est généralement définie comme l'inverse de son étalement temporel, soit

$$B_{coh} = \frac{1}{\tau_{rms}}. \quad (1.16)$$

Elle correspond à une largeur moyenne de sous-bande dans laquelle le canal peut être approximé comme non sélectif en fréquence. Étant donnés les étalements temporels typiques des canaux ASM, les bandes de cohérence correspondantes sont généralement de l'ordre de quelques dizaines à la centaine de Hz.

La dispersion fréquentielle introduite par les effets Doppler, la diffusion aux interfaces de réflexions rugueuses, et par les variabilités intrinsèques du canal se traduit par de l'interférence entre sous-porteuses (IESP) lorsque l'on considère une transmission multiplexée en fréquence, comme dans le cadre des modulations orthogonal frequency division multiplexing (OFDM). L'étalement fréquentiel (ou étalement Doppler) se définit de manière analogue à l'étalement temporel, soit

$$D_{rms} = \sqrt{\frac{\int (f - f_g)^2 S_\Delta(f) df}{\int S_\Delta(f) df}}, \quad (1.17)$$

où  $S_\Delta(f)$  est défini à l'équation (1.14) et  $f_g = \frac{\int f S_\Delta(f) df}{\int S_\Delta(f) df}$ . On définit alors le temps de cohérence par

$$T_{coh} = \frac{1}{D_{rms}}. \quad (1.18)$$

Ce temps correspond à la durée moyenne pour laquelle le canal peut être considéré comme non-sélectif en temps<sup>2</sup>. Les canaux de communications ASM typiques présentent des temps de cohérence de l'ordre de quelques dizaines à quelques centaines de millisecondes, voire de l'ordre de la seconde pour les canaux les plus stables [van Walree, 2011].

---

2. Le temps de cohérence est à distinguer du temps pour lequel le canal est considéré stationnaire

## 1.3 Le canal acoustique sous-marin comme canal d'interférence

Contrairement aux canaux de communications terrestres, l'activité temporelle et fréquentielle des sources acoustiques peuplant le canal ASM n'est pas réglementée. De nombreuses entités hétérogènes transmettent leur signaux dans des bandes de fréquences similaires. De plus, il n'existe pas d'infrastructure de réseaux pré-établie ou de protocoles standardisés et communément accepté permettant d'organiser et de contrôler l'accès aux ressources physiques du canal. Lorsque que plusieurs sources de signaux acoustiques cohabitent dans une même zone géographique, une congestion du spectre ASM peut vraisemblablement se produire, et se produira d'autant plus souvent que les activités humaines liées au monde de la mer se développent [McCarthy, 2004; NAS, 1970].

Jusqu'en avril 2017, le seul standard existant pour les systèmes de communication était le standard STANAG 1074 de l'OTAN pour les téléphones sous-marins utilisant des modulations analogiques à bande latérale unique. Depuis, le protocole de communications numériques JANUS a été standardisé par l'OTAN [NATO, 2017]. Ce protocole a pour but de permettre l'inter-opérabilité entre systèmes ASM hétérogènes en leur offrant un cadre pour communiquer grâce à la définition d'une modulation, d'un codage, et d'une mise en trame particulière. La couche physique mise en œuvre dans JANUS a été conçue afin d'assurer d'une part, une certaine robustesse des liens de communications, et d'autre part une intégration facile dans les modems commerciaux existants [Potter et al., 2014]. Dans cette première version, le standard propose une modulation à étalement de spectre par saut de fréquences (FHSS-FSK) à 13 paires de tons réparties sur une bande de 4160 Hz centrée sur 11520 Hz. Le protocole de la couche accès au média, basé sur l'écoute de la porteuse (CSMA, pour carrier sense multiple access), est également orienté vers la simplicité d'intégration et de mise en œuvre. Il est cependant limité en terme de nombre de sources pouvant cohabiter sans collisions et sensible aux longs délais de propagation typiques du canal ASM (étant donnée la faible célérité de l'onde dans le milieu).

Bien que la standardisation de JANUS marque vraisemblablement le début d'une nouvelle période dans le développement des communications ASM, le protocole n'apporte pas en lui-même toutes les réponses au problème de la coexistence entre sources ASM hétérogènes. Cela est évident par exemple pour les sources biologiques. D'autre part, la modulation proposée par ce protocole est d'une applicabilité restreinte aux applications bas débit ou ne nécessitant que l'échange sporadique de petites quantités de données. Il en va de même pour la méthode d'accès au médium CSMA, pour laquelle le nombre de source pouvant cohabiter sans collision est limité. Il est connu que de longs délais de propagation mettent facilement en défaut ce type de méthode, car, selon les distances de transmission entre les utilisateurs, l'écoute du canal avant émission peut être réalisée en retard ou en avance par rapport au signaux qui vont entrer en collision au récepteur destinataire des messages. Éventuellement, JANUS pourrait être utilisé pour permettre aux utilisateurs d'établir eux-même un partage de ressources équitable, préalablement à leurs transmissions respectives. Cependant, les délais de propagations peuvent être problématiques dans un cadre distribué, où des échanges

répétés de messages entre les utilisateurs sont nécessaires pour diffuser l'information sur les états du canal perçus par chacun. La centralisation de ce partage au niveau d'un seul récepteur qui enverrait ensuite des directives aux différents émetteurs du réseau est quant à elle d'une applicabilité restreinte, tous les réseaux n'étant pas centralisés par nature. Enfin, JANUS peut souffrir d'un manque de capacité d'adaptation, au sens où la variabilité du canal sous-marin n'est pas prise en ligne de compte explicitement.

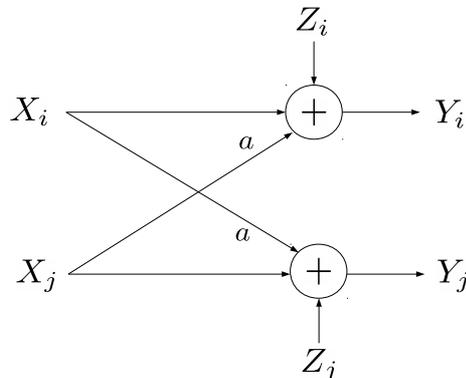


FIGURE 1.8 – Canal d'interférence gaussien scalaire à deux utilisateurs.

D'une manière très générale, les communications ASM peuvent être considérées comme non-coopératives : l'émission d'ondes ASM peut se faire sans grande restriction, l'interopérabilité entre systèmes d'origines différentes est quasiment inexistante à l'heure actuelle, et des interférences sont provoquées et subies inconsciemment dès lors que plusieurs sources entrent en compétition pour l'utilisation de la ressource physique offerte par le canal. Il est par conséquent très difficile d'aboutir à un partage efficace de cette ressource.

Du point de vue de la théorie de l'information, les modèles de communications sous-jacents aux travaux de cette thèse entrent dans la classe des *canaux d'interférences*. La figure 1.8 représente schématiquement le canal d'interférence gaussien pour deux couples émetteur-récepteur. Ce modèle formalise des situation où  $M$  couples émetteur-récepteur partagent un même canal et où chaque récepteur traite l'interférence des émetteurs qui ne lui sont pas appairés comme du bruit additif. La théorie de l'information s'intéresse à l'étude de la *région de capacité* lorsqu'elle se positionne dans le cadre des communications multi-utilisateurs et des réseaux. Cette région désigne ici l'ensemble des  $M$ -tuples de débits atteignables simultanément par les  $M$  utilisateurs du canal d'interférence lorsque ceux-ci optimisent conjointement leurs stratégies de transmission. Dans le cas général, seules des bornes de cette région sont connues [Motahari and Khandani, 2009]. La détermination de stratégies optimales de transmission pour l'atteindre reste également un problème ouvert. Dans le cadre du partage de spectre, il est possible de montrer que lorsque l'interférence mutuelle est suffisamment forte, le partage en sous-bandes orthogonales (FDMA) est optimal pour la somme des débits atteignables [Hayashi and Luo, 2009; Zhao and Pottie, 2013]. De plus, sous ces conditions, l'allocation FDMA permet d'atteindre tous les points de la région de capacité tels qu'aucun utilisateur ne peut obtenir un débit plus grand sans que celui d'au moins un autre utilisateur diminue<sup>3</sup>. Néanmoins, la définition de la région de capacité

3. Les allocations de puissance FDMA sont ici *optimales au sens de Pareto*, voir les définitions données au chapitre 2, section 2.1.6.

sous-tend une forme de coopération entre les utilisateurs afin que ceux-ci s'accordent pour mettre en œuvre les stratégies de transmission optimales. Lorsque cette entente est impossible ou que certains utilisateurs adoptent un comportement égoïste de maximisation de leur seule performance individuelle, un concept plus adapté est nécessaire. Les outils de la théorie des jeux peuvent alors compléter ceux de la théorie de l'information, notamment la notion d'équilibre de Nash (voir au chapitre 2), grâce à laquelle la *région d'équilibre de Nash* peut être définie : celle correspond à tous les débits atteignables simultanément par les utilisateurs du canal d'interférence lorsque chacun d'entre eux adopte une stratégie d'optimisation individuelle de sa propre performance [Berry and Tse, 2011].

Ce manuscrit ne se consacre pas à l'étude des méthodes d'allocation de ressources dans le canal d'interférence ASM sous le point de vue de la théorie de l'information, mais plutôt à identifier les outils permettant de mettre en œuvre des stratégies pragmatiques vis-à-vis des hypothèses que le contexte impose. Néanmoins, les conclusions tirées chapitre 3 pourront être analysées à la lumière des quelques travaux en théorie de l'information qui viennent d'être évoqués ci-dessus.

## 1.4 Adaptation au canal et gestion de l'interférence

Les paragraphes précédents ont exposé les caractéristiques physiques du canal ASM et les limitations qu'elles imposent aux systèmes de communications. L'environnement sous-marin variant dans le temps exige une certaine forme d'adaptation des émetteurs et des récepteurs ASM afin d'assurer des communications fiables et d'une qualité aussi constante que possible. Cette capacité d'adaptation est d'autant plus nécessaire lorsque le canal est partagé. En effet, les faibles largeurs de bandes passantes et les longs délais de propagations rendent difficile la cohabitation entre sources acoustiques hétérogènes. En général, ces sources, évoluant en dehors d'une structure de réseau et ne disposant pas d'un protocole de communication commun, sont rarement en mesure de se concerter de manière autonome pour établir un schéma équitable de partage de la ressource. Ceci ne peut être vérifié que dans le cas particulier d'entités mises en réseau par le même concepteur pour une application spécifique. Dans le cas contraire, les sources en activité simultanée sur une bande de fréquence commune entrent inévitablement en interférence les unes avec les autres, au détriment de leurs performances et de de leur fiabilité.

Dans cette section est présenté un bref état de l'art des communications ASM, où on se concentrera sur les techniques d'adaptation au canal, de gestion de l'interférence multi-utilisateurs et de mise en réseau en lien avec les questions qui viennent d'être soulevées.

### 1.4.1 Communications ASM point-à-point et adaptatives

Les années 1990 ont vu le début des recherches sur l'utilisation, en communications en ASM, des techniques de réception cohérentes basées sur l'estimation de la

phase et de l'amplitude des symboles transmis (PSK et QAM). Ces avancées avaient été anticipées dans les années 1980 comme conséquences de l'augmentation des puissances de calculs offertes par le développement des circuits intégrés [Baggeroer, 1984]. Avant cela, les communications ASM numériques se sont longtemps limitées aux modulations par sauts de fréquences (FSK) et aux récepteurs non-cohérents basés sur la détection d'énergie [Baggeroer, 1984; Kilfoyle and Baggeroer, 2000]. Ce type de couche physique a l'avantage d'être facilement réalisable avec des processeurs peu performants et de présenter une certaine robustesse face aux distorsions introduites par le canal, notamment les trajets multiples.

Les travaux de M. Stojanovic [Stojanovic et al., 1993, 1994] sont parmi les premiers à faire état de communications PSK cohérentes réalisées avec succès dans des canaux ASM, et sur de longues distances. Le récepteur proposé met en œuvre un égaliseur de canal adaptatif, nécessitant l'estimation régulière des paramètres du canal au moyen de symboles pilotes insérés au début des trames émises. Ce type de récepteur peut être considéré comme une première forme d'adaptation des liens de communication au canal ASM. De nombreuses recherches y ont donné suite pendant plus d'une décennie [Singer et al., 2009], aussi bien sur l'égalisation du canal ASM [Geller et al., 1995] et les récepteurs itératifs [Laot and LeBidan, 2011] que sur les techniques de communication à étalement de spectre [Stojanovic et al., 1998; van Walree et al., 2008b; Yang and Yang, 2008] ou l'exploitation de la diversité spatiale et fréquentielle dans les systèmes multi-porteuses [Bejjani and Belfiore, 1996; Berger et al., 2010; van Walree and Leus, 2009] et/ou multi-antennes [Li et al., 2009].

Les travaux qui viennent d'être évoqués se sont intéressés à l'acquisition et l'exploitation optimales de la connaissance de l'état du canal à la réception. Cependant, il est établi que la capacité des liens de communications est conditionnée par la connaissance du canal à la fois au récepteur et à l'émetteur [Biglieri et al., 1998], le cas le plus favorable étant celui où elle est disponible des deux côtés. D'un point de vue pratique, cela signifie que si l'émetteur connaît le canal alors il est capable d'optimiser de transmissions vis-à-vis du débit ou de la robustesse de la communication. Un exemple simple connu en théorie de l'information est celui de l'allocation de puissance par « water-filling » dans des canaux parallèles à bruit gaussien coloré [Cover and Thomas, 2006]. Dans cet exemple, l'émetteur exploite sa connaissance de la matrice de covariance du bruit pour allouer, sous contrainte de puissance moyenne, d'autant plus de puissance sur un canal que son niveau de bruit est faible. Cette allocation est optimale pour la capacité. Enfin, disposer de connaissances sur le canal à l'émetteur peut permettre de mettre en place des stratégies d'adaptation aux variations temporelles de son état ou de ses statistiques.

Parmi les premiers résultats de recherches sur l'adaptation de lien à l'émetteur en communication ASM, on notera ceux publiés par Qarabaqi et Stojanovic [Qarabaqi and Stojanovic, 2011a] où un contrôle de puissance adapté aux variations lentes des statistiques du canal est proposé. Le but est ici d'assurer un certain rapport signal à bruit à l'entrée du récepteur. L'information sur le canal est estimée par le récepteur et transmise à l'émetteur via un lien de retour, affecté d'un certain délai de propagation fonction de la distance séparant les terminaux. Afin de prendre en compte ce délai, les estimations du gain du canal sont utilisées afin de former une prédiction avec une avance temporelle d'au moins le délai de propagation. La faisabilité d'une telle

méthode est discutée dans [Radosevic et al., 2011]. Sur cette base, [Radosevic et al., 2014] proposent deux algorithmes d'adaptation de l'ordre de modulation et de la puissance alloués sur les sous-porteuses d'une modulation OFDM, dans le but de maximiser le débit atteignable sous contrainte de taux d'erreurs binaires (TEB) et de puissance moyenne. Les débits atteignables dans des canaux ASM pour différentes stratégies d'allocation de puissance dans les systèmes OFDM sont étudiées par [Aval et al., 2015]. La connaissance du canal à l'émetteur est supposée acquise de manière analogue à [Qarabaqi and Stojanovic, 2011a; Radosevic et al., 2011, 2014]. Ce travail étudie les bénéfices pouvant être apportés par une adaptation de l'allocation de puissance de l'émetteur aux évanouissements rapides du canal. Du fait des imperfections de prédiction du canal, des longs délais de propagation et des faibles temps de cohérence des évanouissements rapides, ce travail conclut qu'une allocation de type « water-filling » basée sur les prédictions des réalisations instantanées du canal n'apporte pas de gain significatifs. Ainsi, allouer uniformément la puissance sur la bande utile reste une stratégie satisfaisante d'un point de vue pratique et simple à mettre en œuvre. Il sera vu au chapitre 3 qu'une information statistique sur l'état du canal - plutôt qu'une prédiction de ses réalisations - permet de contourner les difficultés relevées par [Aval et al., 2015] dans la mise en œuvre d'une allocation de puissance adaptative pour les systèmes OFDM.

### 1.4.2 Communications multi-utilisateurs, accès cognitif et gestion de l'interférence

La faisabilité de communications point-à-point cohérentes démontrée dans [Stojanovic et al., 1994] et leurs extension à des récepteurs multi-antennes [Stojanovic et al., 1993] ont conduit peu à peu la communauté des communications ASM à s'intéresser aux communications multi-utilisateurs, et notamment à la possibilité pour un même récepteur de décoder des messages d'émetteurs différents grâce à la diversité spatiale. Il ne s'agit pas ici de réseaux, mais de configurations où un récepteur unique est capable de décoder les signaux d'émetteurs différents, ces derniers n'étant pas en mesure de communiquer entre eux et se trouvant éventuellement en concurrence pour l'accès au récepteur. Cet accès est géré au niveau de la couche physique.

Les résultats présentés dans [Stojanovic and Zvonar, 1996] ont initié le développement de recherches sur des techniques de réception multi-utilisateurs variées. C'est par exemple le cas des travaux sur l'annulation successive d'interférences multi-utilisateurs présentés par [Cho et al., 2011, 2012; Huang et al., 2013]. Afin de répondre aux difficultés de synchroniser les émetteurs dans le contexte des communications ASM multi-utilisateurs, [Wang et al., 2013] étudient le problème de la réception asynchrone. Les différentes méthodes citées ici pré-supposent une réception centralisée au niveau d'un seul récepteur ayant de forts a priori sur les signaux transmis. Outre la connaissance des paramètres élémentaires de la transmission, tels que la rapidité de modulation ou la fréquence porteuse, cet a priori prend notamment la forme de la connaissance de signaux pilotes permettant d'estimer le canal de chaque utilisateur et de distinguer son signal des autres (à condition que les rapports signaux sur interférence plus bruit (RSIB) soient suffisants).

Les débuts des recherches sur la radio cognitive [Mitola, 2000] ont inspiré quelques travaux en ASM. On peut notamment citer [Baldo et al., 2008], qui pointent des similarités entre les canaux radios et les canaux ASM au sens où les ressources sont limitées, et où leur disponibilité dépend de l'utilisation qu'en font d'autres systèmes présents dans le canal. On s'attend ainsi à ce que les méthodes d'accès dynamique au spectre développées pour la radio soient efficaces sous l'eau. L'algorithme d'allocation de spectre proposé est mis en œuvre aux moyens d'un contrôleur centralisé. Le modèle de canal considéré se base sur les modèles simples d'atténuation, présentés ici à la Section 1.2.1, et néglige la sélectivité fréquentielle et la variabilité de l'environnement. Biagi *et al.* notent cependant une différence majeure entre le canal ASM et les canaux radio, du point de l'accès cognitif au spectre. Celle-ci réside dans le fait que, dans la plupart des cas, les systèmes de radio-communications ne subissent d'interférences que de la part de signaux de communications, tandis que le canal ASM est peuplé de sources acoustiques de natures variées, incluant les mammifères notamment. Une méthode d'accès cognitif aux ressources du canal prenant en compte, dans une certaine mesure, la diversité des sources ASM est proposée dans [Biagi et al., 2016]. Les interférences provenant de signaux de communications sont différenciées de celles provoquées par d'autres sources en se basant sur un apprentissage automatique de la signature acoustique de ces sources. L'apprentissage est réalisé sur des données d'entraînement qui contiennent, entre autres, des vocalises de mammifères, des bruits de bateaux et signaux de communications OFDM. Les ressources du canal sont divisées en blocs temps-fréquence. L'algorithme d'allocation des ressources décide de ne pas transmettre sur un bloc temps-fréquence seulement lorsqu'un signal de communication y est détecté. Cette méthode, discutable sur le plan de la coexistence entre sources acoustiques humaines et mammifères marins, se base donc sur une forte connaissance a priori des interférences pouvant être rencontrées.

La gestion de l'interférence issue de sources autres que des signaux de communication a été étudiée dans [McGee et al., 2014; Wang et al., 2012]. L'algorithme d'annulation d'interférences proposé repose sur la connaissance a priori de la durée et de la bande de fréquence. L'interférence est représentée de manière paramétrique par ses coefficients de Fourier. L'algorithme utilise un test de rapport de vraisemblance généralisé afin de détecter la présence de l'interférence tout en estimant les valeurs de ses coefficients, puis l'annule si elle est présente.

### 1.4.3 Réseaux sous marins

L'utilisation des modulations cohérentes et des techniques d'égalisation adaptatives ont permis un meilleur usage du spectre ASM que ne le permettaient les modulations non-cohérentes comme la FSK et ainsi, les débits atteints par les modems ont pu être augmentés. Les premières études sur les réseaux ASM ont émergées à la fin des années 1990, l'accent étant mis sur l'étude des méthodes d'accès multiple et les protocoles d'accès au canal [Kilfoyle and Baggeroer, 2000; Sozer et al., 2000]. Encore aujourd'hui, mettre en place des réseaux sous-marins fiables et performants pose de nombreuses difficultés, et il existe assez peu d'applications pratiques des solutions apportées par les études de la recherche académique.

Des états de l'art spécifiquement centrés sur la mise en réseau de systèmes de communication ASM peuvent être trouvés dans [Kilfoyle and Baggeroer, 2000; Sozer et al., 2000; Stojanovic et al., 2008] ainsi que dans le livre de [Otnes et al., 2012]. Les techniques d'accès multiple et les protocoles étudiés jusqu'à maintenant sont souvent statiques et dédiés à des systèmes opérant au sein d'un réseau homogène. Les problématiques d'interférences « externes » ne sont pas prises en compte. La vaste majorité des techniques répertoriées proviennent encore des communications terrestres, ou en sont des variantes plus ou moins bien adaptées aux spécificités du canal ASM. Les difficultés de leur mise en œuvre proviennent essentiellement des limites physiques du canal ASM en termes de délai de propagation et de bande passante. La plupart des travaux présentés ici et au paragraphe 1.4.2 ci-dessus se basent sur une connaissance plus ou moins détaillée de l'interférence et un contexte centralisé. Ces hypothèses sont assez restrictives, et valides seulement dans des contextes particuliers. Sans réglementation ni protocole permettant d'obtenir de l'information de la part de la source interférente, acquérir une connaissance suffisamment détaillée pour pouvoir l'annuler *in situ* et « en aveugle » est chose difficile.

## 1.5 Conclusion

Le canal de communication ASM est souvent présenté comme difficile, en raison de ses particularités physiques et des hétérogénéités du milieu. Ces difficultés sont une des raisons pour lesquelles le développement des communications ASM a été plus lent que celui des communications radiofréquences, car les techniques employées dans ce domaine ne peuvent se transposer directement au milieu sous-marin. La nature large bande du canal, les différentes échelles de variabilité du milieu, ainsi que l'absence de protocoles et d'infrastructures réseaux standardisées pour organiser l'accès au canal posent de nouvelles difficultés.

En dépit des efforts récents de standardisation de protocoles de communication, la gestion de l'interférence multi-utilisateur et le partage efficace de la ressource reste un problème actuel, qui nécessite des solutions appropriées aux conditions dans lesquelles les systèmes de communication ASM évoluent. Quelle que soit la nature des sources d'interférences, la situation d'accès anarchique au canal ASM fait entrer les systèmes de communications dans une compétition pour l'utilisation de la ressource. La suite de cette thèse entend proposer des solutions raisonnables et pragmatiques pour résoudre cette compétition aux moyens d'outils issus de la théorie des jeux et de l'apprentissage automatique. Le soin sera pris d'éviter le plus possible de poser des hypothèses qui contreviennent à ce qui a été exposé dans ce chapitre. Le chapitre suivant est consacré aux notions de base de la théorie des jeux, qui nous servira d'outil de formalisation de problèmes plus concrets traités aux chapitres 3 et 4.



# Chapitre 2

## Notions de théorie des jeux

La théorie des jeux est une branche des mathématiques dont l'objet d'étude concerne les interactions entre agents visant à l'optimisation d'objectifs interdépendants et, éventuellement, contradictoires. Un *jeu* est décrit comme l'ensemble de ces agents munis chacun d'une fonction objectif et de stratégies représentant les variables de ces fonctions.

L'origine de cette théorie remonte au XIX<sup>ème</sup> siècle avec les travaux de Cournot sur les *Principes mathématiques de la théorie des richesses*. Il y expose le modèle connu aujourd'hui sous le nom de duopole de Cournot et qui entend décrire, de manière abstraite et simplifiée, la compétition de deux producteurs sur un marché vis-à-vis des volumes et des prix de vente de leur production [Cournot, 1838]. Bien qu'il n'y soit pas encore fait mention de « jeux » ni d'« agents », Cournot y propose la notion d'un *équilibre* entre les volumes et les prix de vente des deux producteurs, déterminé en fonction des conditions initiales du marché. Ce sont les situations d'équilibre, leurs propriétés, et leur apprentissage par les agents qui intéresseront la théorie des jeux par la suite. Les travaux de von Neumann et Morgenstern sur la résolution des jeux à somme nulle [von Neumann and Morgenstern, 1944], ainsi que le développement, par John Nash, du concept d'équilibre portant son nom [Nash, 1951], ont été fondateurs de la théorie des jeux moderne qui est devenu un champs de recherche à part entière. Dans un premier temps, ses outils mathématiques ont surtout servi à l'analyse économique, et souvent dans des situations simples et idéalisées telles que le duopole Cournot. Ils ont ensuite touché des champs d'applications divers et variés, allant de la biologie [Maynard-Smith and Price, 1973; Vincent and Brown, 2005] aux systèmes de communications numériques [Han et al., 2012; Lasaulce and Tembine, 2011], en passant par l'étude théorique de la course à l'armement durant la Guerre Froide [Aumann and Maschler, 1995]. La théorie des jeux se révèle également utile dans le domaine de l'apprentissage automatique dans le contexte des systèmes multi-agents [Leyton-Brown and Shoham, 2009].

Ce chapitre a pour objectif de présenter les notions de bases de la théorie des jeux qui permettront une meilleure compréhension des autres chapitres de ce manuscrit. Seuls les jeux non-coopératifs sont étudiés ici et on ne présentera que les jeux sous forme normale et les jeux de satisfaction. Bien qu'elles ne soient pas présentées dans cette

thèse, il convient de mentionner qu'il existe d'autres formes de jeux non-coopératifs : jeux bayésiens, jeux répétés, stochastiques et jeux évolutionnaires. Celles-ci servent principalement à raffiner le concept d'équilibre de Nash, ce dont nous n'aurons pas besoin ici étant donnés les objectifs de cette thèse. Les jeux sous forme stratégique seront étudiés dans un premier temps afin de présenter les concepts de base communs à toutes les autres formes de jeux : actions, stratégies, fonction d'utilité et équilibre. Ces notions seront principalement utilisées dans le chapitre 3. Les jeux de satisfaction, qui sont une forme particulière de jeux non-coopératifs, seront présentés en deuxième partie de ce chapitre pour les besoins du chapitre 4. Il sera expliqué comment résoudre un jeu, à travers des exemples simples dont la littérature sur la théorie des jeux fait souvent usage. On essaiera également, lorsque cela est utile, de faire le lien avec les sujets qui seront traités dans la suite.

## 2.1 Jeux sous forme stratégique

### 2.1.1 Définition et exemples introductifs

**Définition 2.1.1** (Jeu sous forme normale). *Un jeu sous forme stratégique (ou normale) à  $I$  joueurs est un triplet*

$$\mathcal{G} = (\mathcal{I}, \mathcal{A}, u) \tag{2.1}$$

où

- $\mathcal{I} = \{1, \dots, I\}$  est l'ensemble des joueurs du jeu, aussi appelés agents,
- $\mathcal{A} = \mathcal{A}_1 \times \dots \times \mathcal{A}_I$  est l'ensemble des actions jouables simultanément par les joueurs, où  $\mathcal{A}_i$ , est l'ensemble des actions jouables par le joueur  $i \in \mathcal{I}$ .
- $u : \mathcal{A} \rightarrow \mathbb{R}^I$  où la  $i^{\text{ème}}$  composante  $u_i : \mathcal{A} \rightarrow \mathbb{R}$  est la fonction d'utilité du joueur  $i$ , et on note  $u_i(a)$  son degré de satisfaction vis-à-vis du profil d'action  $a \in \mathcal{A}$ .

Dans le contexte des télécommunications, les joueurs peuvent être un ensemble d'émetteurs ou de couples émetteur-récepteur, ayant des actions telles que la puissance d'émission, le schéma de codage et de modulation ou la bande de fréquence et les fonctions d'utilités seraient définies par des métriques de performances telles que la capacité, la probabilité d'erreur, l'énergie du signal émis, ou le rapport signal à bruit.

Les *profils d'actions*  $a = (a_1, a_2, \dots, a_I)$  représentent des choix d'actions simultanés faits par les joueurs. Du point de vue d'un joueur  $i \in \mathcal{I}$  donné, on notera  $a = (a_i, a_{-i})$  où  $a_i \in \mathcal{A}_i$  est l'action de ce joueur et  $a_{-i} \in \mathcal{A}_{-i}$ , avec  $\mathcal{A}_{-i} = \mathcal{A} \setminus \mathcal{A}_i$ , représente les actions choisies par tous les joueurs excepté  $i$ . Par exemple, si  $i$  n'est ni le joueur 1 ni le joueur  $I$ , on aura  $a_{-i} = (a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_I)$ .

**Définition 2.1.2** (Jeu fini). *On dit d'un jeu qu'il est fini si et seulement si tous les joueurs disposent d'un nombre fini d'alternatives. Autrement dit, l'ensemble  $\mathcal{A}$  est fini et dénombrable. Au contraire, dans un jeu infini il existe au moins un joueur dont l'ensemble des actions est infini (dénombrable ou non).*

Les exemples didactiques dont la théorie des jeux fait usage pour illustrer ses notions de base sont souvent des jeux finis, à 2 ou 3 joueurs et à peu d'actions par joueur, de sorte que les jeux soient représentables facilement par des matrices [Fudenberg and Tirole, 1991; Osborne and Rubinstein, 1994]. Un classique est celui du « chifoumi » donné à la figure 2.1. Les lignes de la matrice représentent les stratégies disponibles au joueur 1 et les colonnes celles du joueur 2. Elle sont notées  $P$ ,  $F$  et  $C$  ici pour « Pierre », « Feuille » et « Ciseaux ». Chaque entrée de la matrice du jeu est un résultat possible en terme de choix d'actions réalisés par les joueurs avec les utilités associées, où, par convention, l'utilité du joueur 1 est à gauche et celle du joueur 2 à droite. Ainsi, lorsque le profil d'actions choisi est par exemple  $(P, F)$ , le joueur 1 reçoit un gain de  $-1$  et le joueur 2 de  $+1$ . Résumé avec la définition 2.1.1, on a  $\mathcal{I} = \{1, 2\}$ ,  $\mathcal{A}_i = \{P, F, C\}$ ,  $i = 1, 2$ , et les utilités  $u_i$  sont données dans la matrice pour chaque combinaison d'actions parmi  $\mathcal{A} = \{(P, P), (P, F), \dots, (C, C)\}$ . Le chifoumi est un exemple de jeu à somme nulle.

	$P$	$F$	$C$
$P$	0, 0	-1, 1	1, -1
$F$	1, -1	0, 0	-1, 1
$C$	-1, 1	1, -1	0, 0

FIGURE 2.1 – Matrice du jeu de « chifoumi ».

**Définition 2.1.3** (Jeu à somme nulle). *Un jeu est à somme nulle si et seulement si  $\forall a \in \mathcal{A}$ ,  $\sum_i u_i(a) = 0$ . Autrement dit, la somme des gains des joueurs est nulle pour tout résultat du jeu. Dans le cas contraire le jeu est à somme non-nulle.*

Un autre exemple est le jeu d'anti-coordination donné à la figure 2.2. On peut associer cette matrice à un problème simple de partage de spectre en communications numériques. Supposons que les actions  $S$  et  $B$  représentent des bandes de fréquences orthogonales et que les joueurs soient des émetteurs ayant le choix de la bande sur laquelle leur transmission s'opère. Les deux agents partagent une préférence pour l'une des deux bandes, et il est important qu'ils choisissent une bande différente l'un et l'autre pour que leurs communications n'interfèrent pas. Dans le cas contraire on considère que les communications ne peuvent pas s'établir et une utilité de 0 est ainsi attribuée aux deux joueurs lorsque leurs actions ne sont pas coordonnées.

Enfin, on donne le « dilemme du prisonnier » à la figure 2.3 qui est un exemple à la fois classique et important en théorie des jeux. Deux prisonniers supposés complices

	$B$	$S$
$B$	0, 0	2, 1
$S$	1, 2	0, 0

FIGURE 2.2 – Matrice d'un jeu d'anti-coordination.

	$C$	$D$
$C$	$-1, -1$	$-4, 0$
$D$	$0, -4$	$-3, -3$

FIGURE 2.3 – Matrice du jeu « dilemme du prisonnier ».

	$C$	$D$
$C$	$R, R$	$T, N$
$D$	$N, T$	$P, P$

FIGURE 2.4 – Matrice d'un dilemme social.

d'un crime dont ils sont accusés font face aux choix suivants : si l'un des prisonniers dénonce son complice ( $a_i = D$ ) tandis que ce dernier se tait ( $a_j = C$ ), il est libéré ( $u_i = 0$ ) et son complice est condamné à 4 ans de prison ( $u_j = -4$ ). Si les deux prisonniers coopèrent et se taisent ( $a_i = a_j = C$ ), ils sont condamnés à 1 an de prison ( $u_i = u_j = -1$ ). Si les deux prisonniers se dénoncent ( $a_i = a_j = D$ ), ils sont tous deux condamnés à 3 ans de prison ( $u_i = u_j = -3$ ).

Une version de ce dernier jeu peut être donné un contexte de communication numériques [Bacci et al., 2016]. Supposons que les deux joueurs soient des systèmes de radio cognitive en activité sur la même portion de spectre, et ayant chacun deux stratégies  $C$  et  $D$  correspondant respectivement à transmettre en bande étroite ou en large bande. Posons également l'hypothèse que les transmissions en bande étroite s'établissent sur deux bandes orthogonales, une étant attribuée à chaque joueur. Les utilités peuvent être rapportées à des débits de transmission. Ainsi, lorsque les deux systèmes ont la possibilité de partager le spectre de manière orthogonale en jouant la stratégie  $(C, C)$ , les débits sont meilleurs pour chacun car les transmissions n'interfèrent pas. Au contraire, les débits sont faibles pour la stratégie  $(D, D)$  car les deux systèmes interfèrent l'un sur l'autre. Dans le cas où un joueur émet en bande étroite tandis que l'autre émet en large bande, ses performances sont mauvaises du fait à la fois de la petite bande de transmission et de l'interférence subie.

Le dilemme du prisonnier peut être généralisé par la matrice de la figure 2.4 représentant ce que [Leibo et al., 2017] nomment un dilemme social. Ce genre de jeu modélise des situations où deux agents (ou deux entités plus abstraites) se trouvent face au choix de la coopération ( $C$ ) avec l'autre agent ou de l'individualisme ( $D$ ). La coopération est le résultat bénéfique pour tous mais son occurrence est mise en péril par des intérêts individuels vers lesquels les agents peuvent être tentés de converger. Les caractéristiques générales des dilemmes sociaux peuvent être résumées par les conditions suivantes :

- $R > P$ . La coopération mutuelle est plus intéressante que l'individualisme mutuel.
- $R > N$ . La coopération mutuelle est plus intéressante que la coopération unilatérale.
- $2 \times R > T + N$  de sorte que la coopération mutuelle est toujours plus intéressante que la coopération unilatérale et l'individualisme unilatéral jouées à égales probabilités.
- — Ou bien  $T > R$ , l'individualisme pour exploiter une coopération unilatérale de l'adversaire est individuellement plus intéressant que la coopération mutuelle.

- Ou bien  $P > N$ , l'individualisme mutuel est individuellement préférable à se laisser exploiter.

Dans le dilemme du prisonnier, les deux dernières conditions sont vérifiées ensemble. On expliquera mieux le caractère de dilemme de ce jeu dans les paragraphes suivants.

### 2.1.2 Stratégies pures et randomisées, utilité espérée

Il convient dans un premier temps de faire la distinction entre la notion d'*action*, qui est l'objet résultant d'un choix stratégique d'un joueur, et la *stratégie*, qui correspond à la manière dont ce choix est réalisé. Reprenons pour commencer l'exemple du jeu de chifoumi. L'expérience de chacun et l'observation de la matrice de la figure 2.1 montrent qu'un joueur recherchant une issue victorieuse à ce jeu doit nécessairement faire en sorte que ses choix d'actions paraissent aléatoires pour son adversaire. Ceci nous amène à la notion de *stratégie randomisée*.

**Définition 2.1.4** (Stratégie randomisée). *Une stratégie randomisée du joueur  $i$  est un choix de distribution de probabilités sur son ensemble de stratégies  $\mathcal{A}_i$  que l'on note  $\pi_i \in \Delta\mathcal{A}_i$ . Le profil de stratégies randomisées est noté  $\pi = (\pi_1, \dots, \pi_I) \in \Delta\mathcal{A}$ .*

Soit  $(\Omega, \mathcal{F}, \mathbb{P})$  l'espace probabilisé usuel. On note  $\mathcal{P}(\mathcal{A}_i)$  l'ensemble des parties de  $\mathcal{A}_i$ . De cette façon, le choix d'action résultant de la stratégies du joueur  $i$  est modélisable par une variable aléatoire

$$A_i : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (\mathcal{A}_i, \mathcal{P}(\mathcal{A}_i), \pi_i) \quad (2.2)$$

prenant ses valeurs dans l'espace  $\mathcal{A}_i$  avec la mesure de probabilité  $\pi_i$ . On notera simplement  $A_i \sim \pi_i$  dans la suite. Un profil de stratégies randomisées est alors la mesure de probabilité  $\pi : \mathcal{P}(\mathcal{A}) \rightarrow [0, 1]$ , distribuant conjointement les actions  $A = (A_1, \dots, A_I) \sim \pi$ . En toute généralité, les stratégies randomisées des différents joueurs peuvent être corrélées les unes aux autres. Un profil de *stratégies mixtes* est tel que les stratégies des joueurs sont indépendantes. Il est bien sur possible de choisir une action de façon déterministe, ce qui est défini par la notion de *stratégie pure*. Ainsi, l'ensemble des stratégies pures de chaque joueur s'identifie à son espace d'actions  $\mathcal{A}_i$ .

**Définition 2.1.5** (Stratégie mixte). *Un profil de stratégies mixtes est un profil de stratégies randomisées  $\pi \in \Delta\mathcal{A}$  tel que les variables aléatoires  $A_1, A_2, \dots, A_I$  des choix d'actions des joueurs soient indépendantes.*

**Définition 2.1.6** (Stratégie pure). *Une stratégie pure du joueur  $i$ , notée  $a_i$  est une stratégie randomisée dégénérée telle que  $\pi_i(a_i) = 1$ , où  $a_i \in \mathcal{A}_i$ , et  $\pi_i(a'_i) = 0$  pour toute action  $a'_i \in \mathcal{A}_i$  telle que  $a'_i \neq a_i$ .*

Les matrices de jeux données au paragraphe précédent ne font état que des actions et des utilités déterministes associées. Dès lors que les joueurs ont la possibilité d'introduire de l'aléa par une stratégie randomisée  $\pi$ , le critère pertinent pour évaluer la satisfaction des joueurs vis-à-vis de cette stratégie est l'*utilité espérée*.

**Définition 2.1.7** (Utilité espérée d'une stratégie randomisée). *Étant donné un jeu sous forme normale  $\mathcal{G} = (\mathcal{I}, \mathcal{A}, u)$ , l'utilité espérée du joueur  $i$  pour un profil de stratégies randomisées  $\pi \in \Delta \mathcal{A}$  est définie par*

$$\begin{aligned} u_i(\pi) &\triangleq \mathbb{E}_{A \sim \pi} [u_i(A)] \\ &= \int_a u_i(a) d\pi(a) \end{aligned} \tag{2.3}$$

Dans le cas particulier d'une stratégie mixte, on a donc

$$u_i(\pi) = \int_a u_i(a) \prod_{j=1}^I d\pi_j(a_j). \tag{2.4}$$

### 2.1.3 Jouer de manière rationnelle

Le but principal de la théorie des jeux est de formuler des prédictions quant à la manière dont un jeu devrait être joué. Ces prédictions consistent à déduire, en fonction des règles du jeu (nombre de joueurs, actions et utilités), quelles sont les stratégies d'équilibre des joueurs, où l'équilibre doit être compris comme une forme de stabilité : une stratégie d'équilibre est jouée car aucun joueur n'a intérêt à dévier unilatéralement de celle-ci. La recherche des stratégies d'équilibre se fait à la lumière des concepts de connaissance commune, de rationalité et de stratégie dominée ou de meilleure réponse.

Les joueurs sont supposés avoir la *connaissance commune* du jeu  $\mathcal{G}$  auquel ils jouent. Cela signifie que tous les agents connaissent  $\mathcal{G}$ , et que tout les agents savent que tous les agents connaissent  $\mathcal{G}$ , et que tous savent que tous savent que tous les agents connaissent  $\mathcal{G}$ , ..., et ainsi de suite *ad infinitum*. Il existe des définitions mathématiquement formalisées de la notion de connaissance commune. Le lecteur intéressé peut se reporter au livre [Fudenberg and Tirole, 1991]. Sous ces hypothèses, les agents du jeu sont supposés agir de manière *rationnelle*, ce qui signifie qu'en toutes circonstances un joueur ne joue pas de stratégie dominée.

**Définition 2.1.8** (Stratégie dominée). *Soit un jeu sous forme normale  $\mathcal{G} = (\mathcal{I}, \mathcal{A}, u)$ . La stratégie pure  $a_i$  du joueur  $i$  est dite strictement dominée si il existe une stratégie randomisée  $\pi_i \in \Delta A_i$  telle que*

$$u_i(\pi_i, a_{-i}) > u_i(a_i, a_{-i}) \quad \forall a_{-i} \in \mathcal{A}_{-i}. \tag{2.5}$$

*Elle est faiblement dominée si il existe  $\pi_i$  telle que (2.5) se vérifie avec une inégalité faible, et qu'elle est vérifiée avec une inégalité stricte pour au moins une stratégie  $a_{-i}$ .*

Comme pour la connaissance commune, il est possible de définir formellement la rationalité à partir de la notion de stratégie dominée et du processus d'élimination itérative des stratégies dominées pour chaque joueur [Fudenberg and Tirole, 1991]. Une autre alternative est d'utiliser le concept de meilleure réponse.

**Définition 2.1.9** (Meilleure réponse). *Soit un jeu sous forme normale  $\mathcal{G} = (\mathcal{I}, \mathcal{A}, u)$ . La meilleure réponse d'un joueur  $i$  au profil de stratégie adverse  $\pi_{-i}$  est toute stratégie*

appartenant à l'ensemble retourné par la correspondance  $BR_i : \Delta\mathcal{A}_{-i} \rightarrow \Delta\mathcal{A}_i$  définie  $\forall \pi_{-i} \in \Delta\mathcal{A}_{-i}$  par

$$BR_i(\pi_{-i}) = \{\pi_i \in \Delta\mathcal{A}_i \mid u_i(\pi_i, \pi_{-i}) \geq u_i(\pi'_i, \pi_{-i}), \forall \pi'_i \in \Delta\mathcal{A}_i\}. \quad (2.6)$$

Les stratégies qu'un joueur rationnel devrait jouer sont celles qui sont des meilleures réponses à ce que ce joueur croit ou observe des stratégies adverses. Par exemple dans le dilemme du prisonnier de la figure 2.3, il n'est pas rationnel de jouer  $C$  si un l'autre joue  $D$ . La meilleure réponse n'est pas nécessairement unique, et cela dépend de la fonction d'utilité du joueur évaluée pour le profil de stratégies de ses adversaires. Dès lors que la meilleure réponse ne renvoie pas une unique stratégie pure face à  $\pi_{-i}$ , il existe une infinité de stratégies possibles pour y répondre de manière optimale. En effet, supposons que l'ensemble des stratégies renvoyées par  $BR_i(\pi_{-i})$  soit constitué de deux éléments, disons  $\alpha_i$  et  $\beta_i$ . Le joueur doit être indifférent à ces deux stratégies, qui produisent nécessairement la même valeur d'utilité espérée vis-à-vis de la stratégie mixte  $\pi_{-i}$ . Notons  $v_i$  cette valeur. S'il en avait été autrement, la meilleure réponse aurait retourné la stratégie pour laquelle l'utilité espérée est la plus élevée. Maintenant, supposons que le joueur assigne une probabilité  $p_\alpha$  à  $\alpha_i$  et  $p_\beta$  à  $\beta_i$ . L'utilité espérée pour cette stratégie mixte est  $p_\alpha u_1(\alpha_i, \pi_{-i}) + p_\beta u_1(\beta_i, \pi_{-i}) = p_\alpha v_i + p_\beta v_i = v_i$ . Toute combinaison convexe des utilités pour les deux éléments de l'ensemble  $BR_i(\pi_{-i})$  est égale à  $v_i$ . Ainsi, on voit qu'il est possible d'assigner n'importe quelle probabilité à l'une ou l'autre de ces stratégies pour obtenir une nouvelle stratégie qui soit une meilleure réponse à  $\pi_{-i}$ , ce qui prouve que l'ensemble des meilleures réponses est infini dès lors qu'il n'est pas un singleton.

### 2.1.4 Équilibre de Nash

L'équilibre de Nash [Nash, 1951] est le concept de solution central de la théorie des jeux non-coopératifs. On peut le définir à l'aide des meilleures réponses.

**Définition 2.1.10** (Équilibre de Nash). *Un profil de stratégies  $\pi^* = (\pi_1^*, \pi_2^*, \dots, \pi_I^*)$  est un équilibre de Nash si, pour tout joueur  $i$ ,  $\pi^*$  est une meilleure réponse à  $\pi_{-i}^*$ .*

Selon que la meilleure réponse de chacun des agents est unique ou non, on peut distinguer deux types d'équilibres de Nash : les équilibres stricts et faibles.

**Définition 2.1.11** (Équilibre de Nash strict). *Un profil de stratégies  $\pi^*$  est un équilibre de Nash fort si, pour tout joueur  $i \in \mathcal{I}$ ,*

$$u_i(\pi_i^*, \pi_{-i}^*) > u_i(\pi_i, \pi_{-i}^*) \quad \forall \pi_i \in \Delta\mathcal{A}_i \quad (2.7)$$

**Définition 2.1.12** (Équilibre de Nash faible). *Un profil de stratégies  $\pi^*$  est un équilibre de Nash faible si, pour tout joueur  $i \in \mathcal{I}$ ,*

$$u_i(\pi_i^*, \pi_{-i}^*) \geq u_i(\pi_i, \pi_{-i}^*) \quad \forall \pi_i \in \Delta\mathcal{A}_i \quad (2.8)$$

Un équilibre de Nash en stratégies mixtes est alors nécessairement un équilibre faible (selon les mêmes arguments que ceux avancés au paragraphe précédent concernant les meilleures réponses dont l'ensemble est constitué de plusieurs éléments). Au

contraire, un équilibre strict est nécessairement un équilibre en stratégies pures bien qu'un équilibre en stratégies pures puisse être strict ou faible.

On a la propriété évidente suivante stipulant que la stratégie d'équilibre de chaque joueur est un point fixe de sa correspondance de meilleure réponse.

**Proposition 2.1.1.** *À tout équilibre de Nash  $\pi^* \in \Delta\mathcal{A}$ , les stratégies d'équilibre vérifient  $\forall i \in \mathcal{I}$*

$$\pi_i^* \in BR_i(\pi_{-i}^*). \quad (2.9)$$

En notant la correspondance

$$BR : \begin{cases} \Delta\mathcal{A} & \longrightarrow & \mathcal{P}(\Delta\mathcal{A}) \\ \pi & \longrightarrow & \begin{pmatrix} BR_1(\pi_{-1}) \\ \vdots \\ BR_I(\pi_{-I}) \end{pmatrix} \end{cases} \quad (2.10)$$

on peut identifier l'équilibre de Nash à un point fixe de  $BR : \pi^* \in BR(\pi^*)$ . Ceci est utile pour l'étude de la convergence d'algorithmes de calcul de l'équilibre de Nash basés sur les meilleures réponses.

**Théorème 2.1.1** ([Nash, 1951]). *Tout jeu fini possède au moins un équilibre en stratégies mixtes.*

*Démonstration.* Le théorème de Nash se prouve à l'aide du théorème du point-fixe de Kakutani pour les correspondances entre ensembles, qui donne les conditions suffisantes suivantes pour l'existence d'un point-fixe de  $BR$  :

1.  $\Delta\mathcal{A}$  est un sous-ensemble non-vide, compact et convexe d'un espace euclidien de dimension finie.
2.  $BR(\pi)$  est non-vide pour tout  $\pi$ .
3.  $BR(\pi)$  est convexe.
4.  $BR$  a un graphe clos : si la suite  $(\pi^n, \beta^n) \rightarrow (\pi, \beta)$ , où, pour tout  $n$ ,  $\beta^n \in BR(\pi^n)$ , alors  $\beta \in BR(\pi)$ .

Les conditions 1 à 3 sont simples à vérifier :  $\Delta\mathcal{A}$  respecte bien (1) en tant qu'ensemble de distribution de probabilités, et formant donc un simplexe inclus dans  $\mathbb{R}^I$ . Les fonctions d'utilités espérées  $u_i(\pi_i, \pi_{-i})$  sont des fonctions linéaires de  $\pi$ , donc continues en  $\pi_i$ . On peut ainsi en déduire que  $BR$  est toujours non-vide car les fonctions continues sur un compact atteignent toujours un maximum. La condition (3) est vérifiée car toute combinaison convexe de meilleures réponses est une meilleure réponse. La condition (4) se vérifie en conséquence de la continuité des  $u_i(\pi)$  vis-à-vis des stratégies mixtes  $\pi$ . Pour s'en convaincre, supposons  $(\pi^n, \beta^n) \rightarrow (\pi, \beta)$ , où, pour tout  $n$ ,  $\beta^n \in BR(\pi^n)$  et  $\beta \notin BR(\pi)$ . Il existe alors au moins un joueur  $i$  pour lequel  $\beta_i \notin BR_i(\pi_{-i})$ . Donc il existe pour ce joueur une stratégie mixte  $\beta'_i$  telle que

$$\exists \epsilon > 0 \quad u_i(\beta'_i, \pi_{-i}) > u_i(\beta_i, \pi_{-i}) + \epsilon. \quad (2.11)$$

Comme  $u_i$  est continue vis-à-vis des stratégies mixtes et que  $(\pi^n, \beta^n) \rightarrow (\pi, \beta)$  on a

$$\lim_{n \rightarrow \infty} u_i(\beta_i^n, \pi_{-i}^n) = u_i(\beta_i, \pi_{-i}) \quad (2.12)$$

$$\lim_{n \rightarrow \infty} u_i(\beta_i', \pi_{-i}^n) = u_i(\beta_i', \pi_{-i}), \quad (2.13)$$

et ainsi, pour  $n$  suffisamment grand, les inégalités suivantes se vérifient :

$$u_i(\beta_i', \pi_{-i}^n) > u_i(\beta_i', \pi_{-i}) - \frac{\epsilon}{2} > u_i(\beta_i, \pi_{-i}) + \frac{\epsilon}{2} > u_i(\beta_i^n, \pi_{-i}^n). \quad (2.14)$$

Ce qui signifie que  $\beta_i'$  constitue une meilleure réponse que  $\beta_i^n$  face  $\pi_{-i}^n$  et contredit l'hypothèse de départ  $\beta^n \in BR(\pi^n) \forall n$ .  $BR$  a donc un graphe clos.  $\square$

Le théorème 2.1.1 garantit l'existence d'un équilibre dans tout jeu dont les espaces d'actions sont finis. Cependant, rien n'est dit quant à l'existence d'un équilibre en stratégies pures. L'existence de tels équilibres est souvent la conséquence d'une structure de jeu particulière. C'est le cas par exemple des jeux modulaires [Fudenberg and Tirole, 1991, Chapitre 12], et des jeux potentiels [Monderer and Shapley, 1996]. En communications numériques, les espaces d'actions des systèmes de transmissions peuvent souvent être modélisés comme des espaces continus. Ce serait le cas, par exemple, d'actions identifiées à des puissances de transmissions. Le théorème ne traite pas ce cas, et, pour contourner la difficulté, on pourra toujours considérer une discrétisation des espaces d'actions si nécessaire. Néanmoins, le théorème de Debreu-Glicksberg-Fan garantit, sous certaines conditions, l'existence d'un équilibre de Nash en stratégies pures dans les jeux à espaces d'actions et utilités continus.

**Théorème 2.1.2.** *Soit le jeu  $\mathcal{G} = (\mathcal{I}, \mathcal{A}, u)$  où  $\mathcal{A} = \mathcal{A}_1 \times \dots \times \mathcal{A}_I$  est l'ensemble des profils d'actions, avec  $\forall i \in \mathcal{I}, \mathcal{A}_i \subseteq \mathbb{R}^{n_i}$ . Si les espaces  $\mathcal{A}_i$  sont compacts et convexes et que les utilités  $u_i(a_i, a_{-i})$  sont continues en  $a \in \mathcal{A}$  et quasi-concaves<sup>1</sup> en  $a_i \in \mathcal{A}_i$ , alors il existe un équilibre de Nash en stratégies pures pour  $\mathcal{G}$ .*

*Démonstration.* Ce théorème se démontre également à l'aide du théorème de Kakutani. Les espaces d'actions sont compacts, convexes et non-vides par hypothèses. La continuité des fonctions d'utilités implique que les meilleures réponses sont non-vides et ont un graphe clos, tandis que leur quasi-concavité implique que les meilleures réponses forment un ensemble convexe.  $\square$

En général, l'unicité de l'équilibre n'est pas garantie et il est souvent difficile d'obtenir une telle caractérisation. Seuls certains jeux présentant des propriétés bien précises possèdent un unique équilibre de Nash. Les jeux concaves au sens de Rosen [Rosen, 1965] ou les jeux résolubles par la suppression itérative des stratégies dominées, tels que le dilemme du prisonnier, en sont des exemples.

---

1. Une fonction  $f : \mathcal{E} \rightarrow \mathcal{F}$  est *quasi-concave* si  $\forall (x, y) \in \mathcal{E}^2$  et  $\forall \lambda \in [0, 1]$  on a  $f(\lambda x + (1-\lambda)y) \geq \min(f(x), f(y))$ .

### 2.1.4.1 Solutions d'équilibre des exemples de la section 2.1.1

Illustrons maintenant le concept d'équilibre de Nash grâce aux exemples de la section 2.1.1. De manière informelle, il est utile de se dire que l'équilibre de Nash est une stratégie de laquelle aucun joueur n'a intérêt de dévier unilatéralement, ou encore, que si tous les joueurs jouent un équilibre de Nash, aucun joueur ne pourra choisir d'autre stratégie sans voir son utilité diminuer.

On se concentre tout d'abord sur le jeu d'anti-coordination donné à la matrice 2.2. Rappelons que dans ce jeu, deux émetteurs choisissent sur quelle bande de fréquence émettre leur signaux parmi deux bandes orthogonales  $S$  ou  $B$ . On constate que ce jeu possède deux équilibres de Nash en stratégies pures,  $(B, S)$  et  $(S, B)$ . Il possède également un équilibre en stratégies mixtes, que l'on peut déterminer simplement comme suit. Choisissons le joueur 2 et posons  $p = \mathbb{P}[a_2 = B]$  et  $1 - p = \mathbb{P}[a_2 = S]$ . Étant donné que l'on joue un équilibre mixte, le joueur 2 doit faire en sorte que le joueur 1 soit indifférent quant à ces choix d'actions - en d'autres termes, les deux actions doivent produire la même utilité espérée par rapport à la stratégie du joueur 2. En effet, si il joue également cet équilibre mixte, ses deux actions appartiennent au support de sa meilleure réponse à la stratégie mixte du joueur 2. On a alors

$$\mathbb{E}_{A_2 \sim \pi_2} [u_1(B, A_2)] = \mathbb{E}_{A_2 \sim \pi_2} [u_1(S, A_2)] \quad (2.15)$$

$$1 - p = 2p \quad (2.16)$$

$$p = 1/3. \quad (2.17)$$

En répétant la même procédure pour le joueur 1 on trouve l'équilibre mixte  $\pi^* = (\pi_1^*, \pi_2^*)$  avec  $\pi_1^* = \pi_2^* = (1/3, 2/3)$ , et dont l'utilité espérée est égale à  $2/3$  pour chaque joueur. Les émetteurs ont donc la possibilité de randomiser leurs choix de bande de fréquence, le risque étant pris de ne pas pouvoir établir de communications, ou de jouer l'un des deux équilibres en stratégies pures. Dans ce dernier cas, il est important que les deux joueurs puissent anticiper lequel des deux équilibres devrait être joué. Ceci peut se faire par un algorithme d'apprentissage itératif ou par un mécanisme extérieur permettant aux joueurs de s'accorder sur l'équilibre à jouer. Le paragraphe 2.1.5 reviendra plus en détail sur cette dernière solution.

Dans l'exemple du chifoumi, l'examen de la matrice 2.1 fait à la lumière des concepts introduits jusqu'à maintenant permettra de confirmer l'intuition selon laquelle l'équilibre est un profil de stratégies mixtes assignant une probabilité  $1/3$  pour chaque action de chaque joueur.

Enfin, le dilemme du prisonnier de la matrice 2.3 possède un unique équilibre de Nash en stratégie pures qu'il est possible de déterminer par la suppression itérative des stratégies dominées. L'examen de la matrice montre que pour les deux joueurs la stratégie de coopération  $C$  est dominée strictement par la stratégie de trahison  $D$ , ainsi l'équilibre de Nash de ce jeu est  $(D, D)$ . Le caractère de dilemme de ce jeu est clair dans la mesure où ce profil de stratégies est celui offrant le moins d'utilité aux joueurs, la coopération étant le résultat intuitivement préférable. Cette coopération ne peut cependant être le fruit que d'une discussion faite entre les prisonniers en amont du jeu, pourvu qu'ils aient chacun la connaissance commune de la matrice 2.3. C'est une

condition nécessaire mais non suffisante car l'un des deux prisonniers pourrait toujours décider ensuite de trahir son compagnon pour maximiser son gain. Dans le contexte du partage de ressources en communications, cela est différent dans la mesure où la coopération entre systèmes devrait être mise en place par des entités extérieures, par exemple, le nœud central d'un réseau capable de calculer les points de coopération intéressants, ou un organisme de normalisation du spectre qui fixe a priori les bandes de fréquences à utiliser.

### 2.1.5 Équilibre corrélé

L'équilibre de Nash du paragraphe précédent offre des solutions en stratégies mixtes ou pures. Les équilibres corrélés complètent les notions précédentes en offrant des solutions basées sur des stratégies randomisées où les variables aléatoires des choix d'actions des joueurs sont corrélées.

**Définition 2.1.13** (Équilibre corrélé). *Un équilibre corrélé de  $\mathcal{G} = (\mathcal{I}, \mathcal{A}, u)$  est un profil de stratégies randomisées  $\pi^* \in \Delta \mathcal{A}$  tel que  $\forall i \in \mathcal{I}$  on ait*

$$\mathbb{E}_{A \sim \pi^*} [u_i(A)] \geq \mathbb{E}_{A_i \sim \pi_i} \left[ \mathbb{E}_{A_{-i} \sim \pi_{-i}^*} \left[ u_i(a_i, A_{-i}) \mid A_i = a_i \right] \right] \quad \forall \pi_i \in \Delta \mathcal{A}_i. \quad (2.18)$$

De manière plus intuitive, les joueurs sont supposés prendre leur décision conditionnellement à la réalisation d'une variable aléatoire représentant un signal (privé ou public) qu'ils auraient observé avant de jouer. Ceci correspond au formalisme de [Aumann, 1986]. Dans ce contexte, les stratégies corrélées sont représentées comme des fonctions associant un profil d'action à chaque valeur prise par la variable de signalisation. Notons  $\Omega = \Omega_1 \times \cdots \times \Omega_I$  l'espace probabilisé des signaux envoyés aux joueurs. Il n'est pas nécessaire que chaque joueur observe le même signal, mais pour conserver un formalisme simple il sera considéré que c'est le cas ici. Un élément  $\omega$  est choisi au hasard afin de suggérer à chaque joueur  $i \in \mathcal{I}$  de choisir l'action donnée par la fonction  $A_i(\omega)$ , cette fonction étant définie par l'équation (2.2). Si les joueurs décident tous de suivre la recommandation, alors le tuple de fonctions  $A(\omega) = (A_1(\omega), \dots, A_I(\omega))$  est une stratégie corrélée. On notera  $\mathcal{F}(\mathcal{A}_i)$  l'ensemble des fonctions de  $\Omega$  dans  $\mathcal{A}_i$ .

**Définition 2.1.14** (Équilibre corrélé [Aumann, 1986]). *Un équilibre corrélé de  $\mathcal{G} = (\mathcal{I}, \mathcal{A}, u)$  est un tuple de choix d'action  $A^* = (A_1^*, \dots, A_I^*)$  tel que  $\forall i \in \mathcal{I}$  on ait*

$$\mathbb{E} [u_i(A^*)] \geq \mathbb{E} [u_i(A_i, A_{-i}^*)] \quad \forall A_i \in \mathcal{F}(\mathcal{A}_i) \quad (2.19)$$

Ainsi, à l'équilibre corrélé, tous les joueurs ont intérêt à suivre la recommandation donnée par la fonction  $A^*$ . On peut facilement appliquer ces définitions au jeu d'anti-coordination 2.2. Les équilibres de Nash en stratégies pures  $(B, S)$  et  $(S, B)$  présentent le défaut d'être inéquitables, tandis que l'équilibre en stratégies mixtes fait prendre le risque aux émetteurs de choisir la même bande et d'obtenir ainsi une utilité nulle. Supposons que l'on soit capable de fournir à ces émetteurs un signal commun  $\omega \in \{0, 1\}$  distribué selon la loi de Bernouilli  $\mathcal{B}(1/2)$ , et les recommandations suivantes aux joueurs :

- si  $\omega = 0$ , les émetteurs choisissent de jouer l'équilibre  $A(\omega) = (B, S)$ ,
- si  $\omega = 1$ , les émetteurs choisissent de jouer l'équilibre  $A(\omega) = (S, B)$ .

Ainsi, chacun des deux équilibres de Nash en stratégie pure est joué avec probabilité  $1/2$  si les joueurs suivent la recommandation, ce qui constitue un équilibre corrélé car aucun joueur n'a intérêt à s'en détourner (les deux joueurs auraient une utilité nulle). L'utilité espérée de cet équilibre corrélé est égale à  $1,5$  pour chaque joueur. En pratique, le signal  $\omega$  doit être observé par les deux émetteurs avant le choix de leur bande de fréquence. Dans un réseau cellulaire par exemple, ce signal pourrait-être fourni par une station de base. Cette station de base doit néanmoins être capable de déduire cet équilibre corrélé, ce qui ne peut se faire qu'en ayant la connaissance complète et parfaite des éléments du jeu : actions, joueurs, et fonctions d'utilité. En communication ASM, la mise en place d'un dispositif de signalisation de ce genre reste à ce jour assez difficile, sauf, éventuellement, dans certaines applications spécifiques où une structure de réseau bien définie est nécessaire. Ainsi, de tels équilibres peuvent être difficiles à obtenir à moins qu'ils ne soient le résultat d'un algorithme d'apprentissage qui respecte le caractère à la fois non-coopératif et décentralisé des communications ASM dans leur cadre le plus général (voir à ce sujet le paragraphe 1.3).

### 2.1.6 Pareto-optimalité

La notion d'optimalité en théorie des jeux est relativement floue et dépend du point de vue selon lequel les jeux sont étudiés. Les équilibres de Nash et les équilibres corrélés sont optimaux au sens où les joueurs sont considérés comme des agents maximisateurs égoïstes de leur seule fonction d'utilité, sous contrainte de la connaissance du jeu dont ils disposent (observation d'un signal, des stratégies adverses, connaissance des fonctions d'utilité, etc). L'optimalité au sens de Pareto est, au contraire, une notion relative aux joueurs considérés dans leur ensemble plutôt qu'individuellement.

**Définition 2.1.15** (Domination au sens de Pareto). *Soit  $\mathcal{G} = (\mathcal{I}, \mathcal{A}, u)$ . Un profil de stratégies  $\pi \in \Delta\mathcal{A}$  domine  $\pi' \in \Delta\mathcal{A}$  au sens de Pareto si  $\forall i \in \mathcal{I}, u_i(\pi) \geq u_i(\pi')$  et si  $\exists j \in \mathcal{I}$  tel que  $u_j(\pi) > u_j(\pi')$ .*

**Définition 2.1.16** (Pareto-optimalité). *Soit  $\mathcal{G} = (\mathcal{I}, \mathcal{A}, u)$ . Un profil de stratégies  $\pi^* \in \Delta\mathcal{A}$  est optimal au sens de Pareto si il n'existe aucun profil de stratégie qui le domine au sens de Pareto.*

En d'autres termes, il n'est pas possible d'augmenter les gains d'un joueur sans que cela soit au détriment des gains d'un autre lorsqu'ils dévient d'un profil de stratégies Pareto-optimal. Pour jouer des stratégies Pareto-optimales, et à moins qu'elles ne soient des équilibres de Nash, il est nécessaire que les joueurs aient la possibilité de coopérer et dépassent les comportements égoïstes de maximisation de l'utilité individuelle.

**Définition 2.1.17** (Optimum et bien-être social). *Un optimum social du jeu  $\mathcal{G} = (\mathcal{I}, \mathcal{A}, u)$  est un profil de stratégies  $\pi \in \Delta\mathcal{A}$  tel que le bien-être social défini par*

$$SW(\pi) = \sum_{i \in \mathcal{I}} u_i(\pi) \tag{2.20}$$

est maximal. Par définition, un optimum social est Pareto-optimal.

À titre d'exemple, le dilemme du prisonnier possède trois stratégies Pareto-optimales qui sont  $(C, D)$ ,  $((D, C)$ , et  $(C, C)$ , où cette dernière stratégie est un *optimum social* - défini comme un profil de stratégie maximisant la somme des gains des joueurs. Pour jouer la stratégie  $(C, C)$ , il est nécessaire que les deux prisonniers passent un accord avant l'interrogatoire. Bien sur, ceci n'empêche pas l'un des prisonniers de trahir son complice pour tirer le meilleur parti de la situation. Dans le jeu d'anti-coordination 2.2, on pourra noter que les deux équilibres de Nash en stratégies pures et l'équilibre corrélé sont des optima sociaux.

### 2.1.7 Apprentissage distribué des équilibres de Nash

Jusqu'ici, les définitions données des équilibres de Nash et des équilibres corrélés supposent que les joueurs sont en mesure de prédire les stratégies rationnelles de leurs adversaires et d'en déduire leurs propres stratégies rationnelles grâce au processus d'introspection itératif impliqué par la connaissance commune du jeu. Cette approche échoue néanmoins à expliquer comment les joueurs pourraient jouer un équilibre dès lors que le jeu en possède plusieurs ou que la connaissance commune ne peut plus être supposée. Une approche alternative permet de considérer que les équilibres sont appris par les joueurs au cours d'interactions répétées dans le même jeu. Ces interactions sont dictées par des règles itératives, souvent simples, qui, à partir de retours d'informations de la part de l'environnement, permettent d'inférer les stratégies optimales à mettre en œuvre. Les hypothèses de connaissance commune et de rationalité, souvent trop restrictives, peuvent alors être relâchées. Dans ce contexte, la théorie des jeux et l'apprentissage automatique permettent d'étudier l'apprentissage multi-agents de stratégies jugées « bonnes » au sens d'un certain critère.

Il existe de nombreux algorithmes d'apprentissage pour les systèmes composés de multiples agents en interaction dans un cadre défini par la théorie des jeux non-coopératifs. La notion d'optimalité peut donc être rapportée à la maximisation individuelle de l'utilité de chaque joueur. Par conséquent, les points de convergence attendus de ces algorithmes sont des équilibres de Nash (ou corrélés, éventuellement). L'apprentissage dans les jeux est un sujet vaste avec une littérature riche, et on ne détaillera dans ce paragraphe que l'algorithme classique de *dynamique de meilleures réponses*, celui-ci étant central dans le chapitre suivant. L'apprentissage par renforcement ou par regret sont d'autres méthodes possibles pour l'apprentissage multi-agent d'équilibres donnés par la théorie des jeux. Le lecteur intéressé est invité à consulter les ouvrages [Fudenberg and Levine, 1998] et [Leyton-Brown and Shoham, 2009, Chapitre 7].

#### Dynamique de meilleures réponses

La dynamique de meilleures réponses est un algorithme au cours duquel les joueurs jouent successivement leurs meilleures réponses aux profils de stratégies des

itérations précédentes. Dans sa version séquentielle, l'algorithme opère de la manière suivante : au temps  $t$ , si c'est au joueur  $i$  de mettre à jour sa stratégie, on a

$$a_i(t+1) = BR_i(a_1(t+1), \dots, a_{i-1}(t+1), a_{i+1}(t), \dots, a_I(t)). \quad (2.21)$$

En version simultanée, à toute itération  $t$  de l'algorithme, tous les joueurs mettent à jour leur stratégie selon

$$a_i(t+1) = BR_i(a_{-i}(t)). \quad (2.22)$$

Enfin, on peut en donner une version asynchrone où l'ordre de mise à jour n'est pas pré-déterminé et où les connaissances qu'entretiennent chaque joueur sur les stratégies adverses peuvent être périmées. Le cadre asynchrone généralise les algorithmes simultanés et séquentiels. En notant  $\mathcal{T} = 1, 2, \dots$ , les itérations de l'algorithme,  $\mathcal{T}_i$  l'ensemble des itérations où le joueur  $i$  joue, et  $\tau_i(t)$  le dernier temps avant  $t$  où  $i$  a obtenu une connaissance de la stratégie  $a_{-i}$ , on a

$$a_i(t+1) = BR_i(a_{-i}(\tau_i(t))) \quad \forall t \in \mathcal{T}_i. \quad (2.23)$$

Ce type d'algorithme converge vers des équilibres de Nash en stratégies pures [Bacci et al., 2016]. Néanmoins, sa convergence est, en général, soumise à conditions. Comme stipulé par la propriété 2.1.1, tout équilibre de Nash est un point-fixe des correspondances de meilleure réponse  $BR$ . Définissons la suite  $\{a(t)\}_{t \geq 0}$  par

$$a(t+1) = BR(a(t)) \quad (2.24)$$

où les composantes de  $a(t+1)$  sont mises à jour selon les règles des schémas (2.21), (2.22), ou (2.23), avec  $BR$  donnée par l'équation (2.10). Si cette suite tend vers un certain  $a^*$  tel que  $a^* = BR(a^*)$ , l'algorithme de meilleure réponse converge bien vers l'équilibre de Nash  $a^*$  pour un certain  $a(0)$  initial.

Cet algorithme d'apprentissage sera utilisé au chapitre 3 pour le partage de spectre entre systèmes de communications ASM non-coopératifs. L'annexe C donne un théorème de convergence général pour cet algorithme, et une application au cas particulier exposé dans le chapitre 3.

## 2.2 Jeux de satisfaction

Les jeux dits de satisfaction sont un autre modèle de jeux non-coopératifs, où les agents sont considérés non plus comme maximisateurs égoïstes de leur fonction d'utilité respective, mais comme recherchant un certain degré de satisfaction exprimé par cette utilité. L'idée originale de ce type de jeu a été proposée en premier par [Ross and Chaib-draa, 2006]. Elle va être exposée succinctement dans ce qui suit, avant d'être plus détaillée dans les paragraphes suivants pour les besoins du chapitre 4.

Reprenons l'exemple du dilemme du prisonnier de la figure 2.3. L'équilibre de Nash de ce jeu est le résultat le plus néfaste vis-à-vis de l'optimum social, atteint pour la stratégie  $(C, C)$ . On aimerait faire en sorte que les joueurs coopèrent implicitement, sans remettre en cause l'hypothèse de rationalité. Supposons qu'à chaque joueur soit

	$C$	$D$
$C$	1, 1	0, 1
$D$	1, 0	0, 0

 FIGURE 2.5 – Matrice du jeu dilemme du prisonnier sous forme de satisfaction avec  $\Gamma_i = -2$ .

assigné un niveau de satisfaction. Posons-le égal à  $\Gamma_i = -2$  pour chaque joueur, et écrivons la matrice du jeu résultant de la comparaison à ce seuil des utilités du dilemme du prisonnier selon la règle

$$\begin{cases} u_i(a_1, a_2) \geq \Gamma_i \Rightarrow u'_i(a_1, a_2) = 1 \\ u_i(a_1, a_2) < \Gamma_i \Rightarrow u'_i(a_1, a_2) = 0 \end{cases} \quad (2.25)$$

où  $u'_i$  sont les utilités du nouveau jeu. La matrice résultante est donnée par la figure 2.5.

Des joueurs rationnels du jeu 2.5 n'ont donc plus aucun intérêt à trahir leur compagnon. L'équilibre de Nash devient  $(C, C)$  et les joueurs peuvent le jouer de manière non-coopérative, en maximisant leur nouvelle utilité. Dans ce cas précis, cet équilibre les amène à jouer le profil de stratégies socialement optimal. Cette idée est à la base de la théorie des jeux de satisfaction qui a été développée par la suite par [Goonewardena et al., 2017; Mériaux et al., 2012; Perlaza et al., 2012] pour le domaine des télécommunications.

## 2.2.1 Jeu et équilibre de satisfaction

**Définition 2.2.1.** *Un jeu de satisfaction est défini par le tuple*

$$\mathcal{G} = (\mathcal{I}, \mathcal{A}, (u_i)_{i \in \mathcal{I}}, (\varphi_i)_{i \in \mathcal{I}}) \quad (2.26)$$

où

- $\mathcal{I} = \{1, \dots, I\}$  est l'ensemble des joueurs impliqués dans le jeu,
- $\mathcal{A} = \mathcal{A}_1 \times \dots \times \mathcal{A}_I$  où  $\mathcal{A}_i$  est l'ensemble de stratégies  $i \in \mathcal{I}$ . De manière usuelle, les choix de stratégies du joueur  $i$  sont notés  $a_i \in \mathcal{A}_i$ , ceux des autres joueurs que  $i$  par  $a_{-i} = [a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_I] \in \mathcal{A}_{-i} = \mathcal{A} \setminus \mathcal{A}_i$  et les profils de stratégies  $a \in \mathcal{A}$ .
- $u_i : \mathcal{A} \rightarrow \mathbb{R}$  est la fonction d'utilité du joueur  $i$ .
- La correspondance de satisfaction  $\varphi_i : \mathcal{A}_{-i} \rightarrow \mathcal{P}(\mathcal{A}_i)$  associe à tout profil de stratégie  $a_{-i} \in \mathcal{A}_{-i}$  des adversaires de  $i$  une partie de l'ensemble de stratégies de ce dernier contenant tous les  $a_i \in \mathcal{A}_i$  qui respectent une certaine contrainte, pouvant être exprimée en terme d'utilité.

Dans la suite, on se restreint à des correspondances de satisfaction de la forme

$$\varphi_i(a_{-i}) = \{a_i \in \mathcal{A}_i : u_i(a_i, a_{-i}) \geq \Gamma_i\} \quad \forall a_{-i} \in \mathcal{A}_{-i} \quad (2.27)$$

Le joueur  $i$  est donc satisfait par toute stratégie  $a_i \in \varphi_i(a_{-i})$  car, pour le profil  $a = (a_i, a_{-i})$ , son utilité  $u_i(a)$  est au moins aussi grande que la contrainte  $\Gamma_i$ . Notons qu'éventuellement la correspondance  $\varphi_i$  peut retourner l'ensemble vide auquel cas le joueur  $i$  ne peut pas être satisfait.

En communications numériques, un exemple de jeu de satisfaction simple peut être défini en posant des paires émetteur-récepteurs comme joueurs de l'ensemble  $\mathcal{I}$ , possédant des actions  $\mathcal{A}_i = \{a_i^1, \dots, a_i^{N_i}\}$ , où les  $a_i \in \mathcal{A}_i$  correspondent à des niveaux de puissance ordonnés, et des fonctions d'utilité de type rapport signal à bruit plus interférence (RSIB) :

$$u_i(a_i, a_{-i}) = \frac{h_{ii}a_i}{\sigma^2 + \sum_{j \neq i} h_{ji}a_j}, \quad (2.28)$$

où  $\sigma^2$  est la puissance du bruit, et  $h_{ji}$  est le gain du canal entre le récepteur  $i$  et l'émetteur  $j$ , ce dernier étant considéré comme source d'interférence pour le récepteur  $i$ . De nombreuses métriques pertinentes dans les modèles théoriques de communication numérique peuvent s'exprimer comme fonction de ce RSIB, c'est le cas par exemple de la capacité ou de la probabilité de transmission sans erreur. Enfin on peut définir un seuil de RSIB au delà duquel les joueurs sont satisfaits, par exemple  $\Gamma_i = 10$  dB. Ce seuil peut représenter par exemple une qualité de lien minimale nécessaire à l'établissement d'une communication fiable. On considère que si cette contrainte n'est pas satisfaite, les transmissions échouent. Du point de vue d'un joueur  $i$  quelconque et pour une stratégie  $a_{-i} \in \mathcal{A}_{-i}$  donnant les puissances de transmission des autres couples émetteur-récepteur, la correspondance  $\varphi_i(a_{-i})$  renvoie toutes les puissances  $a_i \in \mathcal{A}_i$  telles que le couple  $i$  puisse communiquer avec un RSIB en réception excédant 10 dB.

Pour tout profil de stratégies  $a \in \mathcal{A}$ , on peut diviser l'ensemble des joueurs du jeu en trois partitions disjointes dont l'union recouvre  $\mathcal{I}$ . On a

$$\mathcal{I}_S(a) = \{i \in \mathcal{I} : a_i \in \varphi_i(a_{-i})\}, \quad (2.29)$$

les joueurs satisfaits pour lesquelles la stratégie jouée satisfait la contrainte (*ex.* la transmission s'est établi avec un RSIB supérieur à 10 dB),

$$\mathcal{I}_N(a) = \{i \in \mathcal{I} : a_i \notin \varphi_i(a_{-i}), \varphi_i(a_{-i}) \neq \emptyset\}, \quad (2.30)$$

les joueurs insatisfaits ayant un ensemble non-vide de stratégies satisfaisantes, et qui possèdent donc une déviation de  $a_i$  qui leur est profitable face à  $a_{-i}$  restant fixé (*ex.* la transmission a échoué mais le couple émetteur-récepteur  $i$  possède une puissance  $a_i \in \mathcal{A}_i$  telle que le RSIB supérieur 10 dB pourvu que les autre couples conservent leur stratégie),

$$\mathcal{I}_U(a) = \{i \in \mathcal{I} : \varphi_i(a_{-i}) = \emptyset\}, \quad (2.31)$$

les joueurs insatisfaits qui n'ont aucune stratégie satisfaisante au regard des actions  $a_{-i}$  adverses (la transmission a échoué et le couple émetteur-récepteur  $i$  n'a aucune puissance  $a_i \in \mathcal{A}_i$  permettant d'obtenir un RSIB suffisant). Il apparaît clairement que si les joueurs sont rationnels et partagent une connaissance complète et commune du jeu  $\mathcal{G}$ , alors aucun d'entre eux ne jouera de stratégies qui le conduiraient dans l'ensemble  $\mathcal{I}_N(a)$ . Ainsi, un profil d'*équilibre de satisfaction*, notons-le  $a^*$ , doit être tel que  $\mathcal{I}_N(a)$  soit vide. Ceci nous conduit à la définition suivante de l'équilibre de satisfaction :

**Définition 2.2.2** (Équilibre de satisfaction généralisé [Goonewardena et al., 2017]).  
 Un profil de stratégies  $a^* = (a_1^*, \dots, a_I^*)$  est un équilibre de satisfaction généralisé si et seulement si  $\mathcal{I}_N(a^*) = \emptyset$ .

La définition originale de [Goonewardena et al., 2017] parle d'équilibre de satisfaction *généralisé* au sens où elle généralise en effet l'équilibre de [Ross and Chaib-draa, 2006], repris par [Perlaza et al., 2012], aux cas où l'on accepte que certains joueurs puissent ne pas satisfaire leurs contraintes. Dans ce manuscrit, on parlera d'équilibre de satisfaction sans mentionner la généralité du concept. On voit facilement que la définition 2.2.2 implique qu'aucun joueur n'a de déviation profitable de sa stratégie  $a_i^*$ , ce qui correspond bien à ce que l'on attend d'un équilibre.

Il est démontré dans les travaux de [Goonewardena et al., 2017] que si les espaces de stratégies des joueurs sont finis et que les  $\varphi_i$  s'expriment comme à l'équation (4.12), alors il existe au moins un équilibre de satisfaction en stratégies mixtes pour le jeu  $\mathcal{G}$ . Ceci s'explique naturellement par l'existence garantie d'un équilibre de Nash dans tout jeu fini et, donc, en particulier pour le jeu sous forme normale  $\mathcal{G}' = (\mathcal{I}, \mathcal{A}, (u_i)_{i \in \mathcal{I}})$  défini avec  $u_i$  pour fonctions d'utilité à maximiser. Cet équilibre de Nash est nécessairement un équilibre de satisfaction car, pour tout équilibre de Nash, l'ensemble des joueurs est en effet divisible de telle manière que la définition 2.2.2 soit respectée, et, par définition de l'équilibre de Nash, aucun joueur n'a de déviation profitable en terme d'utilité. On notera également que tout jeu à espaces de stratégies infini qui possède un équilibre de Nash possède par conséquent un équilibre de satisfaction. Enfin, si on note  $NE$  l'ensemble des équilibres de Nash d'un jeu quelconque et  $SE$  l'ensemble des équilibres de satisfaction du même jeu tels que définis par la définition 2.2.2, on a la relation suivante

$$NE \subseteq SE, \quad (2.32)$$

qui signifie que tout équilibre de Nash est nécessairement un équilibre de satisfaction.

Pour garantir l'existence d'un équilibre de satisfaction en stratégies pures et la convergence des algorithmes d'apprentissage, des hypothèses supplémentaires doivent être posées. Les ensembles  $\mathcal{A}_i$  sont supposés bornés et munis d'une relation d'ordre total notée  $\preceq$ . On rappelle qu'une relation d'ordre total est réflexive, transitive et anti-symétrique et que deux éléments de l'ensemble sur lequel elle est définie sont toujours comparables. Les ensembles  $\mathcal{A}_i$  ont (au moins) un majorant  $\bar{a}_i$  et un minorant  $\underline{a}_i$ . L'ordre des stratégies des joueurs dans leurs espaces  $\mathcal{A}_i$  respectifs peut être vu comme une forme d'effort ou de coût associé à chaque stratégie et que le joueur doit payer pour jouer. Il est également considéré que pour tout  $i \in \mathcal{I}$ , les fonctions  $u_i$  sont monotones croissantes en  $a_i \in \mathcal{A}_i$  pour tout  $a_{-i} \in \mathcal{A}_{-i}$  fixé. Sous ces nouvelles hypothèses, un équilibre de satisfaction en stratégies pures est garanti : cet équilibre est l'équilibre de Nash  $\bar{a} = (\bar{a}_1, \dots, \bar{a}_I)$  du jeu sous forme normale  $\mathcal{G}'$  où tous les joueurs utilisent leur stratégie la plus coûteuse. Le jeu de communications numériques pris en exemple ici respecte toutes ces hypothèses.

## 2.2.2 Apprentissage décentralisé des équilibres de satisfaction

Comme pour les équilibres de Nash des jeux sous forme normale, les équilibres de satisfaction peuvent être appris de façon décentralisée au moyen d’algorithmes distribués dans les contextes où les hypothèses de rationalité et de connaissance commune ne s’appliquent pas, ou que les équilibres recherchés ne sont pas uniques. L’algorithme présenté ici a été proposé par [Goonewardena et al., 2017]. L’idée générale est assez simple. On se place dans un cadre où chaque joueur doit produire un certain *effort* pour jouer une stratégie et où, lorsque les adversaires produisent plus d’efforts, il est nécessaire de produire également plus d’effort pour être satisfait<sup>2</sup>. Les joueurs commencent à jouer l’action requérant le moins d’effort dans leur ensemble de stratégies. Au cours du processus, les joueurs ayant une déviation profitable en termes de satisfaction jouent la stratégie satisfaisante demandant le plus petit effort. Ainsi, et considérant que les joueurs ont une limite quant à l’effort qu’ils peuvent produire, l’algorithme converge nécessairement vers un équilibre de satisfaction.

La réponse de satisfaction d’un joueur  $i$  donné à un profil de stratégies  $a_{-i}$  est définie comme étant la sélection d’une stratégie particulière dans l’ensemble retourné par sa correspondance de satisfaction  $\varphi_i(a_{-i})$ . Cet ensemble n’est en général pas un singleton, et la question du choix d’une stratégie parmi celles qui sont satisfaisantes peut être posée. On introduit le concept de réponse de satisfaction de moindre coût, identifiable à la plus petite stratégie dans l’ensemble  $\varphi_i(a_{-i})$  au sens de la relation d’ordre  $\preceq$ , que l’on définit par :

$$SR_i(a_{-i}) \in \min_{a_i \in \varphi_i(a_{-i})} a_i. \quad (2.33)$$

Dans l’exemple du paragraphe précédent, ceci permet d’exprimer, dans une certaine mesure, un objectif de minimisation de la puissance, en plus de l’objectif de satisfaction de la contrainte de RSIB. Les mêmes mécanismes seront réutilisés au chapitre 4.

---

### Algorithme 1 Dynamique de réponses de satisfaction

---

- 1:  $t = 0$
- 2:  $a(t) = (a_i)_{i \in \mathcal{I}}$
- 3: **tant que**  $\mathcal{I}_N(a(t)) \neq \emptyset$  **faire**  $\forall i \in \mathcal{I}$
- 4:     **si**  $t \in \mathcal{T}_i$  **alors**
- 5:

$$a_i(t+1) = \begin{cases} SR_i(a_{-i}(t)) & \text{if } i \in \mathcal{I}_N(a(t)) \\ a_i(t) & \text{otherwise} \end{cases}$$

- 6:     **fin si**
  - 7:      $t = t + 1$
  - 8: **fin tant que**
- 

Soient  $\mathcal{T} = \{0, 1, 2, \dots\}$  des indices dans un ensemble de temps représentés par une séquence croissante de nombres réels, espacés d’un pas quelconque. À tout temps

---

2. En théorie des jeux, les jeux modulaires présentent une structure similaire [Altman and Altman, 2003; Fudenberg and Tirole, 1991].

ou itération  $t \in \mathcal{T}$ , et tant qu'un équilibre n'est pas atteint par les joueurs, certains de ces derniers auront à jouer une réponse de satisfaction  $a_i \in \varphi_i(a_{-i}(t))$ . On se place ici dans un cadre asynchrone, où le profil  $a_{-i}(t)$  perçu par le joueur  $i$  au temps  $t$  peut ne pas correspondre aux  $(a_j)_{j \neq i}$  jouées effectivement par ses adversaires parce que la connaissance acquise par  $i$  est périmée. On note  $\mathcal{T}_i \subseteq \mathcal{T}$  l'ensemble des temps où le joueur  $i$  peut jouer une nouvelle stratégie. L'algorithme 1 offre une procédure permettant l'apprentissage décentralisé, distribué et asynchrone d'un équilibre de satisfaction en stratégies pures. Dans le jeu de satisfaction du RSIB précédent, cet algorithme consiste, pour chaque joueur, à augmenter graduellement sa puissance de transmission jusqu'à ce qu'à l'équilibre où les joueurs seront soit satisfaits avec un RSIB supérieur à 10 dB, soit insatisfaits avec aucune puissance assez grande ne permettant d'établir une communication avec un RSIB suffisant. La preuve formelle de la convergence de l'algorithme et de son extension aux stratégies mixtes est donnée dans [Goonewardena et al., 2017, Section III.A et III.B].

## 2.3 Conclusion

Ce chapitre a passé en revue les concepts de théorie des jeux nécessaires à la bonne compréhension des chapitres suivants. Les jeux sous forme normale et les jeux de satisfaction ont été étudiés en détails et illustrés à travers des exemples de jeux à matrice ou des problèmes de communications numériques simplifiés. Les algorithmes d'apprentissage distribués des équilibres qui seront mis en œuvre par la suite ont également été présentés.

Les chapitres 3 et 4 vont développer, sur cette base, des solutions à des problèmes plus concrets de partage de ressources en communications ASM. Le chapitre 3 va proposer une modélisation d'un problème de partage de spectre par un jeu sous forme stratégique où les joueurs ont une connaissance incomplète des éléments du jeu et une observation imparfaite de certains paramètres. Il sera proposé une méthode d'apprentissage basée sur les meilleures réponses capable de conduire les joueurs vers un partage de spectre efficace, malgré leur faible connaissance de leur environnement. Le chapitre 4 étudie des problèmes de partage de ressource sous contraintes de qualité de service et formalise le problème comme un jeu de satisfaction. Il est montré que les joueurs de ce jeu peuvent apprendre des équilibres de satisfaction avec une connaissance du jeu réduite à 1 bit d'information.



# Chapitre 3

## Partage de spectre robuste et décentralisé

Dans ce chapitre, un problème de partage de spectre décentralisé pour les communications ASM est formulé à l'aide des outils de la théorie des jeux. On se place dans un cadre non-coopératif où il est entendu que les systèmes de communications entrent en compétition pour l'utilisation du spectre ASM, et où le recours à l'échange de messages pour aboutir à un partage équitable est interdit. Nous allons montrer qu'il est possible de partager le spectre efficacement en ne requérant qu'une connaissance locale et limitée de l'environnement de la part des systèmes impliqués dans le problème.

De nombreux travaux, issus des technologies des communications, ont proposé des modèles à base de théorie des jeux pour résoudre des problèmes de partage de ressources [Bacci et al., 2016; Belmega et al., 2009; Mertikopoulos et al., 2012]. Dans le cas précis du canal d'interférence, on peut en particulier citer [Scutari et al., 2008a,b] dont certains résultats sont utilisés à la section 3.2. La plupart des travaux similaires en théorie des communications ne s'appliquent pas directement au cas particulier des communications ASM. Dans [Scutari et al., 2008a,b], les auteurs se concentrent surtout sur la caractérisation théorique des équilibres de Nash et sur la convergence d'algorithmes de meilleures réponses permettant de les calculer, les canaux de communications étant considérés déterministes et fixes sur la durée des transmissions. Comme on l'a vu au premier chapitre, cela est rarement vérifié en ASM. Des déclinaisons de ces travaux existent pour des canaux aléatoires et/ou une connaissance imparfaite de l'état du canal [Yang et al., 2016], cependant les algorithmes proposés nécessitent des échanges d'informations entre les couples émetteurs-récepteurs impliqués dans le problème de partage de la ressource. Cette hypothèse d'une connaissance partagée de l'environnement, également difficile à vérifier dans le contexte de cette thèse, est présente plus ou moins explicitement dans de nombreux autres travaux de théorie des communications sur l'allocation optimale de ressources [Feng et al., 2004; Schmidt et al., 2009; Scutari et al., 2014]. Enfin, et afin de permettre une adaptation des stratégies de transmission, il est souvent supposé que les émetteurs acquièrent la connaissance des réalisations instantanées du canal et de l'environnement multi-utilisateur. Cette connaissance étant en pratique fournie par le récepteur, cela implique que le lien de retour du récepteur vers l'émetteur n'admette pas de délai de propagation, ce qui, encore une fois, est irréaliste

en ASM.

Nous nous intéresserons ici à des liens de communication ASM utilisant une modulation OFDM. Le choix de cette modulation est justifié par la flexibilité qu'elle offre pour le partage du spectre, la possibilité étant laissée aux émetteurs de choisir les niveaux de puissance alloués sur un ensemble de sous-bandes fréquentielles. À la section 3.1, un jeu non-coopératif est formulé, où les joueurs sont donc les liens de communications OFDM. Leurs stratégies sélectionnent des allocations de puissance possibles sur le spectre partagé, et les fonctions d'utilité prennent la forme d'un critère d'optimisation robuste aux interférences mutuelles. L'équilibre de Nash du jeu est donné à la section 3.2, où les meilleures réponses des joueurs correspondent à une allocation de puissance de type water-filling, calculées par les émetteurs sur la base de statistiques sur le canal fournies par leur récepteur. Les erreurs d'estimation de ces statistiques sont prises en compte par un jeu robuste à la section 3.3. Des résultats numériques obtenus grâce à des canaux synthétiques et des données issus de sondage de canaux en mer illustrent, en section 3.4, l'efficacité de la méthode de partage de spectre proposée, avant de conclure le chapitre.

## 3.1 Formulation du problème

### 3.1.1 Modèle de la transmission et hypothèses

On considère un ensemble  $\mathcal{I} = \{1, \dots, I\}$  de  $I$  paires émetteur-récepteur (TX-RX) fixées communiquant en modulation OFDM dans le canal ASM. Les communications sont supposées avoir lieu en même temps et dans la même bande de fréquence  $B$ , divisée en  $N$  sous-porteuses espacées de  $\Delta f$  Hertz. Cet ensemble de liens de communications forme les joueurs du jeu que l'on va formuler par la suite. Le canal de communication est sélectif en fréquence et varie aléatoirement dans le temps. On pose l'hypothèse que son étalement temporel n'est pas supérieur à  $1/\Delta f$ . Le temps symbole est donné par  $T = T_s + T_g$  où  $T_s = 1/\Delta f$  et  $T_g$  est la durée du préfixe cyclique. Soit  $\mathbf{x}_{i,n} \in \mathbb{C}^L$  un vecteur de  $L$  symboles, i.i.d. et de moyennes nulles, transmis par l'émetteur  $i \in \mathcal{I}$  sur la sous-porteuse  $n$ . Pour tout  $i$  et quelle que soit la sous-porteuse  $n' \in \mathcal{N} = \{1, \dots, N\}$ , les symboles  $\mathbf{x}_{j \neq i, n'}$  transmis par les autres émetteurs sont indépendants de ceux transmis par  $i$ . Les émetteurs sont chacun soumis à leur propre contrainte de puissance

$$\sum_{n=1}^N p_{i,n} \leq P_i^{max}, \quad (3.1)$$

où  $p_{i,n}$  est la puissance allouée par  $i$  sur la  $n^{\text{ème}}$  sous-porteuse. Après compensation de l'effet Doppler, suppression du préfixe cyclique et transformation de Fourier discrète du signal reçu, la sortie du canal observée sur  $L$  symboles OFDM par le récepteur  $i$  sur la sous-porteuse  $n$  s'écrit

$$\mathbf{y}_{i,n} \triangleq \mathbf{h}_{ii,n} \odot \mathbf{x}_{i,n} + \sum_{j \neq i} \mathbf{h}_{ji,n} \odot \mathbf{x}_{j,n} + \mathbf{w}_{i,n} \quad (3.2)$$

où  $\mathbf{w}_{i,n}$  sont  $L$  réalisations du bruit  $W_{i,n} \sim \mathcal{CN}(0, \sigma_{i,n}^2)$  et  $\mathbf{h}_{ji,n} \in \mathbb{C}^L$  sont  $L$  réalisations d'une variable aléatoire  $H_{ji,n}$  modélisant le gain complexe du canal entre le récepteur  $i$  et l'émetteur  $j$  à la sous-porteuse  $n$ . Les variables aléatoires du bruit, des canaux et des symboles transmis par les émetteurs sont supposées indépendantes les unes des autres.

En présence d'interférences dues à l'accès simultané et anarchique à la même ressource, il se peut que certains liens TX-RX ne puissent pas établir de communications fiables. On aimerait faire en sorte que les émetteurs soient capables de partager le spectre de manière efficace et autonome, en répartissant leur puissance judicieusement sur les sous-porteuses OFDM. Cette adaptation de la puissance transmise doit être effectuée en fonction à la fois du canal et de l'interférence. Pour que l'émetteur ait la connaissance de l'état du canal entre lui et son récepteur, il est nécessaire de mettre en place un lien de retour depuis ce récepteur qui est alors en charge d'estimer le canal à partir de signaux pilotes placés en début de trame. Cependant, les longs délais de propagation du canal ASM rendent inutile le retour des réalisations instantanées du canal, du fait des temps de cohérence relativement courts. En d'autres termes, cela revient à fournir à l'émetteur une connaissance périmée de l'environnement. On préférera donc faire le retour d'une information statistique, plus robuste car issue de l'accumulation des réalisations estimées sur un temps très supérieur aux temps de cohérence du canal ASM. De manière générale, des systèmes différents évoluant hors d'un réseau et ne disposant pas de protocole de communications commun ne peuvent obtenir les estimations des canaux des autres. Les communications sont donc non-coopératives, au sens où aucune paire TX-RX ne peut échanger de messages avec les autres afin de négocier une stratégie de partage de ressource. Ainsi, chaque récepteur traite les interférences générées par les émetteurs qui ne lui sont pas appairés comme du bruit coloré de distribution inconnue, ce qui correspond au modèle de canal d'interférence parallèle (ou vectoriel) de la théorie de l'information [Cover and Thomas, 2006]. On se place donc sous les hypothèses suivantes :

1. Les distributions de probabilité des canaux  $\mathbf{H}_{ji} = (H_{ji,n})_{n \in \mathcal{N}}$  sont inconnues et rien n'est connu de la part du joueur  $i \in \mathcal{I}$  à propos de  $\mathbf{x}_{j,n}$ ,  $j \neq i$ .
2. Tout émetteur  $i \in \mathcal{I}$  n'a qu'une connaissance statistique de son propre canal  $\mathbf{H}_{ii}$  et de l'agrégation de l'interférence plus bruit, acquise par lien de retour depuis le récepteur.
3. Les  $\mathbf{H}_{ji}$  sont considérés stationnaires sur la durée d'un bloc de  $L$  symboles.
4. La durée  $L \times T$  est considérée très supérieure au temps de cohérence du canal.
5. Les délais de propagation sont supérieurs au temps de cohérence des évanouissements rapides du canal.
6. Les émetteurs sont synchronisés.

Considérant ces hypothèses, nous allons formuler un jeu qui capture le problème de partage de spectre décentralisé dans le canal ASM entre émetteurs OFDM non-coopératifs dont l'objectif est de mettre en œuvre une stratégie d'allocation de puissance optimisée vis-à-vis d'un certain critère de performance, qui reste encore à définir.

### 3.1.2 Jeu de maximisation du débit moyen

La manière dont un émetteur répartit sa puissance sur la bande de fréquence partagée impacte les possibilités laissées aux autres émetteurs pour leurs propres allocations de puissance. Mathématiquement, on a  $I$  problèmes d'optimisation à résoudre simultanément. Pour chaque problème, la variable d'optimisation est le vecteur de puissance  $\mathbf{p}_i$  de l'émetteur correspondant, et les solutions optimales  $\mathbf{p}_i^*$  de chaque problème dépendent des solutions optimales  $\mathbf{p}_{-i}^*$  des autres émetteurs de la manière suivante

$$\mathbf{p}_i^* = \operatorname{argmax}_{\mathbf{p}_i} u_i(\mathbf{p}_i, \mathbf{p}_{-i}^*) \quad \forall i \in \mathcal{I} \quad (3.3)$$

sous la contrainte  $\mathbf{p}_i \in \mathcal{P}_i$

où  $u_i(\mathbf{p}_i, \mathbf{p}_{-i})$  est la fonction objectif que le joueur  $i$  cherche à maximiser. Le vecteur  $\mathbf{p}_i = [p_{i,1}, \dots, p_{i,N}]^T \in \mathbb{R}_+^N$  est son vecteur d'allocation de puissance et  $\mathbf{p}_{-i} = [\mathbf{p}_1^T, \dots, \mathbf{p}_{i-1}^T, \mathbf{p}_{i+1}^T, \dots, \mathbf{p}_M^T]^T$  sont les allocations des autres joueurs. Dans la suite, on notera le profil d'actions  $\mathbf{p} = [\mathbf{p}_1^T, \dots, \mathbf{p}_M^T]^T$ . L'espace d'actions du joueur  $i$  est l'ensemble des allocations de puissance faisables sous sa contrainte de puissance moyenne. Il est noté

$$\mathcal{P}_i \triangleq \left\{ \mathbf{p}_i \in \mathbb{R}_+^N \mid \sum_{n=1}^N p_{i,n} \leq P_i^{\max} \right\}. \quad (3.4)$$

On formule donc le jeu sous forme stratégique avec des fonctions d'utilité  $u_i$  par

$$\mathcal{G} \triangleq (\mathcal{I}, \mathcal{P}, u) \quad (3.5)$$

avec  $\mathcal{P} = \mathcal{P}_1 \times \dots \times \mathcal{P}_I$  et  $u = (u_i)_{i \in \mathcal{I}}$ , où  $u_i : \mathcal{P} \rightarrow \mathbb{R}$ . Les problèmes d'optimisation donnés par l'équation (3.3) forment un problème de recherche de l'Équilibre du Nash du jeu (3.5).

En tant que joueurs rationnels du jeu  $\mathcal{G}$ , les émetteurs vont mettre en place une stratégie d'allocation de puissance dans le but de maximiser leurs fonctions d'utilité respectives  $u_i$ . Ce critère doit être choisi de manière à respecter les hypothèses posées au paragraphe précédent quant à la connaissance disponible sur l'environnement - canal et utilisateurs. Une fonction d'utilité basée sur le débit d'information théorique est choisie pour ce problème, en cohérence avec l'usage courant de la modulation OFDM pour les transmissions ASM hauts débits [Zhou and Wang, 2014]. Étant données les contraintes sur le lien de retour, elle doit représenter un critère de performance moyen sur la durée des  $L$  symboles OFDM transmis. En vertu des hypothèses 3 et 4 du paragraphe précédent, l'information mutuelle moyenne du lien TX-RX  $i$  peut s'exprimer par la capacité ergodique [Cover and Thomas, 2006]

$$R_i(\mathbf{p}_i, \mathbf{p}_{-i}) = \alpha \sum_{n=1}^N \mathbb{E} \left[ \log \left( 1 + \frac{|H_{ii,n}|^2 p_{i,n}}{\sigma_{i,n}^2 + \sum_{j \neq i} |H_{ji,n}|^2 p_{j,n}} \right) \right] \quad (3.6)$$

où l'espérance est prise par rapport à la distribution conjointe des canaux  $(\mathbf{H}_{ji})_{j \in \mathcal{I}}$  et  $\alpha = (NT\Delta f)^{-1}$ . L'expression de  $R_i$  telle qu'elle est donnée par (3.6) néglige l'interférence entre symboles OFDM (ISI) et entre sous-porteuses (ICI) Ces hypothèses sont valides dans la mesure où, (i) le préfixe cyclique est assez long pour faire que la

durée symbole  $T$  soit plus grande que l'étalement temporel du canal, et (ii) les effets Doppler sont compensés et l'étalement fréquentiel du canal est supposé suffisamment inférieur à  $\Delta f$ . Il est de plus nécessaire que les signaux des différents émetteurs soient alignés lorsque reçus aux récepteurs afin de garantir l'orthogonalité des sous-porteuses des signaux interférents lors de la démodulation. Ceci justifie l'hypothèse de synchronisation de la section précédente. Notons cependant que les travaux de [Scutari et al., 2007] montrent que la prise en compte de l'ISI et de l'ICI dues à la non-synchronisation ne change pas fondamentalement la formulation du problème dans la mesure où ces interférences s'additionnent simplement au bruit et à l'interférence externe. Les résultats donnés à la section 3.4 seraient en revanche différents.

La résolution du problème (3.3) considérant l'objectif  $u_i = R_i$ , requiert la connaissance complète et parfaite des données du jeu de la part des joueurs. Plus précisément, chaque joueur, afin de pouvoir calculer sa capacité ergodique, doit avoir la connaissance des distributions de  $(\mathbf{H}_{ji})_{j \in \mathcal{I}}$  et être capable d'observer les stratégies ou actions adverses  $\mathbf{p}_j$ ,  $j \neq i$ . Le choix de  $R_i$  comme fonction d'utilité entre alors en contradiction avec toutes les hypothèses posées au paragraphe précédent, et qui définissent finalement un jeu à connaissance incomplète.

### 3.1.3 Connaissance incomplète de la fonction d'utilité

En théorie des jeux, une approche classique adoptée pour traiter la connaissance incomplète repose sur la formulation du problème comme un jeu bayésien [Fudenberg and Tirole, 1991]. Celle-ci a d'ailleurs été mise en pratique en communications numériques par [He et al., 2009, 2010; Noam et al., 2010]. Cette méthode est cependant discutable à plusieurs égards. D'une part, la complexité mathématique du formalisme employé semble amener les auteurs à simplifier la modélisation de la transmission. Par exemple, dans [He et al., 2009, 2010], le canal est représenté par une variable aléatoire discrète de Bernoulli, tandis que [Noam et al., 2010] se limite à un jeu à deux joueurs ayant chacun deux actions. D'autre part, les jeux bayésiens introduisent des distributions de probabilités a priori sur les croyances qu'entretiennent les joueurs les uns sur les autres, ces probabilités étant supposées connaissance commune. Se pose alors le problème pratique, au moins dans le contexte ASM, de savoir comment un a priori partagé par des systèmes hétérogènes peut être raisonnablement défini, ne sachant pas combien d'autres sources acoustiques vont être rencontrées, quelle sera leur nature, ni quels seront leurs actions, stratégies ou objectifs.

Une approche robuste, basée sur la maximisation d'une fonction d'utilité calculée pour la pire interférence envisageable, peut en revanche nous aider à résoudre le jeu (3.5) sans avoir recours à des hypothèses irréalistes vis-à-vis de la connaissance de l'environnement. Cet objectif dans le « pire cas » peut être obtenu en minimisant l'information mutuelle moyenne (3.6) vis-à-vis de la distribution des canaux inconnus  $H_{ji,n}$ . En posant  $f_{\mathbf{H}_{ji}}$  ces distributions, où  $\mathbf{H}_{ji} = (H_{ji,n})_{n \in \mathcal{N}}$ ,  $\forall j \neq i$ , on peut alors

exprimer  $u_i$  par

$$u_i(\mathbf{p}_i, \mathbf{p}_{-i}) = \min_{f_{\mathbf{H}_{ji}}} \alpha \sum_{n=1}^N \mathbb{E} \left[ \log \left( 1 + \frac{|H_{ii,n}|^2 p_{i,n}}{\sigma_{i,n}^2 + \sum_{j \neq i} |H_{ji,n}|^2 p_{j,n}} \right) \right] \quad (3.7)$$

$$= \alpha \sum_{n=1}^N \mathbb{E} \left[ \log \left( 1 + \frac{|H_{ii,n}|^2 p_{i,n}}{\sigma_{i,n}^2 + \sum_{j \neq i} \mathbb{E}[|H_{ji,n}|^2] p_{j,n}} \right) \right] \quad (3.8)$$

$$\approx \alpha \sum_{n=1}^N \log \left( 1 + \frac{e^{\mathbb{E}[\log(|H_{ii,n}|^2)]} p_{i,n}}{\sigma_{i,n}^2(n) + \sum_{j \neq i} \mathbb{E}[|H_{ji,n}|^2] p_{j,n}} \right). \quad (3.9)$$

L'équation (3.8) est conséquence de l'inégalité de Jensen. On peut montrer que l'égalité entre (3.7) et (3.8) est vérifiée lorsque l'interférence est gaussienne. L'approximation (3.9) sera utile pour exprimer les meilleures réponses des joueurs de façon analytique dans la section suivante, ainsi que pour établir des conditions d'unicité des solutions du jeu et de convergence des algorithmes distribués qui les calculent. On peut d'ailleurs donner expression analytique de (3.9) lorsque le canal  $\mathbf{H}_{ii}$  suit une loi de Rice. Les détails du calcul de la fonction d'utilité est donné en Annexe A.1.

On définit l'état statistique du canal du joueur  $i$  à la sous-porteuse  $n$  par

$$\gamma_{i,n}(\mathbf{p}_{-i}) \triangleq \frac{e^{\mathbb{E}[\log(|H_{ii,n}|^2)]}}{\sigma_{i,n}^2(n) + \sum_{j \neq i} \mathbb{E}[|H_{ji,n}|^2] p_{j,n}}, \quad (3.10)$$

ainsi la fonction d'utilité du jeu  $\mathcal{G}$  défini à l'équation (3.5) peut s'écrire

$$u_i(\mathbf{p}_i, \mathbf{p}_{-i}) = \alpha \sum_{n=1}^N \log \left( 1 + \gamma_{i,n}(\mathbf{p}_{-i}) p_{i,n} \right). \quad (3.11)$$

La forme de la fonction d'utilité qui est proposée ici ne requiert pas de connaissance sur les distributions de probabilité des canaux, ce qui permet de se passer de poser un modèle statistique des évanouissements du canal sur lequel la communauté des communications ASM n'admet pas de consensus. Résoudre le jeu (3.5) de manière non-coopérative avec des fonctions d'utilité  $u_i$  exprimées comme dans l'équation (3.11) ne demande aux joueurs qu'une connaissance locale et minimale de leur environnement. Pour un joueur  $i$  donné, seules les statistiques du canal  $\mathbf{H}_{ii}$  et la densité spectrale de puissance (DSP) de l'interférence agrégée au bruit sont nécessaires. L'observation des stratégies d'allocation de puissance des autres joueurs  $j \neq i$  n'est pas requise et on peut traiter toute l'interférence comme du bruit additif coloré. En situation réelle, il suffit que chaque récepteur mette en œuvre une méthode quelconque d'estimation des statistiques du canal et de la puissance du bruit (incluant les interférences) sur la bande de communication et renvoie périodiquement ces statistiques via le lien de retour. De plus, il est facilement envisageable d'utiliser des statistiques calculées sur une longue période afin de réduire l'activité sur le lien de retour et ainsi d'apporter une certaine robustesse vis-à-vis de la décorrélation rapide des réalisations du canal comparée au délai de propagation du lien de retour, pourvu que le canal soit stationnaire suffisamment longtemps.

On notera enfin que  $u_i$  ne doit pas être interprétée comme un débit atteignable ou une borne de débit atteignable, mais plutôt comme un critère d'optimisation utile pour conduire des émetteurs OFDM ASM vers un partage de spectre efficace de manière décentralisée et avec une connaissance minimale de leur environnement.

## 3.2 Solution de water-filling itératif

### 3.2.1 Meilleure réponse et équilibre de Nash

Le problème de maximisation compétitive (3.3) exprimé par le jeu  $\mathcal{G}$  peut être résolu grâce à l'équilibre de Nash, qui est optimal au sens de la maximisation égoïste de l'utilité de chaque joueur. On le définit comme suit dans le jeu étudié ici :

**Définition 3.2.1.** *Un équilibre de Nash du jeu  $\mathcal{G}$  défini par (3.5) est un profil d'allocations de puissance  $\mathbf{p}^*$  tel qu'on ait  $\forall i \in \mathcal{I}$  et  $\forall \mathbf{p}_i \in \mathcal{P}_i$*

$$u_i(\mathbf{p}_i^*, \mathbf{p}_{-i}^*) \geq u_i(\mathbf{p}_i, \mathbf{p}_{-i}^*). \quad (3.12)$$

Dans le contexte du présent problème, cette définition signifie que les joueurs sont en équilibre de Nash dès lors qu'aucun d'entre eux ne peut dévier unilatéralement du profil d'équilibre  $\mathbf{p}^* = [\mathbf{p}_1^{*T}, \dots, \mathbf{p}_I^{*T}]^T$  sans que sa fonction d'utilité ne décroisse.

Le jeu  $\mathcal{G}$  étant infini, on ne peut pas appliquer directement le résultat du théorème de Nash [Nash, 1951] garantissant l'existence d'un équilibre en stratégie mixte dans tout jeu fini. Dans notre cas particulier, on peut en revanche faire appel au théorème de Debreu-Glicksberg-Fan, énoncé au Chapitre 2. On vérifie facilement que les espaces  $\mathcal{P}_i$  définis à l'équation (3.4) respectent toutes les hypothèses posées par le théorème 2.1.2, de même que les fonctions d'utilités  $u_i(\mathbf{p})$  de l'équation (3.11) sont continues en  $\mathbf{p}$  et concaves en  $\mathbf{p}_i$  (propriété plus forte que la quasi-concavité). L'existence d'un équilibre en stratégies pures, c'est à dire déterministes, est donc garantie. On notera qu'il n'y a pas besoin de rechercher des équilibres en stratégies mixtes pour le jeu que nous traitons ici, en vertu de la concavité de  $u_i(\mathbf{p})$  en  $\mathbf{p}_i$ . En effet, posons  $\pi = (\pi_1, \dots, \pi_I) \in \Delta\mathcal{P}$  où  $\pi_i \in \Delta\mathcal{P}_i$  est la stratégie mixte du joueur  $i$ . Ses vecteurs d'allocation de puissances sont alors aléatoirement distribués par  $\pi_i$  et son utilité espérée est donnée par

$$\mathbb{E}_\pi [u_i(\mathbf{p})] \leq \mathbb{E}_{\pi_{-i}} [u_i(\mathbb{E}_{\pi_i} [\mathbf{p}_i], \mathbf{p}_{-i})], \quad (3.13)$$

où on applique l'inégalité de Jensen, qui se vérifie avec égalité dès lors que  $\mathbf{p}_i$  est déterministe. Autrement dit, toute stratégie mixte est une stratégie dominée. Ceci étant vérifié  $\forall i \in \mathcal{I}$ , on peut donc se restreindre aux stratégies pures lors de la recherche d'un équilibre de Nash du jeu  $\mathcal{G}$ .

On introduit maintenant la notion de meilleure réponse :

**Définition 3.2.2.** *Soit le jeu  $\mathcal{G} = (\mathcal{I}, \mathcal{P}, (u_i)_{i \in \mathcal{I}})$  et un joueur  $i \in \mathcal{I}$ . Étant donné un profil de stratégies (pures) des adversaires de  $i$  noté  $\mathbf{p}_{-i} \in \mathcal{P}_{-i}$ , les meilleures réponses de  $i$  à  $\mathbf{p}_{-i}$  sont dans l'ensemble  $BR_i(\mathbf{p}_{-i})$  défini par*

$$BR_i(\mathbf{p}_{-i}) \triangleq \{\mathbf{p}_i^* \in \mathbb{R}_+^N : u_i(\mathbf{p}_i^*, \mathbf{p}_{-i}) \geq u_i(\mathbf{p}_i, \mathbf{p}_{-i}), \forall \mathbf{p}_i \in \mathcal{P}_i\}. \quad (3.14)$$

À l'équilibre de Nash, tous les joueurs de  $\mathcal{I}$  jouent donc une meilleure réponse aux meilleures réponses jouées par leurs adversaires, c'est à dire que le profil d'équilibre  $\mathbf{p}^*$  vérifie  $\mathbf{p}_i^* \in BR_i(\mathbf{p}_{-i}^*)$ . Dans le jeu de partage de spectre traité ici, avec la fonction

d'utilité donnée par (3.11), la meilleure réponse d'un joueur  $i$  quelconque à la stratégie  $\mathbf{p}_{-i}$  de ses adversaire est l'unique solution de

$$\begin{aligned} \mathbf{p}_i^* &= \operatorname{argmax}_{\mathbf{p}_i} u_i(\mathbf{p}_i, \mathbf{p}_{-i}) \\ &\text{sous la contrainte } \mathbf{p}_i \in \mathcal{P}_i \end{aligned} \quad (3.15)$$

et peut être obtenue en résolvant le système d'équations données par les conditions de Karush-Kuhn-Tucker associées au problème (3.15). Ceci conduit à la solution de water-filling donnée  $\forall i \in \mathcal{I}$ , et pour un  $\mathbf{p}_{-i}$  fixé, par

$$p_{i,n}^* = \left[ \lambda_i - \frac{1}{\gamma_{i,n}(\mathbf{p}_{-i})} \right]^+, \quad \forall n \in \mathcal{N}, \quad (3.16)$$

où  $[x]^+ = \max(0, x)$ ,  $\gamma_{i,n}(\mathbf{p}_{-i})$  est donné par l'équation (3.10), et  $\lambda_i$  est le multiplicateur de Lagrange devant être choisi de manière à saturer la contrainte de puissance (3.1), et est aussi appelé « niveau d'eau » dans ce contexte.

On notera  $\boldsymbol{\gamma}_{-i} = (\gamma_{i,n}(\mathbf{p}_{-i}))_{n \in \mathcal{N}}$ , où l'indice  $-i$  exprime que l'état du canal de  $i$  dépend de  $\mathbf{p}_{-i}$ , et on défini

$$\begin{aligned} WF_i[\boldsymbol{\gamma}_{-i}] &\triangleq BR_i(\mathbf{p}_{-i}) \\ &= \left( \left[ \lambda_i - \frac{1}{\gamma_{i,n}(\mathbf{p}_{-i})} \right]^+ \right)_{n=1}^N \end{aligned} \quad (3.17)$$

l'opérateur de water-filling du joueur  $i$  étant donné sa connaissance de  $\boldsymbol{\gamma}_{-i}$ .

Un équilibre de Nash du jeu  $\mathcal{G}$  est donc tout vecteur  $\mathbf{p}^*$  tel que tous les joueurs jouent une allocation de puissance  $\mathbf{p}_i^*$  qui vérifient l'ensemble d'équations de water-filling donnée par (3.16), étant donnée l'état statistique de leur canal respectif  $\boldsymbol{\gamma}_{-i}^* = (\gamma_{i,n}(\mathbf{p}_{-i}^*))_{n \in \mathcal{N}}$ .

### 3.2.2 Apprentissage décentralisé d'un équilibre

La connaissance du jeu et de ses conditions initiales (présence et nombre de sources d'interférence, actions possibles, etc) étant incomplète, un équilibre ne peut émerger de l'auto-introspection des joueurs impliquée par les hypothèses de connaissance commune et de rationalité. Dans cette situation, des équilibres peuvent cependant apparaître en tant que résultats d'un processus d'apprentissage au cours duquel les joueurs interagissent en jouant leurs meilleures réponses itérativement [Fudenberg and Levine, 1998]. Dans le contexte du jeu de partage de spectre, cela va consister à mettre en œuvre un algorithme de water-filling itératif et distribué parmi les joueurs.

Posons  $\mathcal{T} = \{1, 2, \dots\}$  un ensemble d'indices associés à des nombres réels représentant des temps croissants d'un pas quelconque. Á chacun de ces temps correspond la mise à jour du vecteur d'allocation de puissance d'un des joueurs. On note  $\mathcal{T}_i \subseteq \mathcal{T}$ , et pour tout  $t \in \mathcal{T}_i$  le joueur  $i$  joue sa meilleure réponse  $WF_i[\boldsymbol{\gamma}_{-i}(t)]$  par rapport à

sa connaissance la plus récente de l'état du canal au temps  $t$ . En toute généralité, les joueurs sont considérés asynchrones et les périodes de mises à jour peuvent être différentes d'un joueur à l'autre. On considère que dans l'état initial les émetteurs allouent leur puissance uniformément sur la bande de communication. Cette stratégie, bien que d'apparence naïve au regard du problème traité ici, correspond à ce qui est le plus souvent mis en œuvre dans le contexte ASM. On note  $\mathbf{p}(0)$  le profil de stratégies initial d'allocation uniforme. Successivement, et au fur et à mesure des périodes de mises à jour de chaque joueur, les allocations de puissance jouées sont sensées converger vers un partage du spectre non-orthogonal où les joueurs sont en équilibre de Nash. L'algorithme 2 formalise cette procédure d'apprentissage basé sur la meilleure réponse.

---

**Algorithme 2** Water-filling distribué

---

```

1:  $t = 0$ 
2:  $\forall i \in \mathcal{I}, \forall n \in \mathcal{N}, p_{i,n}(t) = P_i^{max}/N.$ 
3: pour  $t \in \mathcal{T}$  faire  $\forall i \in \mathcal{I}$ 
4:   si  $t \in \mathcal{T}_i$  alors
5:      $\mathbf{p}_i(t+1) = WF_i[\boldsymbol{\gamma}_{-i}(t)]$ 
6:   fin si
7:    $t = t + 1$ 
8: fin pour

```

---

Conformément aux hypothèse du paragraphe, la connaissance statistique de l'état du canal modélisée ici par  $\boldsymbol{\gamma}_{-i}(t)$  doit être fournie par le lien de retour du récepteur  $i$  vers son émetteur. Posons  $t \in \mathcal{T}_i$  le temps de mise à jour du joueur  $i$  et  $t' = \sup\{\tau \in \mathcal{T}_i : \tau < t\}$  le temps de la mise à jour précédente. Si le délai de propagation entre le récepteur  $i$  et son émetteur est  $\Delta_{ii}$ , alors le récepteur a retourné  $\boldsymbol{\gamma}_{-i}(t')$  et  $\boldsymbol{\gamma}_{-i}(t)$  aux temps  $t' - \Delta_{ii}$  et  $t - \Delta_{ii}$ , respectivement. Il est admis qu'entre ces deux dates, l'émetteur a transmis suffisamment de symboles pilotes pour que le récepteur soit en mesure de réaliser des estimations du canal  $\mathbf{H}_{ii}$  et de l'interférence plus bruit  $\mathbf{Z}_i$  (ou de la puissance de ces variables) afin de former la statistique  $\boldsymbol{\gamma}_{-i}$ .

En général, la convergence des algorithmes de meilleures réponses tels que l'algorithme 2 n'est pas garantie [Bacci et al., 2016]. De plus, selon les conditions du canal, plusieurs équilibres de Nash peuvent exister au sein du jeu  $\mathcal{G}$ . Les deux paragraphes suivants donnent des conditions suffisantes pour l'unicité d'un équilibre et pour la convergence de l'algorithme 2 vers celui-ci.

### 3.2.3 Unicité de l'équilibre et convergence de l'algorithme d'apprentissage

Des conditions suffisantes d'unicité de l'équilibre de Nash ont été données par [Scutari et al., 2008a] dans le cas de canaux déterministes et invariants dans le temps. Il est possible de suivre les mêmes raisonnements pour aboutir à des conditions analogues sur les statistiques des canaux dans le cadre du jeu formulé à la section 3.1. Let  $\mathbf{Q}_n \in$

$\mathcal{M}_{I \times I}(\mathbb{R})$  be defined as

$$[\mathbf{Q}_n]_{ij} \triangleq \begin{cases} \frac{\mathbb{E}[|H_{ji,n}^2|]}{e^{\mathbb{E}[\log(|H_{ii,n}|^2)]}} & \text{if } i \neq j \\ 0 & \text{otherwise} \end{cases} \quad (3.18)$$

**Théorème 3.2.1.** *L'équilibre de Nash du jeu  $\mathcal{G}$  est unique si*

$$\rho(\mathbf{Q}_n) < 1, \quad \forall n \in \{1, \dots, N\}, \quad (3.19)$$

où  $\mathbf{Q}_n$  est définie par (3.18) et  $\rho(\mathbf{Q}(n))$  est son rayon spectral (i.e., le supremum du module de ses valeurs propres).

La preuve de ce théorème est esquissée à l'annexe B, et le lecteur intéressé peut également se reporter à [Scutari et al., 2008a].

De même que pour la condition suffisante d'unicité, des conditions suffisantes de convergence de l'algorithme de water-filling itératif ont été données dans le cas de canaux déterministes et connus par [Scutari et al., 2008b,c]. Il est intéressant de noter que ces conditions sont les mêmes pour la version asynchrone de l'algorithme, qui a été présentée ci-dessus, que pour les versions synchrones. Soit la matrice  $\mathbf{Q}^{max} \in \mathcal{M}_{I \times I}(\mathbb{R})$  définie par

$$[\mathbf{Q}^{max}]_{ij} \triangleq \begin{cases} \max_n \frac{\mathbb{E}[|H_{ji,n}^2|]}{e^{\mathbb{E}[\log(|H_{ii,n}|^2)]}} & \text{if } i \neq j \\ 0 & \text{otherwise} \end{cases} \quad (3.20)$$

**Théorème 3.2.2.** *Si la condition suivante est satisfaite :*

$$\rho(\mathbf{Q}^{max}) < 1, \quad (3.21)$$

alors l'algorithme de water-filling itératif converge géométriquement vers l'unique équilibre de Nash de  $\mathcal{G}$ .

On note que la condition suffisante de convergence est également suffisante pour l'unicité de l'équilibre. La preuve de ce théorème est donnée en annexe C, et il est également suggéré au lecteur intéressé de se reporter aux articles [Scutari et al., 2008b,c] pour plus de détails.

Ces conditions ont une interprétation physique liée au niveau d'interférence perçu par chaque récepteur [Scutari et al., 2008a]. Si l'interférence est élevée, par exemple parce que les liens de communications sont trop proches les uns des autres, alors les deux conditions des théorèmes 3.2.1 et 3.2.2 sont fausses. Ainsi, plusieurs équilibres peuvent exister au sein du jeu et, pourvu que l'algorithme converge néanmoins, les joueurs peuvent atteindre l'un d'entre eux itérativement en jouant successivement des meilleures réponses. L'allocation de puissance initiale détermine alors quel équilibre est atteint lorsque le processus a convergé. On note d'après l'équation de water-filling (3.16) que,

1. le niveau d'eau d'un joueur est fixé par sa contrainte de puissance de sorte que l'on ait  $p_{i,n} + \frac{\sigma_i^2 + \sum_{j \neq i} \mathbb{E}[|H_{ji,n}|^2] p_{j,n}}{e^{\mathbb{E}[\log(|H_{ii,n}|^2)]}} = \lambda_i$  pour toute sous porteuse  $n \in \mathcal{N}_i$ , où  $\mathcal{N}_i$

est l'ensemble des sous porteuses sur lesquelles le joueur  $i$  alloue de la puissance défini par

$$\mathcal{N}_i = \left\{ n \in \mathcal{N} : \lambda_i \geq \frac{\sigma_i^2 + \sum_{j \neq i} \mathbb{E} [|H_{ji,n}|^2] p_{j,n}}{e^{\mathbb{E}[\log(|H_{ii,n}|^2)]}} \right\}. \quad (3.22)$$

2. si le niveau de bruit et d'interférence est trop élevé,  $\sigma_i^2 + \sum_{j \neq i} \mathbb{E} [|H_{ji,n}|^2] p_{j,n} > \lambda_i$  et le joueur  $i$  n'alloue aucune puissance sur la sous porteuse  $n$ .

Ainsi, lorsque le niveau d'interférence mutuelle augmente, les joueurs auront tendance à ne pas utiliser les sous-porteuses sur lesquelles un autre joueur est déjà présent. L'équilibre de Nash du jeu de partage de spectre tend donc vers une allocation d'accès multiple en fréquence orthogonal (FDMA), où chaque utilisateur du canal occupe seul une portion de la bande.

Jusqu'ici, il a été considéré que les émetteurs obtiennent une connaissance parfaite des statistiques de leur canal. La section suivante présente un jeu robuste où une connaissance erronée de  $\gamma_{-i}$  est considérée.

### 3.3 Incertitude quant aux statistiques du canal

La résolution du jeu de partage de spectre tel que présenté jusqu'ici requiert des joueurs une connaissance parfaite des statistiques de leur canal. Dans une situation pratique, les récepteurs auraient à les estimer avec des symboles pilotes insérés en début de chaque trame. Quel que soit l'estimateur de  $\gamma_{-i}$  mis en œuvre par le récepteur  $i$ , une erreur plus ou moins grande viendra entacher la mesure et la connaissance de l'état du canal est alors imparfaite.

Face aux erreurs d'estimation, il est néanmoins possible d'adopter une approche robuste similaire à celle présentée à la section 3.1.3. On va donc formuler ici un nouveau problème d'optimisation robuste contre la pire erreur d'estimation. Notons qu'un jeu robuste aux incertitudes sur le canal a été également proposé par [Anandkumar et al., 2011]. Cependant, un facteur limitant de ces travaux est l'hypothèse posée que les canaux d'interférences  $\mathbf{H}_{ji}$  pour  $j \neq i$  sont déterministes et connus des joueurs  $i$ . Étant donnée la présentation du contexte actuel faite en section 3.1, ceci n'est pas vérifié pour le problème de partage de spectre ASM traité ici.

Soit  $\hat{\gamma}_{-i} = (\hat{\gamma}_{i,n})_{n \in \mathcal{N}}$  où  $\hat{\gamma}_{i,n}$  est l'estimateur de l'état du canal  $\gamma_{i,n}(\mathbf{p}_{-i})$  du joueur  $i$  à la sous porteuse  $n$ . Cet estimateur est considéré comme une variable aléatoire continue de variance finie, et vérifie  $\forall n$

$$\mathbb{E}[\hat{\gamma}_{i,n}] = \gamma_{i,n}(\mathbf{p}_{-i}) \quad (3.23)$$

$$\sigma_{\hat{\gamma}_{i,n}}^2 \triangleq \mathbb{E} [(\hat{\gamma}_{i,n} - \gamma_{i,n}(\mathbf{p}_{-i}))^2]. \quad (3.24)$$

L'incertitude sur l'estimation de l'état statistique du canal est contenue dans la variance d'estimation  $\sigma_{\hat{\gamma}_{i,n}}^2$ . On considère que l'estimateur produit des erreurs bornées, et on définit l'ensemble des états statistiques du canal du joueur  $i$  pour lesquels l'erreur

d'estimation est bornée par

$$\Gamma_i \triangleq \left\{ \gamma_i \in \mathbb{R}_+^N \mid |\hat{\gamma}_{i,n} - \gamma_{i,n}| \leq \epsilon_{i,n} \quad \forall n \in \mathcal{N} \right\}, \quad (3.25)$$

où  $\epsilon_{i,n}$  est la borne d'incertitude à la sous porteuse  $n$ . L'hypothèse posée est raisonnable puisque l'incertitude sur une variable aléatoire quelconque peut être bornée avec une forte probabilité en vertu de l'inégalité de Bienaymé-Chebychev. Pour tout  $\epsilon_{i,n} \in \mathbb{R}_+$  et quelle que soit la distribution de probabilité de  $\hat{\gamma}_{i,n}$ , pourvu que sa variance soit finie, on a

$$\mathbb{P} [|\hat{\gamma}_{i,n} - \mathbb{E}[\hat{\gamma}_{i,n}]| \leq \epsilon_{i,n}] \geq 1 - \frac{\sigma_{\hat{\gamma}_{i,n}}^2}{\epsilon_{i,n}}. \quad (3.26)$$

L'incertitude sur l'estimation de l'état du canal peut donc être prise en compte dans le problème de partage de spectre en maximisant une utilité minimisée vis-à-vis de l'erreur d'estimation la plus préjudiciable. Le jeu robuste, noté  $\tilde{\mathcal{G}}$ , est alors défini par le triplet

$$\tilde{\mathcal{G}} = \{\mathcal{I}, \mathcal{P}, (\tilde{u}_i)_{i \in \mathcal{I}}\} \quad (3.27)$$

où les ensembles de joueurs et d'actions restent inchangés par rapport à  $\mathcal{G}$  et  $(\tilde{u}_i)_{i \in \mathcal{I}}$  sont de nouvelles fonctions d'utilité qu'on exprime au travers de la minimisation

$$\begin{aligned} \tilde{u}_i(\mathbf{p}_i, \mathbf{p}_{-i}) &\triangleq \min_{\gamma_i \in \Gamma_i} \alpha \sum_{n=1}^N \log(1 + \gamma_{i,n} p_{i,n}) \\ &= \alpha \sum_{n=1}^N \log(1 + [\hat{\gamma}_{i,n} - \epsilon_{i,n}]^+ p_{i,n}). \end{aligned} \quad (3.28)$$

La résolution de ce nouveau jeu par l'équilibre de Nash à travers le calcul des meilleures réponses amène à une solution que l'on appellera water-filling robuste, et qui s'exprime  $\forall i \in \mathcal{I}$  et pour un état du canal estimé fixé  $\hat{\gamma}_{-i}$  par

$$\tilde{p}_{i,n}^* = \left[ \lambda_i - \frac{1}{[\hat{\gamma}_{i,n} - \epsilon_{i,n}]^+} \right]^+, \quad \forall n \in \mathcal{N}. \quad (3.29)$$

On notera l'opérateur de water-filling dans le cas du jeu robuste par

$$WF_i[\tilde{\gamma}_{-i}] = \left( \left[ \lambda_i - \frac{1}{[\hat{\gamma}_{i,n} - \epsilon_{i,n}]^+} \right]^+ \right)_{n=1}^N \quad (3.30)$$

où  $\tilde{\gamma}_{-i} = ([\hat{\gamma}_{i,n} - \epsilon_{i,n}]^+)_{n \in \mathcal{N}}$ .

Le jeu robuste présenté ici peut s'interpréter comme une pénalisation de la fonction d'utilité originale  $u_i$  par un terme représentant l'incertitude sur l'état des sous-canaux  $n \in \mathcal{N}$ , la pénalité étant d'autant plus importante qu'allouer une puissance sur la sous-porteuse correspondante est incertain en terme de bénéfices apportés sur l'utilité. En d'autres termes, les joueurs ont tendance à sous-estimer la qualité des sous-canaux selon l'incertitude  $\epsilon_{i,n}$  paramétrée au préalable. Un tel mécanisme a pour effet de créer une "aversion au risque", et les force ainsi à n'allouer leur puissance

disponible que sur les sous-porteuses dont le rapport signal sur interférence plus bruit est le meilleur. Comme les résultats le montreront, ce genre de comportement peut avoir l'avantage de conduire les joueurs vers un partage de spectre plus efficace que l'équilibre de Nash du jeu  $\mathcal{G}$  sous connaissance parfaite de l'état du canal, car lorsqu'un joueur ne gaspille plus sa puissance sur un sous-canal de faible qualité, il génère ainsi moins d'interférences et le libère pour les joueurs adverses.

## 3.4 Études et résultats sur canaux synthétiques et sondés en mer

Cette section est dédiée à l'exposition de résultats d'évaluation de la méthode de partage de spectre décentralisée proposée dans ce chapitre. On présente dans un premier temps la méthode d'évaluation sur des canaux ASM simulés, puis un cas simple de jeu à deux joueurs afin de confirmer les comportements prévus par les études précédentes. On donne ensuite des résultats moyens où le nombre de joueurs et la puissance des interférences varient avant de présenter des résultats obtenus sur des canaux ASM sondés en mer.

### 3.4.1 Méthode d'évaluation

On rappelle qu'un ensemble de  $I$  couples émetteur-récepteur communiquant en OFDM constitue l'ensemble des joueurs  $\mathcal{I}$ . Les communications sont supposées prendre place en même temps sur une bande  $B = 6$  kHz centrée sur  $f_c = 12$  kHz. La bande est divisée en  $N = 256$  sous-canaux correspondant aux sous-porteuses OFDM, dont l'espacement et donc de  $\Delta f = 23.43$  Hz. La durée d'un symbole OFDM est de  $T = N\Delta f^{-1} + T_g = 57.7$  ms où  $T_g = 15$  ms est la durée du préfixe cyclique. Tous joueurs  $i \in \mathcal{I}$  sont contraints par une puissance moyenne de  $P_i^{max} = N$ . À l'état initial  $\mathbf{p}(0)$  de l'algorithme 2, la puissance des joueurs est uniformément distribuée sur la bande  $B$ . Un rapport signal à bruit (SNR) de référence est fixé à 20 dB pour une distance de transmission de 1 km. Lorsque la distance varie dans les simulations, le SNR pour le lien considéré varie également et selon les lois de pertes de propagation exposées au Chapitre 1. La densité spectrale de puissance de bruit est définie par le modèle de Knudsen donné également au Chapitre 1, équation (1.7). Les essais en mer étant très coûteux compte tenu du nombre de couples émetteur-récepteur qu'il faudrait opérer simultanément, les résultats sont obtenus par simulation des algorithmes proposés sur la base de canaux ASM synthétiques ou sondés en mer.

Les canaux de communications ASM sont synthétisés par un simulateur de canal décrit dans les travaux de [Socheleau et al., 2010a,b]. Ce simulateur produit des canaux à évanouissements de Rice et procède comme suit : Étant donnés des paramètres d'entrée donnant la description de la géométrie de la transmission (distances et profondeurs), une réponse impulsionnelle invariante dans le temps est construite à partir d'un modèle de tracé de rayons. Le gain en module au carré de ce canal est considéré comme le profil délai-puissance du futur canal aléatoire, dont les fluctuations

temporelles sont générées en maximisant l'entropie du spectre Doppler sous contrainte d'étalement fréquentiel et de facteur de Rice. Le facteur de Rice du trajet principal est donné comme paramètre d'entrée du simulateur et les ceux des trajets secondaires décroissent avec le retard de façon exponentielle. Pour les scénarios de simulation présentés ici, l'étalement fréquentiel est fixé à  $\sigma_D^2 = 1$  Hz et le facteur de Rice à  $K = 10$  dB. Le simulateur de canal intègre aussi des éléments des modèles physiques de propagation, tels que l'atténuation en fonction de la distance et les pertes par absorption données par la formule de Thorp à l'équation (1.3). Le simulateur génère donc des canaux  $h_{ji}(t, \tau)$  sélectifs en fréquence et variable aléatoirement dans le temps. Les coefficients fréquentiels utilisés pour simuler les réalisations des variables aléatoires  $H_{ji,n}$  pour  $n = \{1, \dots, N\}$  sont obtenus par transformation de Fourier discrète puis les statistiques  $\mathbb{E}[|H_{ji,n}|^2]$  et  $e^{\mathbb{E}[\log |H_{ji,n}|^2]}$  impliqués dans les fonctions d'utilités par l'état statistique du canal de l'équation (3.10). Des exemples de réponses fréquentielles variant dans le temps générées par le simulateur sont illustrés à la figure 3.1.

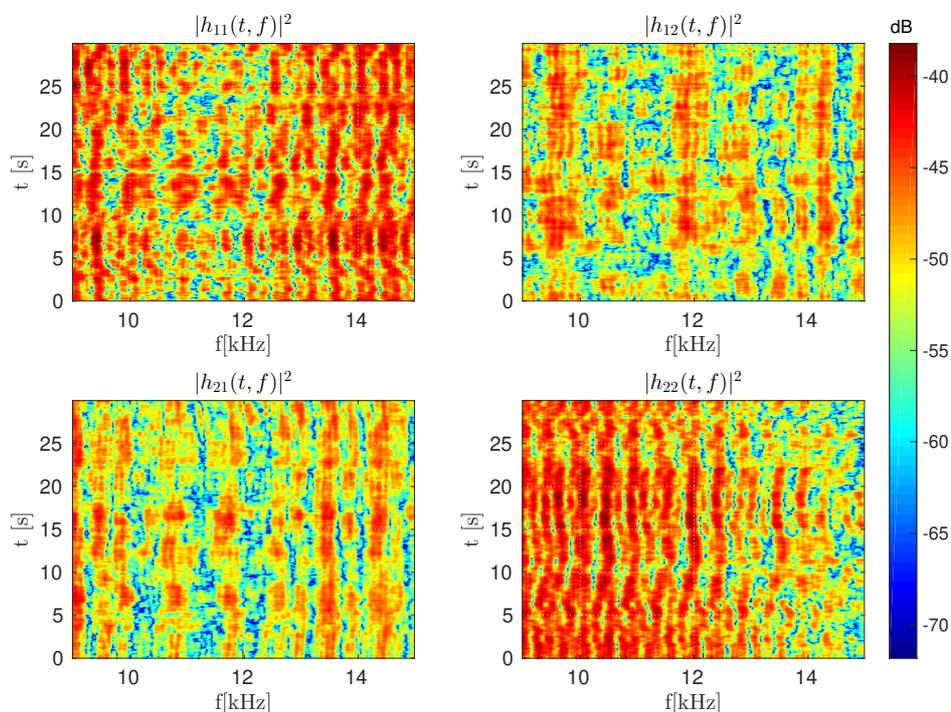


FIGURE 3.1 – Réponses fréquentielles variant dans le temps du simulateur de [Socheleau et al., 2010b].  $d_{ii} = 1$  km,  $d_{ji} = 1.5$  km, profondeurs : TX 1 = 10 m, RX 1 = 10 m, TX 2 = 5, RX 2 = 15 m.

L'efficacité du partage de spectre est évaluée en terme d'utilité atteinte après un certain nombre d'itérations de l'algorithme 2. L'utilité est elle-même évaluée sur la base de l'équation (3.11) quel que soit le jeu, avec pour vecteurs de puissance  $\mathbf{p} = (\mathbf{p}_i, \mathbf{p}_{-i})$  les résultats de l'algorithme de water-filling. Comme mentionné en section 3.1.3 et bien que les fonctions d'utilité soient exprimées et illustrées en bits par secondes et par Hertz (bps/Hz) car homogènes à une efficacité spectrale, ces quantités ne doivent pas être comprises comme des débits atteignables mais plutôt comme une mesure de la satisfaction que les joueurs retirent de leur partage du spectre. Cette interprétation correspond totalement à la signification qu'ont les fonctions d'utilité en théorie des jeux [Fudenberg and Tirole, 1991]. L'opérateur de water-filling est calculé sur la base

des équations (3.17), (3.30) ou

$$WF_i [\hat{\gamma}_{-i}] = \left( \left[ \lambda_i - \frac{1}{\hat{\gamma}_{i,n}} \right]^+ \right)_{n=1}^N \quad (3.31)$$

avec  $\hat{\gamma}_{-i} = (\hat{\gamma}_{i,n})_{n \in \mathcal{N}}$  la variable aléatoire des estimations de l'état du canal, sans prise en compte de l'erreur. On parlera respectivement de jeu avec connaissance parfaite ( $\mathcal{G}$ ), de jeu robuste ( $\hat{\mathcal{G}}$ ) et de jeu non robuste (water-filling avec l'équation (3.31)), ce dernier étant noté  $\tilde{\mathcal{G}}$  par convenance. Dans  $\mathcal{G}$ ,  $\gamma_{-i}$  est connu, tandis que dans  $\hat{\mathcal{G}}$  et  $\tilde{\mathcal{G}}$  cette quantité est estimée. Afin de ne pas restreindre les analyses à une méthode d'estimation particulière du canal, la procédure suivante est employée pour simuler les estimateurs  $\hat{\gamma}_{-i}$ . Pour tout sous-canal  $n$ , les  $\hat{\gamma}_{i,n}$  sont supposés non biaisés et leur variance  $\sigma_{\hat{\gamma}_{i,n}}^2$  peut donc être bornée inférieurement par la borne de Cramér-Rao calculée sur le paramètre  $\gamma_{i,n}(\mathbf{p}_{-i})$ . On note cette borne  $\mathbf{F}^{-1}(\gamma_{i,n})$ , où  $\mathbf{F}(\gamma_{i,n})$  est la matrice d'information de Fisher du paramètre estimé. Les  $\hat{\gamma}_{i,n}$  sont modélisés par des variables aléatoires gaussiennes de moyennes  $\gamma_{i,n}(\mathbf{p}_{-i})$  et de variances  $\sigma_{\hat{\gamma}_{i,n}}^2 = R \times \mathbf{F}^{-1}(\gamma_{i,n})$  où  $R$  est une constante positive, permettant de simuler des estimateurs de différentes performances. L'incertitude sur l'état du canal augmente donc avec  $R$ . Le paramètre  $\gamma_{i,n}(\mathbf{p}_{-i})$  étant nécessairement positif, la distribution gaussienne utilisée pour le modéliser est tronquée pour éviter les valeurs négatives de SINR estimé. La borne d'incertitude est calculée par  $\epsilon_{i,n} = k \times \sigma_{\hat{\gamma}_{i,n}}$ , avec  $k$  une constante non négative (pour  $k = 0$  on est donc dans le cas du jeu non robuste  $\tilde{\mathcal{G}}$ ).

L'apprentissage des équilibres de Nash aux moyens de l'algorithme 2 nécessite que les jeux  $\mathcal{G}$ ,  $\hat{\mathcal{G}}$  et  $\tilde{\mathcal{G}}$  soient répétés un certain nombre d'itérations. On considère une version séquentielle de l'algorithme 2, où les émetteurs adaptent leurs allocations de puissance les uns après les autres sur la base d'une période commune fixée, afin de simplifier les analyses des résultats. La version séquentielle n'est qu'un cas particulier de l'algorithme 2, et les mêmes conditions de convergence s'appliquent. De plus, il n'est pas nécessaire de considérer une synchronisation explicite des joueurs dans la mesure où l'on peut très bien considérer des systèmes ayant initié leurs transmissions respectives les uns après les autres. Dans ce cas, et sur la base d'une période commune de mise à jour, l'ordre de mise à jour des joueurs est donc déterminé. À chaque itération  $l$  de l'algorithme, seul le joueur  $l$  modulo  $I + 1$  joue son water-filling, sur la base de (3.17), (3.31) ou (3.30) selon le jeu joué. La période de mise à jour des stratégies de chaque joueur est considérée égale  $T_{obs} = 30$  secondes. Cette période correspond au temps durant lequel les récepteurs calculent les statistiques de l'interférence et du canal avant de les transmettre à leurs émetteurs respectifs. Ainsi, si un joueur  $i$  donné joue à l'itération  $l$  correspondant à un temps  $t_i \in \mathcal{T}_i$ , alors son prochain tour sera joué à l'itération  $l + I$  qui correspond au temps  $t_i + T_{obs}$ .

### 3.4.2 Étude d'un jeu à 2 joueurs

On va se concentrer ici sur deux scénarios simples de jeux à deux joueurs,  $\mathcal{I} = \{1, 2\}$ , où on fera varier la puissance de l'interférence perçue par les récepteurs d'un jeu à l'autre pour évaluer les différences de comportement engendrées. Le canal est considéré petits fonds et la profondeur d'eau est fixée à 50 mètres. La profondeur des

terminaux est arbitrairement fixée entre 5 et 20 mètres et la distance entre les émetteurs et leur récepteur correspondant est fixée à  $d_{ii} = 1$  km.

### 3.4.2.1 Forte interférence

Afin de simuler un cas où les deux liens de communication sont en situations d'interférence, on fixe  $\forall i, j \in \mathcal{I}, j \neq i$ , la distance entre l'émetteur  $i$  et le récepteur  $j$  à  $d_{ji} = 1.5$  km. La perte de propagation entre un émetteur interférant et un récepteur donné est de 4 dB de moins que celle subie depuis l'émetteur d'intérêt. Les canaux simulés utilisés pour modéliser les  $\mathbf{H}_{ji}$  sont ceux de la figure 3.1. Les variances d'estimation sont fixées  $\forall i \in I$  et  $\forall n = 1, \dots, N$  à  $\sigma_{\hat{\gamma}_{i,n}}^2 = R \times \mathbf{F}^{-1}(\gamma_{i,n})$  avec  $R = 4$ . Les bornes d'incertitude sont fixées  $\forall i$  et  $\forall n$  à  $\epsilon_{i,n} = k \times \sigma_{\hat{\gamma}_{i,n}}$  avec  $k = 3$ . Le choix de  $k$  garanti que l'erreur d'estimation est bornée par  $\epsilon_{i,n}$  avec une probabilité supérieure à 0.88, quelle que soit sa distribution. Pour le modèle gaussien que a été choisit ici, cette probabilité est environ égale à 0.997. Le calcul des fonctions d'utilité, itération après itération des jeux non robustes et robustes, est le résultat d'une moyenne empirique des utilités obtenues pour 100 tirages indépendants des variables  $\hat{\gamma}_{-i}$ .

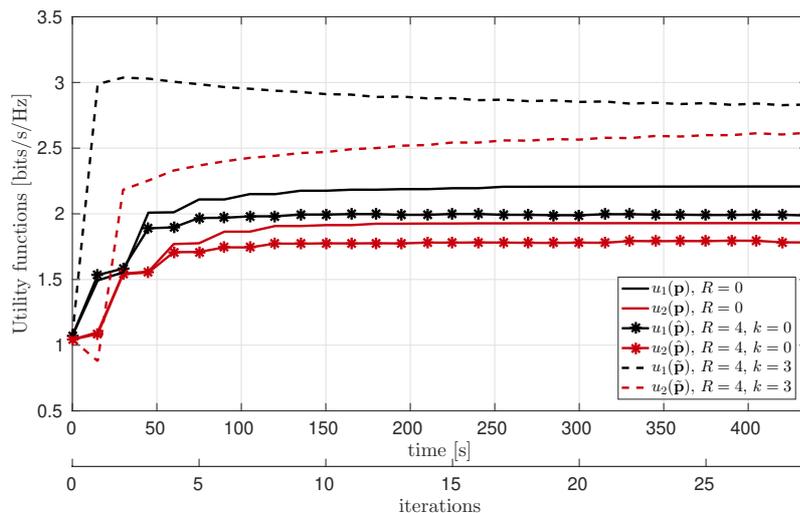


FIGURE 3.2 – Fonctions d'utilité - Forte interférence,  $d_{ji}/d_{ii} = 1.5$ ,  $d_{ii} = 1$  km.

La figure 3.2 illustre les utilités des deux joueurs en fonction des itérations des jeux  $\mathcal{G}$ ,  $\hat{\mathcal{G}}$  ( $R = 4$ ,  $k = 0$ ) et  $\tilde{\mathcal{G}}$  ( $R = 4$ ,  $k = 3$ ). À l'itération  $l = 0$  correspond l'utilité obtenue pour l'allocation de puissance uniforme. Une amélioration est clairement visible au cours du temps pour les deux joueurs, qui multiplient leur utilité initiale par plus de 2 dans le jeu avec connaissance parfaite de l'état du canal. Les 95% de l'utilité atteinte à la convergence sont atteints avant la 6<sup>ème</sup> itération. Considérant la période de mise à jour de  $T_{obs} = 30$  secondes dans une version séquentielle de l'algorithme 2, cela correspond à un temps de moins de 1 min 30. Le temps de convergence peut bien sur être réduit en considérant une période de mise à jour moins longue. Ce ne peut être cependant qu'au prix d'une plus grande variance d'estimation, et du risque de voir les statistiques estimées se décorrélérer de la réalité du canal perçue par le récepteur au moment de la mise à jour de l'émetteur, puisque  $T_{obs}$  correspond au temps d'intégration

de ces statistiques.

Dans cette configuration de forte interférence mutuelle, les conditions d'unicité de l'équilibre de Nash et de convergence de l'algorithme de water-filling données par (3.19) et (3.21) ne sont pas vérifiées. Néanmoins, et en se rappelant que ces conditions sont suffisantes mais non nécessaires, on observe qu'il est clairement possible d'atteindre des stratégies de partage de spectre efficace en cas de fortes interférences.

Dans le jeu non robuste, les joueurs atteignent des utilités plus faibles que dans le jeu à connaissance parfaite de l'état du canal. Ceci est dû à un biais d'estimation s'expliquant par la troncature de la distribution gaussienne utilisée pour les  $\hat{\gamma}_{-i}$  associée à la forte variance d'estimation causée par l'interférence,  $\sigma_{\hat{\gamma}_{i,n}}^2$  étant dépendant de la puissance du bruit et de l'interférence au travers du calcul de sa borne de Cramer-Rao. Le jeu robuste, en revanche, produit des utilités supérieures au jeu à connaissance parfaite. Ce résultat, bien qu'allant contre l'intuition commune, peut s'expliquer par le comportement conservateur des joueurs vis-à-vis des sous-canaux incertains et des puissances qui leur sont allouées. Ce comportement particulier est plus clairement visible aux figures illustrant les profils d'allocation de puissance  $\mathbf{p}^*$  à l'équilibre de Nash pour les jeux  $\mathcal{G}$ ,  $\hat{\mathcal{G}}$  et  $\tilde{\mathcal{G}}$  respectivement. Les niveaux de puissance alloués sur les sous-porteuses sont donnés en pourcentage de la puissance moyenne disponible (à l'allocation uniforme initiale, chacune des sous-porteuses se voit allouer 0.39% de  $P_i^{max}$ ).

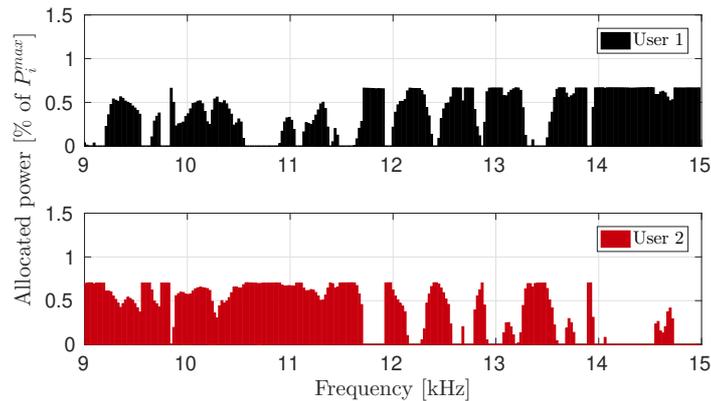


FIGURE 3.3 – Profil d'allocation de puissance à l'équilibre - Jeu à connaissance parfaite, forte interférence.

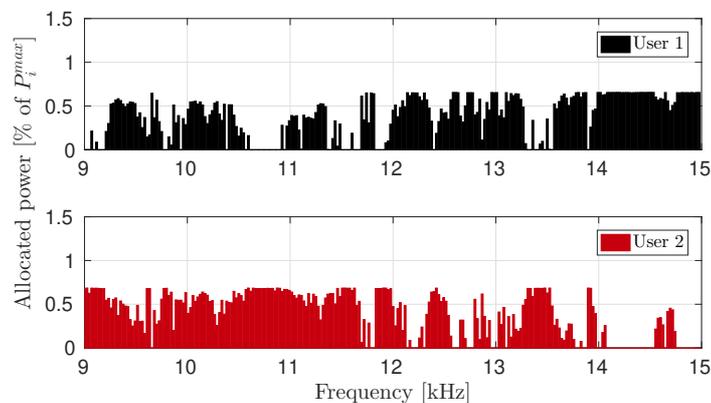


FIGURE 3.4 – Profil d'allocation de puissance à l'équilibre - Jeu non robuste, forte interférence.

Une certaine consistance entre les allocations des joueurs et les gains de leurs canaux (figure 3.1) peut être observée, en particulier dans le jeu avec connaissance parfaite. Par exemple, le joueur 2 a logiquement plus d'intérêt à concentrer sa puissance sur les sous-canaux des fréquences basses étant donné que le gain du canal  $h_{22}(t, f)$  y est le meilleur. Du plus, comme le canal  $h_{11}(t, f)$  du joueur 1 n'offre qu'un faible gain autour des 11 kHz, ce dernier est peu incité à utiliser les sous-porteuses de cette bande et génère ainsi moins d'interférence au joueur 2. L'allocation du joueur 1 sur les porteuses allant de 9 à 10 kHz est expliquée par l'interférence générée par le joueur 2 au travers du canal  $h_{21}(t, f)$ . Les résultats du jeu non robuste sont similaires, néanmoins, les allocations de puissances sont moins corrélées aux canaux du fait des biais et des erreurs d'estimation. On voit en revanche clairement, dans les allocations du jeu robuste, le comportement conservateur des joueurs du fait de la pénalisation de leurs fonctions d'utilité. Ils sont moins enclins à utiliser les canaux rendus incertains par l'interférence mutuelle, et leur stratégie de partage de spectre s'approche d'une division orthogonale de la bande de fréquences (FDMA). Ce fait est également pointé par [Lesherm and Zehavi, 2009; Scutari et al., 2008a].

Il est intéressant de noter que, sous l'hypothèse que l'interférence mutuelle entre utilisateurs d'un même canal est suffisante forte, on peut prouver que le partage de spectre FDMA est optimal au sens de la somme des débits des utilisateurs [Hayashi and Luo, 2009]. Dans le travail de [Hayashi and Luo, 2009], l'algorithme de partage de spectre proposé permet de trouver un partage de spectre FDMA sous-optimal, connaissant le gain du canal de chaque couple émetteur-récepteur, et suppose donc une forme de centralisme ou de coopération. Au contraire, dans le travail de ce chapitre, on montre qu'il est possible de s'approcher de cette solution FDMA de manière décentralisée, avec des systèmes complètement autonomes et disposant d'une connaissance très limitée de leur environnement.

### 3.4.2.2 Faible interférence

Dans cette seconde configuration, la distance entre les récepteurs et les émetteurs indésirables est augmentée à  $d_{ji} = 3$  km, la distance de transmission de chaque couple émetteur-récepteur restant égale à  $d_{ii} = 1$  km. De nouvelles réponses de canal sont générées avec ces nouveaux paramètres. La puissance de l'interférence perçue par les

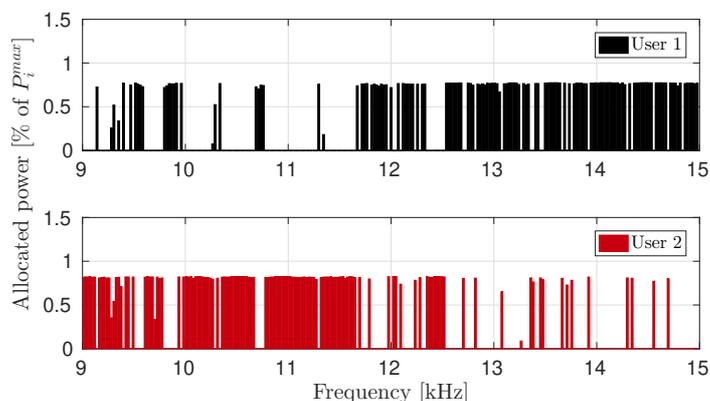


FIGURE 3.5 – Profil d'allocation de puissance à l'équilibre - Jeu robuste, forte interférence.

récepteur est environ 10 dB inférieure à la puissance reçue depuis l'émetteur d'intérêt. Les jeux sont simulés avec les mêmes paramètres que précédemment. Les résultats en termes d'utilité en fonction des itérations de l'algorithme 2 sont illustrés à la figure 3.6

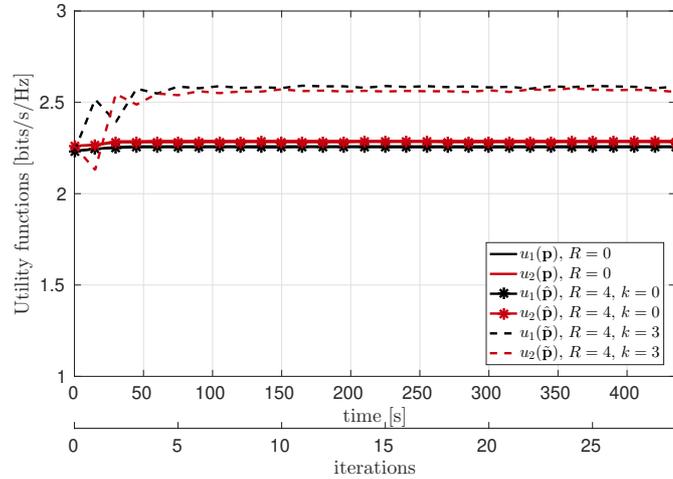


FIGURE 3.6 – Fonctions d'utilité - Faible interférence,  $d_{ji}/d_{ii} = 3$ ,  $d_{ii} = 1$  km.

L'interférence étant diminuée du fait d'une plus grande distance séparant les joueurs, ces derniers sont moins enclins à prendre l'interférence mutuelle en ligne de compte lors du partage du spectre. Les gains obtenus par rapport à l'allocation uniforme à travers ce partage, en termes d'utilité, sont donc amoindris. Néanmoins, les bénéfices apportés par le comportement des joueurs dans le jeu robuste sont toujours visibles. On peut cependant s'attendre à ce que ce bénéfice diminue également à mesure que l'interférence prend moins d'importance dans le partage du spectre.

### 3.4.3 Étude générale

Les bénéfices attendus de la méthode de partage de spectre décentralisée proposée sont évalués ici dans un cadre plus général. Les performances sont évaluées en terme d'utilité moyenne par joueurs, et on évalue l'influence des divers paramètres du problème : puissance de l'interférence, nombre de joueurs, qualité de l'estimateur de canal. On considère des transmissions dans une zone de rayon défini, variant entre 500 m et 8 km. La profondeur de la colonne reste fixée à 50 mètres. Pour chacune des mesures de performance réalisées, l'utilité moyenne par joueur a été calculée de la façon suivante :

1. On tire aléatoirement, et selon une distribution uniforme, la configuration géométrique de la transmission en terme de distances de transmission ( $d_{ji}$ ,  $i, j \in \mathcal{I}$ ) et de profondeurs d'immersions.
2. Pour une réalisation de la géométrie de transmission, on génère les canaux  $\mathbf{H}_{ji}$  à partir du simulateur selon la procédure expliquée au paragraphe 3.4.1.
3. L'algorithme 2 initialisé avec l'allocation uniforme et sur la base des canaux  $\mathbf{H}_{ji}$ . On considère que l'équilibre de Nash est atteint lorsque  $\forall i \in \mathcal{I}$ , jouer une meilleure réponse n'apporte pas un gain supérieur à un certain  $\eta > 0$ . Ceci

correspond à la notion d'équilibre  $\eta$ -Nash [Leyton-Brown and Shoham, 2009], et on fixe  $\eta = 10^{-2}$ .

4. L'utilité des joueurs est calculée pour cette configuration géométrique particulière, et on répète la procédure à partir de l'étape 1 et pour 400 réalisations.
5. L'utilité de chaque joueur est moyennée sur ces 400 réalisations puis on calcule l'utilité moyenne par joueur pour l'affichage des résultats.

La distance minimale séparant deux terminaux est fixée à 250 mètres. Le SNR de référence reste fixé à 20 dB à 1 km et les paramètres de facteur de Rice  $K_{max}$  et d'étalement Doppler  $\sigma_D^2$  sont égaux à 10 dB et 1 Hz, comme dans les simulations des cas à deux joueurs. Enfin, les utilités sont également moyennées sur les réalisations des  $\hat{\gamma}_{-i}$  pour les jeux non robustes et robustes.

Pour comparer notre travail avec des stratégies de partage de spectre utilisées en pratique dans le canal ASM, les fonctions d'utilités  $u_i(\mathbf{p})$  sont évaluées également pour des profils d'actions obtenus par

- le water-filling dans un contexte mono-utilisateur, sans interférence (single user).
- le partage de spectre FDMA, sur la base de l'algorithme centralisé de [Hayashi and Luo, 2009, Algorithm 2] supposant la connaissance des canaux  $\mathbf{H}_{ii}$  de tous les joueurs.
- la stratégie naïve où tous les joueurs utilisent l'allocation uniforme.

### 3.4.3.1 Influence de la puissance de l'interférence

On fixe le nombre de liens de communication à  $I = 2$  et on évalue l'utilité atteinte à l'équilibre  $\eta$ -Nash pour les différents jeux, en fonction du rayon de la zone dans laquelle sont immergés aléatoirement les terminaux. La superficie de la zone détermine alors la puissance moyenne d'interférence perçue par les joueurs.

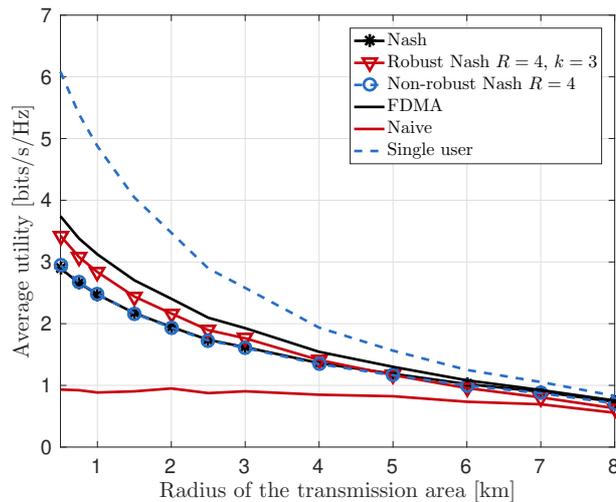


FIGURE 3.7 – Utilité moyenne par joueur en fonction du rayon de la zone d'immersion des terminaux. Deux joueurs,  $SNR = 20$  dB à 1 km.

Les résultats sont illustrés à la figure 3.7. Les avantages de l'allocation de puissance basée sur le water-filling sont notables, comparés à l'allocation uniforme naïve,

ce qui remet en question les conclusions avancées par [Aval et al., 2015] et qui ont été présentées au Chapitre 1 de ce manuscrit (section 1.4.1). Qu'il soit FDMA ou basé sur le water-filling, les bénéfices apportés par le partage du spectre diminuent à mesure que la superficie de la zone d'immersion augmente et on tend à s'approcher des performances du cas mono-utilisateur. Ceci était attendu et déjà observé à la section 3.4.2 : à mesure que l'interférence diminue, il est de moins en moins nécessaire pour les joueurs  $i \in \mathcal{I}$  d'adapter leurs allocations de puissance à autre chose que les canaux  $\mathbf{H}_{ii}$  et le bruit ambiant. Comme prévu précédemment, on voit également que le jeu robuste produit des utilités s'approchant du cas FDMA, du fait de l'aversion au risque des joueurs. Cet avantage est moins présent lorsque l'interférence diminue et, pour les distances les plus longues, il est même plus intéressant d'utiliser l'estimateur de canal sans prendre l'erreur en compte (jeu non robuste). Ce dernier point est vrai dès lors que l'estimateur est non biaisé, ainsi les performances moyennes s'approchent de l'équilibre de Nash calculé avec une connaissance parfaite de l'état statistique du canal.

### 3.4.3.2 Influence de la qualité de l'estimation

Afin de mettre en évidence l'impact de la qualité de l'estimation, on joue des jeux robustes  $\tilde{\mathcal{G}}$  pour plusieurs superficies de la zone de transmission comme précédemment, mais en faisant varier la variance de l'erreur d'estimation  $\sigma_{\hat{\gamma}_{i,n}}^2 = R \times \mathbf{F}^{-1}(\gamma_{i,n})$  à travers le paramètre  $R$ .

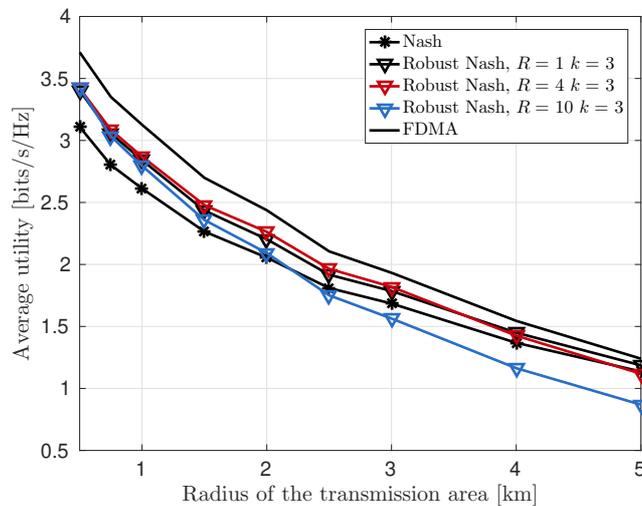


FIGURE 3.8 – Influence de la qualité de l'estimateur  $\hat{\gamma}_{-i}$  dans le jeu robuste. Deux joueurs,  $SNR = 20$  dB à 1 km.

Sur la figure 3.8, on peut voir que l'estimateur de variance minimale n'apporte pas forcément les meilleures utilités. Ceci semble, au premier abord, aller contre l'intuition. Cependant, on peut l'expliquer avec le comportement des joueurs induit par la pénalisation de leur fonction d'utilité : l'incertitude sur le canal augmente avec la variance de l'erreur d'estimation, si bien que les joueurs sont toujours plus réticents à allouer de la puissance sur les canaux les moins bons. Étant donné que la variance d'estimation augmente aussi avec l'interférence, qui est traitée comme du bruit, ce phénomène a logiquement tendance à être observé dans les zones de transmission de

petite superficie. Cependant, et comme les résultats le suggèrent, il existe évidemment un seuil à partir duquel l'augmentation de l'incertitude sur le canal n'apporte aucun bénéfice - et devient même préjudiciable.

### 3.4.3.3 Influence du nombre de joueurs

Les jeux  $\mathcal{G}$ ,  $\hat{\mathcal{G}}$  et  $\tilde{\mathcal{G}}$  sont joués maintenant avec un nombre variable de joueurs, et on évalue l'utilité moyenne par joueur en fonction de leur nombre pour une zone d'immersion de superficie fixée à  $1.5\text{km}$ . Les utilités sont toujours calculées sur la base de 400 jeux dont les profondeurs des terminaux et les distances les séparant sont tirées aléatoirement selon les modalités décrites au début de ce paragraphe. Le nombre d'utilisateur varie de 2 à 8, on fixe les paramètres des jeux non robustes et robustes à  $\sigma_{\hat{\gamma}_{i,n}}^2 = 4 \times \mathbf{F}^{-1}(\gamma_{i,n})$  et  $\epsilon_{i,n} = 3 \times \sigma_{\hat{\gamma}_{i,n}}$ , les autres paramètres conservant leurs valeurs des sections précédentes.

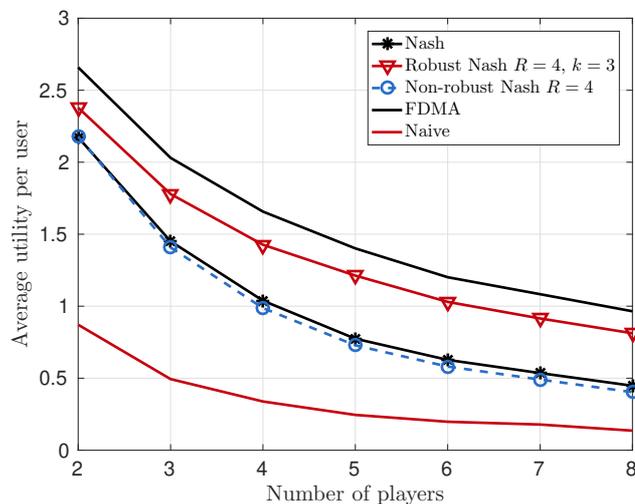


FIGURE 3.9 – Utilité moyenne par joueur en fonction du nombre de joueur dans les jeux  $\mathcal{G}$ ,  $\hat{\mathcal{G}}$  et  $\tilde{\mathcal{G}}$ , ainsi que pour l'allocation FDMA et uniforme (Naive). Les terminaux sont aléatoirement immergé dans une zone de rayon 1.5 km et de 50 m de profondeur. Le  $SNR$  est de 20 dB à 1 km.

Les résultats sont illustrés à la figure 3.9. Les utilités diminuent de manière logique lorsque le nombre de joueurs se partageant le même canal augmente. On voit néanmoins une nouvelle fois que la méthode proposée dans le jeu robuste offre des gains d'utilités importants, non seulement par rapport à l'allocation uniforme, mais aussi par rapport à l'équilibre de Nash du jeu  $\mathcal{G}$ , en s'approchant de la performance du FDMA.

## 3.4.4 Résultats sur canaux sondés en mer

### 3.4.4.1 Conduite de l'expérimentation

On se propose ici d'évaluer la méthode de partage de spectre proposée sur la base de canaux ASM sondés en mer, dans la baie de Toulon en Juillet 2015. Un transducteur

et trois hydrophones ont été utilisés durant ce sondage, où le transducteur a été déplacé à trois positions différentes. Les trois positions du transducteur et les trois hydrophones correspondent à paires émetteur-récepteur dans le futur jeu. Les hydrophones ont été immergés à 4, 12 et 16 m de profondeur, et le transducteur a été immergé à 10 m. La configuration est dépeinte à la figure 10. Les canaux sondés aux différentes positions successives du transducteur correspondent aux canaux  $\mathbf{H}_{ji}$ ,  $j, i = 1, 2, 3$ , intervenant dans le jeu de partage de spectre. Les  $SNR$  sans interférence pour chaque lien du TX au RX correspondant sur la figure étaient d'approximativement de 15, 12 et 9 dB en partant de la position la plus proche du transducteur.

Les mesures des réponses impulsionnelles de canal ont été obtenues par filtrage adapté successif au récepteur d'un signal sonde connu, transmis de manière répétée et ininterrompue pendant 25 secondes. Le signal sonde a été construit à partir d'une séquence binaire de longueur maximale de 511 symboles, modulés en BPSK et transmis à une rapidité de modulation de 8.7 kbits/s. Cette séquence peut capturer des arrivées retardées jusqu'à 58 ms, et les estimations du canal sont mises à jour 17 fois par secondes. La mesure a été réalisée à une fréquence porteuse  $f_c = 10.5$  kHz. Les décalages et compressions/dilatations Doppler ont été compensées par la procédure itérative présentée dans [Socheleau et al., 2015; van Walree et al., 2008a]. Le gain de traitement du signal sonde, étant donnés les 511 symboles de la séquence, est de 27 dB. On considère les canaux mesurés comme la "vérité terrain" du point de vue des jeux qui vont être simulés sur la base de leurs statistiques. En d'autres termes, la connaissance parfaite de l'état du canal se traduit par la connaissance des statistiques mesurées lors du sondage de canal, bien que cette mesure puisse être entachée d'erreur vis-à-vis du canal réel. Les réponses fréquentielles des neuf canaux sondés, sur une bande 6 kHz centrée sur  $f_c$ , sont illustrées à la figure 3.11.

#### 3.4.4.2 Simulation du jeu de partage de spectre

Le jeu de partage de spectre pour trois liens OFDM est simulé sur la base des canaux illustrés à la figure 3.11 et des équations de water-filling (3.17), (3.31) et (3.30). La bande de 6 kHz centrée autour de  $f_c$  est divisée en  $N = 256$  sous-porteuses. L'espacement entre sous-porteuses, la durée symbole et du préfixe cyclique sont les mêmes que dans les simulations sur canaux synthétiques. La contrainte de puissance est fixée à  $P_i^{max} = N$  et l'allocation initiale est la répartition uniforme de cette puissance sur la bande. La période de mise à jour et d'estimation de l'état statistique du canal est

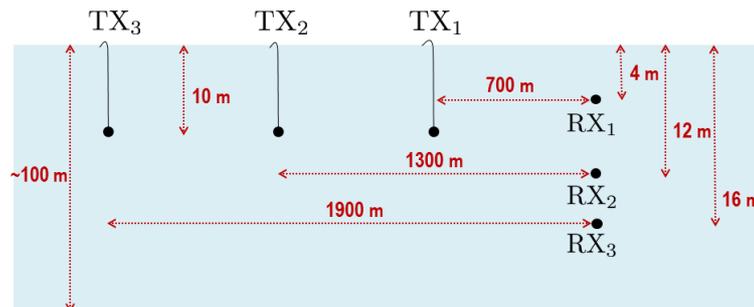


FIGURE 3.10 – Configuration du sondage de canal.

$T_{obs} = 25$  s. Pour les jeux non robustes et robustes, la variance d'estimation est fixée à  $\sigma_{\hat{\gamma}_{i,n}}^2 = 4 \times \mathbf{F}^{-1}(\gamma_{i,n})$  et  $\epsilon_{i,n} = 2 \times \sigma_{\hat{\gamma}_{i,n}}$ . L'estimateur produit des erreurs bornées par  $\epsilon_{i,n}$  avec une probabilité de 0.954. Les jeux sont lancés pour 30 itérations, et aucune garantie n'est assurée pour l'unicité des équilibres de Nash ou la convergence de l'algorithme.

La dernière allocation de puissance du jeu robuste est donnée à la figure 3.12. Les stratégies employées par les joueurs présentent une certaine cohérence avec les canaux de la figure 3.11. D'une part, le joueur 1 bénéficie de sa proximité avec son récepteur, et ainsi est plus « libre » de choisir ses meilleures sous-porteuses. Le joueur 3 au contraire, a tendance à jouer plutôt les fréquences basses de la bande, en évitant les sous-porteuses où se sont placés les joueurs 1 et 2. Le joueur 2 a le désavantage de trouver ses meilleurs gains de canal aux sous-porteuses où à la fois  $h_{11}(t, f)$  et  $h_{21}(t, f)$  sont élevés : celles-ci sont donc utilisées par le joueur 1 qui génère une forte interférence au récepteur 2. Néanmoins, au vu des canaux  $h_{21}(t, f)$  et  $h_{31}(t, f)$ , le joueur 1 perçoit le plus d'interférence de la part des joueurs 2 et 3 sur la bande allant d'environ 8 à 11 kHz,

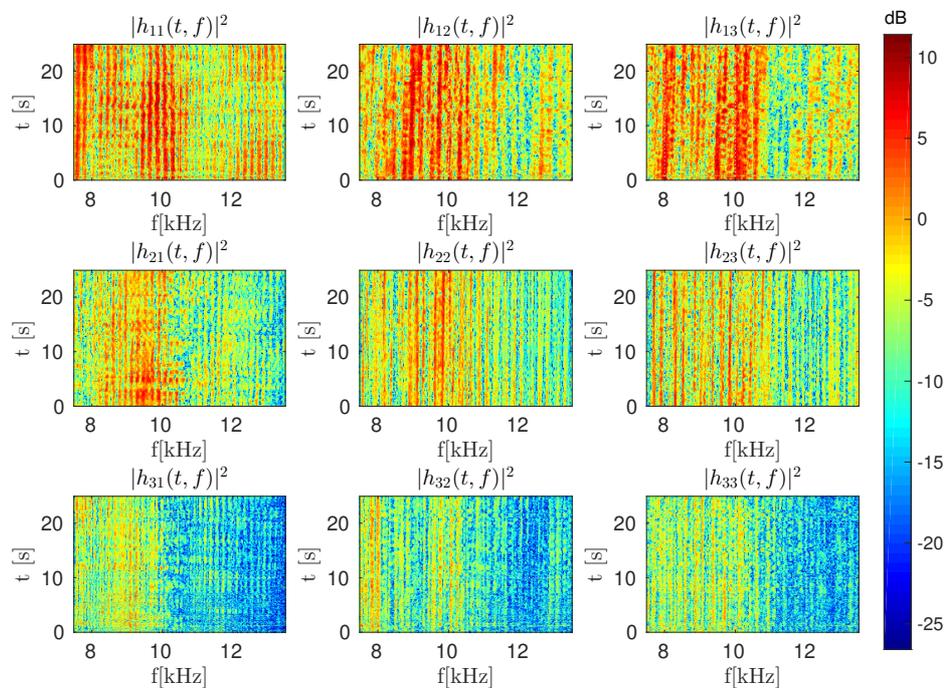


FIGURE 3.11 – Réponses fréquentielles des canaux ASM mesurés.

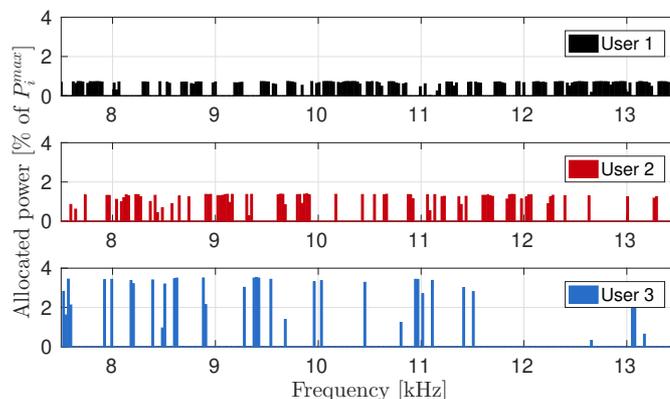


FIGURE 3.12 – Allocation de puissance à la dernière itération du jeu robuste, canaux réels.

ce qui laisse l'opportunité à ces derniers d'utiliser des sous-porteuses de cette bande. Ceci est particulièrement vrai dans le jeu robuste, compte tenu du comportement des joueurs face à l'incertitude.

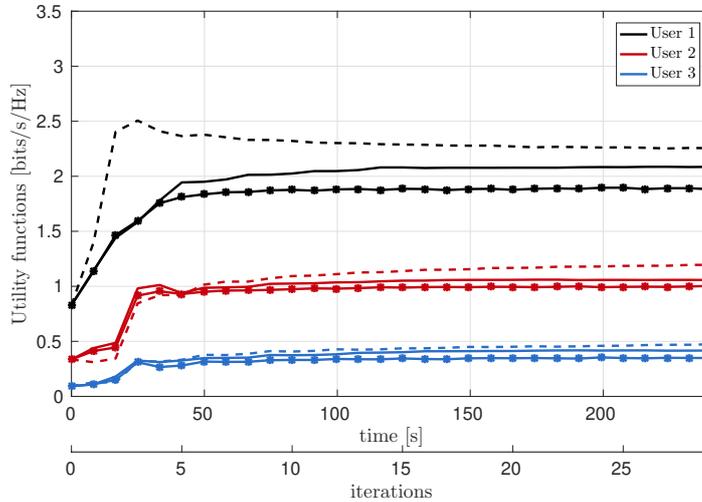


FIGURE 3.13 – Fonctions d'utilité des joueurs, canaux réels. En trait plein : jeu  $\mathcal{G}$ , avec les marqueurs '\*' : jeu  $\hat{\mathcal{G}}$  avec  $R = 4$ , en pointillé : jeu  $\tilde{\mathcal{G}}$  avec  $R = 4$  et  $k = 2$ .

Les utilités au cours des itérations sont affichées à la figure 3.13. À la fin du jeu, les joueurs ont multiplié leurs utilités par au moins 2, en comparaison de l'allocation uniforme. Les erreurs d'estimation semblent moins préjudiciables aux joueurs 2 et 3, ce qui peut s'expliquer par leur situation désavantageuse du fait de la géométrie de la transmission. Dans ce cas particulier, tous les joueurs ont amélioré leur utilité grâce au jeu robuste. Cela confirme que cette méthode est capable d'atteindre des points de la région d'utilité supérieurs à l'équilibre de Nash au sens de Pareto, sans coopération entre joueurs, ni échange d'information ou contrôle centralisé.

### 3.4.5 Aspects pratiques pour une mise en œuvre en situation réelle

Il doit être mis en exergue que le temps de convergence de l'algorithme 2, la stationnarité du canal ASM et la qualité d'estimation de l'état statistique du canal (en terme de variance *et* de biais) sont des paramètres pouvant être mis en relation à travers le choix de la période  $T_{obs}$ . On rappelle que cette période correspond à la fois à la période de mise à jour des joueurs et à la durée durant laquelle les réalisations estimées du canal sont agrégées par les récepteurs pour former l'état statistique du canal  $\gamma_{-i}$  (équation (3.10)). Évidemment, les joueurs pourraient tous avoir un  $T_{obs}$  différent et mettre en œuvre un algorithme asynchrone plutôt que la version séquentielle utilisée dans les simulations, mais cela ne change pas les remarques qui vont être soulevées ici.

De manière générale, le temps de stationnarité du canal doit être plus grand que le temps de convergence des algorithmes calculant l'équilibre de Nash du jeu de partage de spectre. On notera, d'ailleurs, à la lecture de [Bacci et al., 2016], que les

algorithmes d'apprentissage basés sur les dynamiques de meilleures réponses, tels que le water-filling distribué de l'algorithme 2, sont parmi ceux offrant les convergences les plus rapides. Lorsque cette condition sur le temps de convergence est vérifiée, il est alors attendu que la nature itérative de l'algorithme permette de suivre les variations lentes des statistiques du canal. Supposant que le canal est bien stationnaire, augmenter la période  $T_{obs}$  a pour effet de réduire la variance d'estimation des statistiques du canal, car on peut les calculer sur un plus grand nombre de réalisations. Cependant ceci est garanti lorsque l'estimateur est non-biaisé. Dès lors que le canal est non stationnaire sur la période  $T_{obs}$ , l'estimation des  $\gamma_{i,n}$  peut présenter un biais avec lequel l'erreur d'estimation augmente de manière quadratique. On s'attend néanmoins à ce que le jeu robuste puisse lutter contre ce type d'erreur, dans une certaine mesure, au vu des résultats présentés dans le jeu à deux joueurs sur canaux synthétiques. En effet, dans l'exemple de jeu de la section 3.4.2 où l'interférence mutuelle était forte, un biais d'estimation est artificiellement introduit par la modélisation des  $\hat{\gamma}_{i,n}$  par des variables gaussiennes tronquées, et le jeu robuste reste pourtant efficace.

Enfin, augmenter  $T_{obs}$  augmente également les temps de convergence, et cette période doit aussi être choisie grande devant le délai de propagation entre l'émetteur et le récepteur, afin de conserver la cohérence entre l'état du canal, les informations transmises à l'émetteur à travers le lien de retour, et la mise à jour de la stratégie d'allocation de puissance.

### 3.5 Conclusion

Un modèle de transmissions multi-utilisateurs et non-coopératives incluant de nombreuses contraintes réalistes vis-à-vis du contexte ASM a été posé. À partir de ce modèle, une méthode de partage de spectre entre liens de communications OFDM a été élaborée. Les outils de la théorie des jeux ont permis de proposer une solution non-coopérative et décentralisée, ne nécessitant qu'une connaissance statistique limitée de l'environnement de la part des émetteurs. Celle-ci s'accommode de toutes les contraintes posées par le contexte, en particulier : la difficulté d'obtenir de l'information sur l'état du canal ASM étant donnés les délais de propagation, l'hétérogénéité des sources ASM qui rend irréalistes les échanges d'informations entre les entités se partageant la ressource, ainsi que la nature aléatoire et doublement sélective du canal ASM. De plus, la connaissance imparfaite des jeux provenant des erreurs d'estimations du canal a été prise en compte. Un jeu robuste est formulé où les fonctions d'utilité sont maximisées dans le cas de la pire erreur d'estimation, sous contrainte d'incertitude bornée. De nombreux résultats de simulations sur des canaux ASM petits fonds synthétiques et sondés en mer illustrent les bénéfices pouvant être apportés par la méthode proposée<sup>1</sup>. Il est montré, en particulier, que la prise en compte de l'incertitude sur le canal de manière robuste permet de se rapprocher d'un optimum de Pareto donné par une allocation de puissance de type FDMA dans le cas de fortes interférences. L'avantage de ce qui est proposé ici réside dans le fait que l'algorithme d'apprentissage de la solution robuste reste complètement décentralisé et ne nécessite qu'une connaissance

---

1. Il est attendu que celle-ci s'applique sans modification et avec les mêmes bénéfices dans les canaux grands fonds.

locale de l'environnement de la part des joueurs, au contraire des algorithmes connus [Hayashi and Luo, 2009; Zhao and Pottie, 2013] calculant les allocations FDMA optimales dans le canal d'interférence.

Plusieurs extensions de ce travail pourraient être envisagées. En premier lieu, les solutions basées sur l'équilibre de Nash peuvent présenter le défaut d'être inéquitables, selon les conditions initiales des jeux étudiés. Il a été vu à la section 3.4.4 que, au regard des conditions initiales du jeu, le partage de spectre peut conduire à des bons résultats pour certains joueurs mais pas pour d'autres. En dépit des justifications théoriques de l'équilibre de Nash, cette iniquité peut présenter un caractère gênant dans des cas pratiques où, par exemple, des contraintes de qualité de service doivent être impérativement satisfaites. Enfin, bien que, pour des raisons de simplification, la présentation de ce travail ait été faite pour des systèmes utilisant la même modulation paramétrée de la même manière, cela n'est pas une hypothèse restrictive. En effet, on peut tout à fait étendre l'utilisation des outils de la théorie des jeux à un problème plus général de partage de ressource, faisant intervenir des joueurs dont les fonctions d'utilité et les espaces d'actions sont différents. Les difficultés proviennent alors, d'une part, du calcul (analytique ou numérique) des meilleures réponses de chaque joueur, mais aussi des conditions sous lesquelles un équilibre existe et si un algorithme convergent peut le calculer.



# Chapitre 4

## Partage décentralisé du canal sous contraintes de qualité de service

L'efficacité énergétique et la garantie d'un débit minimum d'informations transmises à un taux d'erreurs arbitrairement bas sont des contraintes d'importance capitale dans un grand nombre d'applications du monde marin utilisant les communications ASM. Les drones sous-marins autonomes, les réseaux de capteurs, et, d'une manière générale, les systèmes embarqués fonctionnant sur batterie sont des exemples de telles applications, où les efficacités spectrale et énergétique sont importantes. Dès lors, essayer d'établir des communications où l'interférence mutuelle entre sources acoustiques est potentiellement forte peut résulter à la fois en un gaspillage d'énergie et des communications non fiables. Dans le jeu de partage de spectre du chapitre précédent, la solution basée sur l'équilibre de Nash présente l'inconvénient d'être sous-optimale, du point de vue de l'ensemble des systèmes de communication impliqués, mais aussi d'être inéquitable en terme d'utilité atteignable selon les conditions initiales du jeu. De plus, sans connaissance a priori de la sévérité des canaux de transmission, il est difficile de savoir à l'avance si l'un ou l'autre des joueurs du jeu va atteindre un niveau minimal garanti d'utilité.

Ce chapitre propose donc de traiter le problème de partage du canal sous contrainte de qualité de service (QoS). Cette contrainte sera mesurée par des métriques liées au débit d'information théorique ou au rapport signal à interférence plus bruit (RSIB) moyen, et on mettra l'accent sur la minimisation de l'énergie consommée par les émetteurs lors de la mise en place de leurs stratégies d'utilisation des ressources. Le cadre théorique choisit reste celui des jeux non-coopératifs, dont les joueurs sont des systèmes de communications ASM. On s'intéresse aux solutions de partage de ressources décentralisées et adaptatives, en vertu des mêmes raisons que celles invoquées au chapitre précédent. De plus, et du fait de l'absence de coopération et de moyen de communication entre joueurs, les décisions prises par ces derniers sur leurs stratégies d'utilisation des ressources sont prises localement et à partir uniquement d'informations locales sur l'environnement, canal et environnement multi-utilisateur confondus.

Des problèmes similaires ont été traités par plusieurs travaux en technologie des communications en général, et en communications terrestres et radio-fréquences

en particulier. On peut citer les travaux de [Pang et al., 2008], qui formulent un jeu de partage de spectre sous contrainte de débit, et desquels nous allons nous inspirer dans la première section de ce chapitre pour développer le travail du chapitre précédent. Cependant, la contrainte de débit d'information étant exprimée en fonction des stratégies d'allocation de puissance des sources d'interférences, on verra apparaître une inter-dépendance des espaces de stratégies des joueurs, alors non convexes, de sorte que le théorème de Debreu 2.1.2 garantissant l'existence d'un équilibre de Nash ne s'applique plus.

Selon les gains des canaux des différents joueurs et les débits imposés, le problème de partage de spectre sous contrainte de débit peut ainsi ne pas posséder de solution. Les jeux de satisfaction offrent un autre moyen d'intégrer des contraintes fonctions des stratégies des joueurs dans un jeu. Proposé originellement par [Ross and Chaib-draa, 2006] comme moyen d'introduire une forme de coopération, le concept de l'équilibre de satisfaction a rencontré par la suite un certain succès en théorie des communications, notamment grâce aux travaux de [Goonewardena et al., 2017; Mériaux et al., 2012; Perlaza et al., 2012] où il est utilisé à des fins de partage de ressource sous contraintes de QoS dans les réseaux. Dans sa version originale, utilisée par [Mériaux et al., 2012; Perlaza et al., 2012], l'équilibre de satisfaction requiert que tous les joueurs du jeu soient satisfaits. En d'autres termes, cela revient à ce que toutes les contraintes de QoS soient vérifiées, et on se trouve alors face aux mêmes difficultés que dans le jeu de partage de spectre sous contraintes de [Pang et al., 2008]. L'équilibre de satisfaction généralisé de [Goonewardena et al., 2017], en revanche, n'est pas aussi restrictif et admet qu'à l'équilibre certains joueurs puissent ne pas pouvoir satisfaire leurs contraintes. Nous nous inspirerons de ce concept de solution dans la deuxième section de ce chapitre, où seront également proposés des algorithmes d'apprentissage permettant aux systèmes de communications ASM de jouer ces équilibres avec une connaissance de l'environnement réduite à un signal de 1 bit fourni par les récepteurs à leurs émetteurs respectifs.

## 4.1 Allocation de puissance sous contrainte d'utilité

### 4.1.1 Formulation du problème

On se place dans un cadre similaire à celui du chapitre 3, section 3.1.1. Un ensemble  $\mathcal{I} = \{1, \dots, I\}$  de  $I$  paires fixées d'émetteur et de récepteur (TX-RX), communiquant en modulation OFDM dans le canal ASM, se partagent un spectre d'une bande  $B$  divisée en  $N$  sous porteuses. L'interférence mutuelle est traitée comme du bruit colorée de distribution inconnue. On note  $X_{i,n}$  le symbole aléatoire transmis par l'émetteur  $i \in \mathcal{I}$  sur la sous porteuse  $n$ , les symboles de chaque sous porteuse sont i.i.d et indépendants des symboles  $X_{j,n'}$ ,  $\forall j \in \mathcal{I}$  et  $\forall n' \in \{1, \dots, N\}$ . Pour tout  $i$ , le symbole  $X_{i,n}$  est transmis à la puissance  $p_{i,n} = \mathbb{E}[|X_{i,n}|^2]$  allouée à la sous-porteuse  $n$ . Le symbole reçu au récepteur  $i$  est donné,  $\forall n$ , par

$$Y_{i,n} = H_{ii,n}X_{i,n} + \sum_{j \neq i} H_{ji,n}X_{j,n} + W_{i,n} \quad (4.1)$$

où  $W_{i,n} \sim \mathcal{CN}(0, \sigma_{i,n}^2)$ ,  $H_{ji,n}$  sont les coefficients du canal aléatoire à la sous porteuse  $n$  entre le récepteur  $i$  et l'émetteur  $j$  dont les distributions sont inconnues, et seules les statistiques des  $H_{ii,n}$  sont connues du récepteur  $i$ . Les hypothèses concernant la connaissance de l'environnement et des canaux qu'ont les émetteurs et les récepteurs restent inchangées par rapport au chapitre précédent (voir les points 1 à 5 de la section 3.1.1).

Le but de chaque paire TX-RX, considérée comme joueur du jeu de partage de spectre, est de minimiser la puissance allouée aux sous-porteuses sous contrainte de QoS. On note  $\mathbf{p}_i = [p_{i,1}, \dots, p_{i,N}]^T \in \mathbb{R}_+^N$  le vecteur d'allocation de puissance du joueur  $i$ ,  $\mathbf{p}_{-i} = [\mathbf{p}_1^T, \dots, \mathbf{p}_{i-1}^T, \mathbf{p}_{i+1}^T, \dots, \mathbf{p}_M^T]^T$  les allocations des autres joueurs, et  $\mathbf{p} = [\mathbf{p}_1^T, \dots, \mathbf{p}_M^T]^T$  le profil de stratégies. Le jeu de partage de spectre sous contrainte de QoS peut s'écrire sous la forme de l'ensemble de problèmes d'optimisation suivants

$$\mathbf{p}_i^* = \underset{\mathbf{p}_i}{\operatorname{argmin}} \sum_{n=1}^N p_{i,n} \quad \forall i \in \mathcal{I} \quad (4.2)$$

$$\text{sous la contrainte } \begin{cases} u_i(\mathbf{p}_i, \mathbf{p}_{-i}) \geq R_i^{\min} \\ p_{i,n} \geq 0, \quad \forall n \in \{1, \dots, N\} \end{cases}$$

où  $u_i(\mathbf{p}_i, \mathbf{p}_{-i}) \geq R_i^{\min}$  est la contrainte de débit d'information. Cette contrainte peut s'exprimer à travers la capacité ergodique

$$R_i(\mathbf{p}_i, \mathbf{p}_{-i}) = \alpha \sum_{n=1}^N \mathbb{E} \left[ \log \left( 1 + \frac{|H_{ii,n}|^2 p_{i,n}}{\sigma_{i,n}^2 + \sum_{j \neq i} |H_{ji,n}|^2 p_{j,n}} \right) \right] \quad (4.3)$$

$$\geq u_i(\mathbf{p}_i, \mathbf{p}_{-i}) \quad (4.4)$$

où  $u_i(\mathbf{p}_i, \mathbf{p}_{-i})$  est donnée par l'équation (3.9), l'inégalité ci-dessus étant prouvée selon les mêmes arguments qu'à la section 3.1.3. Ainsi, si la contrainte de QoS est vérifiée avec la fonction  $u_i(\mathbf{p}_i, \mathbf{p}_{-i})$  alors on garanti bien un débit d'information théorique supérieur  $R_i^{\min}$ . Comme au chapitre précédent, cette contrainte d'utilité peut être évaluée aux seuls moyens des statistiques du canal entre le récepteur  $i$  et son émetteur et à la densité spectrale de puissance (DSP) de l'interférence agrégée au bruit.

Le jeu peut alors s'exprimer par

$$\mathcal{G} \triangleq (\mathcal{I}, \mathcal{P}_1(\mathbf{p}_{-1}) \times \dots \times \mathcal{P}_{-I}(\mathbf{p}_{-I}), P^{\text{tot}}), \quad (4.5)$$

où  $P^{\text{tot}} = (P_i^{\text{tot}})_{i \in \mathcal{I}}$  sont les fonctions objectifs que les joueurs ont à minimiser, avec  $P_i^{\text{tot}} = \mathbf{1}_N^T \mathbf{p}_i$ , et les ensembles de stratégies sont donnés par

$$\mathcal{P}_i(\mathbf{p}_{-i}) = \left\{ \mathbf{p}_i \in \mathbb{R}_+^N \mid u_i(\mathbf{p}_i, \mathbf{p}_{-i}) \geq R_i^{\min} \right\}. \quad (4.6)$$

### 4.1.2 Équilibre de Nash et water-filling itératif

Comme au chapitre précédent, on propose de résoudre ce jeu grâce à l'équilibre de Nash. Ici, cet équilibre correspond à un profil de stratégies  $\mathbf{p}^* = [\mathbf{p}_1^{*T}, \dots, \mathbf{p}_I^{*T}]^T$

telles qu'aucun joueur  $i$  ne peut choisir une autre stratégie d'allocation  $\mathbf{p}_i \neq \mathbf{p}_i^*$  qui satisfasse sa contrainte de QoS sans que, pour ce faire, plus de puissance ne soit utilisée. Mathématiquement, on peut alors le définir ainsi :

**Définition 4.1.1.** *Un équilibre de Nash du jeu  $\mathcal{G}$  défini à l'équation (4.5) est un profil de stratégies  $\mathbf{p}^* = [\mathbf{p}_1^{*T}, \dots, \mathbf{p}_I^{*T}]^T$ , où  $\forall i \in \mathcal{I}$   $\mathbf{p}_i^* \in \mathcal{P}_i(\mathbf{p}_{-i}^*)$ , et tel que  $\forall i \in \mathcal{I}$  et  $\forall \mathbf{p}_i \in \mathcal{P}_i(\mathbf{p}_{-i}^*)$  on ait*

$$\sum_{n=1}^N p_{i,n}^* \leq \sum_{n=1}^N p_{i,n}. \quad (4.7)$$

La meilleure réponse d'un joueur  $i$  à un profil de stratégies adverse  $\mathbf{p}_{-i}$  fixé est l'unique solution du problème (4.2), qui peut être exprimée analytiquement par la résolution du système d'équation de Karush-Kuhn-Tucker. Cela amène à une solution de water-filling, s'exprimant  $\forall i$  et  $\forall n = 1, \dots, N$  par

$$p_{i,n}^* = \left[ \lambda_i - \frac{1}{\gamma_{i,n}(\mathbf{p}_{-i})} \right]^+, \quad (4.8)$$

où  $\lambda_i$  est le multiplicateur de Lagrange associé à la contrainte de QoS et choisit de manière à la saturer et  $\gamma_{i,n}(\mathbf{p}_{-i})$  est défini à l'équation (3.10).

En notant

$$WF_i[\mathbf{p}_{-i}] \triangleq \left( \left[ \lambda_i - \frac{1}{\gamma_{i,n}(\mathbf{p}_{-i})} \right]^+ \right)_{n=1}^N \quad (4.9)$$

et

$$WF[\mathbf{p}] = (WF_i[\mathbf{p}_{-i}])_{i \in \mathcal{I}}, \quad (4.10)$$

l'équilibre de Nash du jeu (4.5) se définit alors comme le point fixe des meilleures réponses des joueurs les uns aux autres, c'est à dire que  $\mathbf{p}^* = WF[\mathbf{p}^*]$ . De la même façon qu'au chapitre 3, cet équilibre peut être appris itérativement par la récursion  $\mathbf{p}_i(t+1) = WF_i[\mathbf{p}_{-i}(t)]$  mise en œuvre de manière distribuée et où, pour tout joueur  $i$ , la stratégie adverse  $\mathbf{p}_{-i}$  est indirectement perçue à travers la DSP de l'interférence et du bruit.

Comme il l'a été mentionné précédemment, il se peut qu'aucune solution n'existe au problème (4.2). De plus, rien ne garanti que cette solution soit unique ni qu'un algorithme distribué converge vers un équilibre de Nash. Des conditions suffisantes pour ces trois propriétés (existence, unicité et convergence) ont été néanmoins données par [Pang et al., 2008]. Ces conditions, de manière analogue à celles de la section 3.2.3 dans le jeu de partage de spectre sous contrainte de puissance, possèdent une interprétation physique faisant intervenir l'interférence : il se trouve qu'à mesure que l'interférence augmente, avoir un ensemble non-vide (et à puissance bornée) d'équilibres de Nash nécessite de réduire les contraintes de débit des joueurs. Il en va de même pour les conditions de convergence de l'algorithme de water-filling itératif et l'unicité de l'équilibre, qui requièrent d'avoir assez peu d'interférences. Les problèmes de recherche d'équilibres de Nash tels que celui présenté ici, où les espaces de stratégies des joueurs sont interdépendants, forment la classe des *problèmes d'équilibres de Nash généralisés* [Facchinei and Kanzow, 2007]. En général, ces problèmes sont difficiles à résoudre et,

à moins de poser certaines hypothèses telles que la convexité locale de l'ensemble des profils de stratégies, l'existence d'une solution d'équilibre de Nash n'est pas garantie. Cette difficulté sera contournée grâce aux équilibres de satisfaction à la section 4.2

On donne, dans le paragraphe suivant, quelques résultats du jeu de partage de spectre sous contrainte de QoS obtenus sur des canaux synthétiques.

### 4.1.3 Résultats sur des canaux synthétiques

Le jeu de partage de spectre sous contrainte de QoS est illustré au travers d'un scénario particulier. On se base sur des canaux synthétiques issus du simulateur déjà utilisé au chapitre 3 et les paramètres des modulations OFDM et du simulateur sont les mêmes qu'auparavant. Un jeu à trois joueurs est simulé. Leur contrainte de QoS est fixée à  $R_i^{min} = 1$  bits/s/Hz (on rappelle que  $u_i$  est homogène à une efficacité spectrale). Les distances entre les récepteurs et leurs émetteurs correspondant sont fixées à  $d_{ii} = 1$  km, tandis que les émetteurs interférents sont à  $d_{ji} = 1.5$  km. Les profondeurs des terminaux ont été tirées aléatoirement dans une colonne d'eau de 20 m. Le contexte est celui d'une interférence relativement forte. On notera que dans ce cas, l'existence et l'unicité d'un équilibre, de même que la convergence de l'algorithme de water-filling itératif, ne sont pas garanties par les conditions suffisantes données par [Pang et al., 2008]. Les canaux utilisés sont illustrés à la figure 4.1.

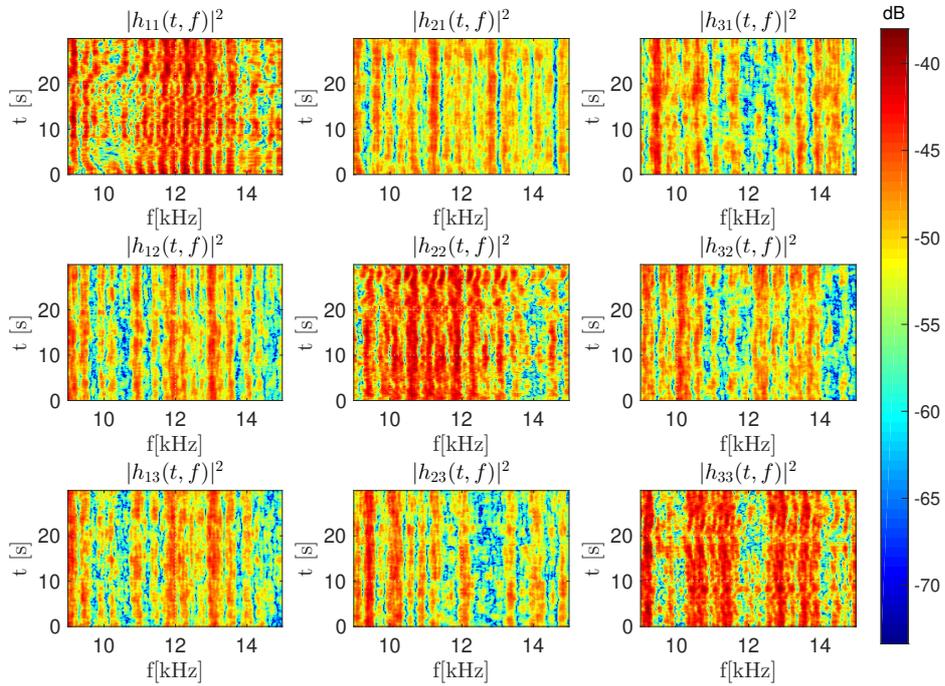


FIGURE 4.1 – Réponses fréquentielles variant dans le temps du simulateur de [Socheleau et al., 2010b]. Les profondeurs, tirées aléatoirement, sont pour les émetteurs de 1 à 3 de 5, 20 et 11 m respectivement. Pour les récepteurs 1 à 3, celles-ci sont de 9, 16 et 13 m respectivement.  $d_{ii} = 1$  km,  $d_{ji} = 1.5$  km,  $j \neq i$ .

La puissance des émetteurs est fixée à  $P_i^{INIT} = 160$  dB ref.  $\mu$ Pa uniformément sur la bande de transmission à l'état initial du jeu, et l'algorithme de water-filling

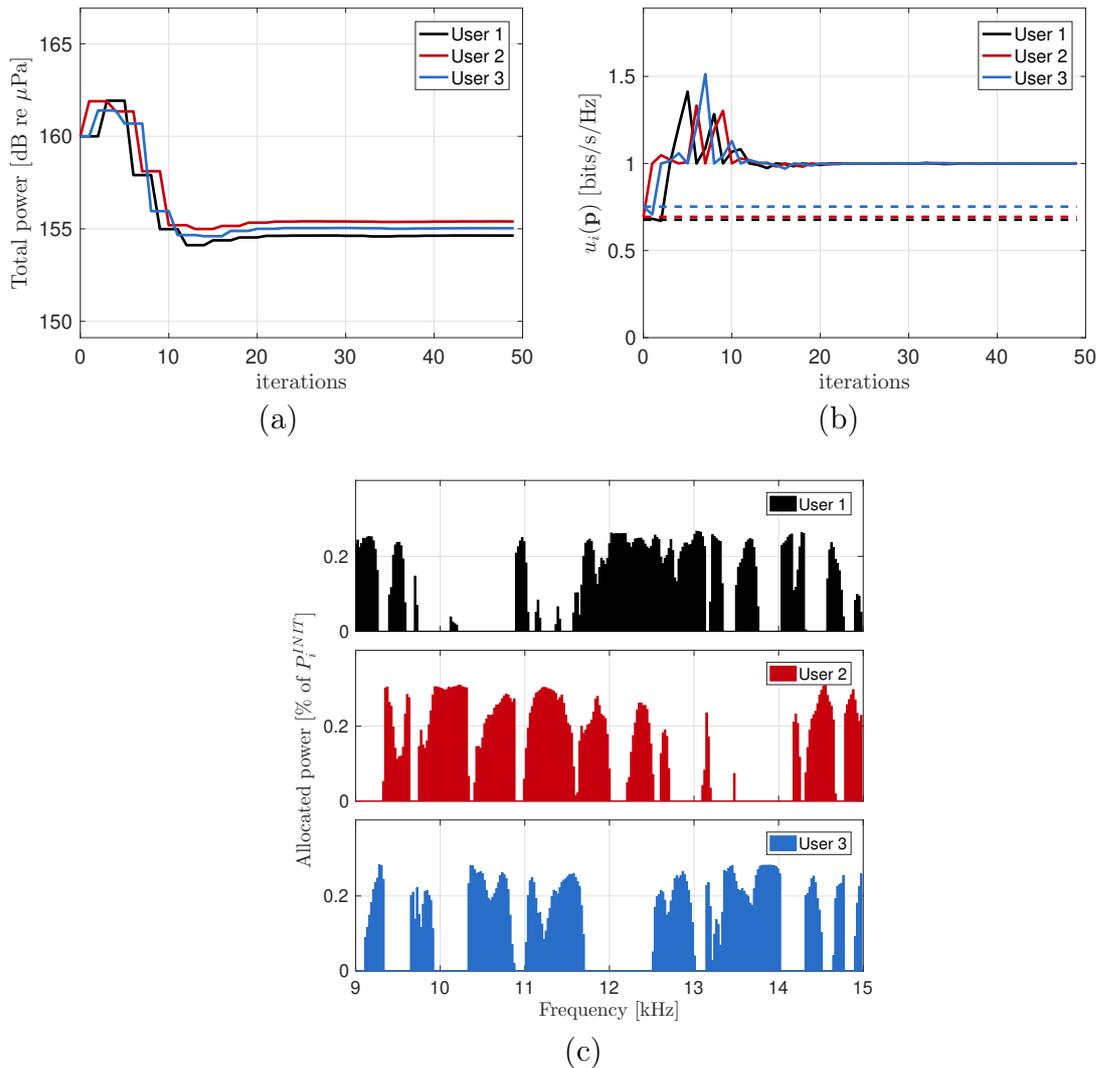


FIGURE 4.2 – (a) Puissance moyenne de transmission en fonctions des itérations. (b) Contrainte de QoS des joueurs. (c) Répartition des joueurs sur le spectre.  $d_{ii} = 1$  km,  $d_{ji} = 1.5$  km pour  $j \neq i$ .

2, dans sa version séquentielle et sur la base de l'opérateur de l'équation (4.9), est utilisé pour le calcul des équilibres de Nash. La période de mise à jour est fixée à  $T_{obs} = 30$  secondes. Les résultats sont évalués en termes de puissance émise et de la contrainte de QoS des joueurs au fil des itérations, illustrées aux figures 4.2-(a) et 4.2-(b) respectivement. La figure 4.2-(c) illustre les allocations de puissance à l'équilibre. Après 15 itérations, soit environ 2.5 minutes, les joueurs sont proches de l'état d'équilibre. Les bénéfices du partage de spectre comparés à l'allocation uniforme à puissance  $P_i^{INIT}$ , en ligne pointillée, sont clairement visibles à la figure 4.2-(b). La contrainte de QoS peut en effet être atteinte avec 5 dB de puissance transmis en moins. Enfin, on pourra analyser la figure 4.2-(c) selon les mêmes procédés qu'au chapitre 3 pour faire ressortir la cohérence des allocations de puissance des trois joueurs avec les canaux de la figure 4.1.

Dans un deuxième scénario, on diminue la distance entre les récepteurs et les émetteurs interférents. Les distances  $d_{ji}$  pour  $j \neq i$  sont fixées à 500 m, si bien qu'à l'allocation uniforme de l'état initial, chaque récepteur perçoit approximativement deux

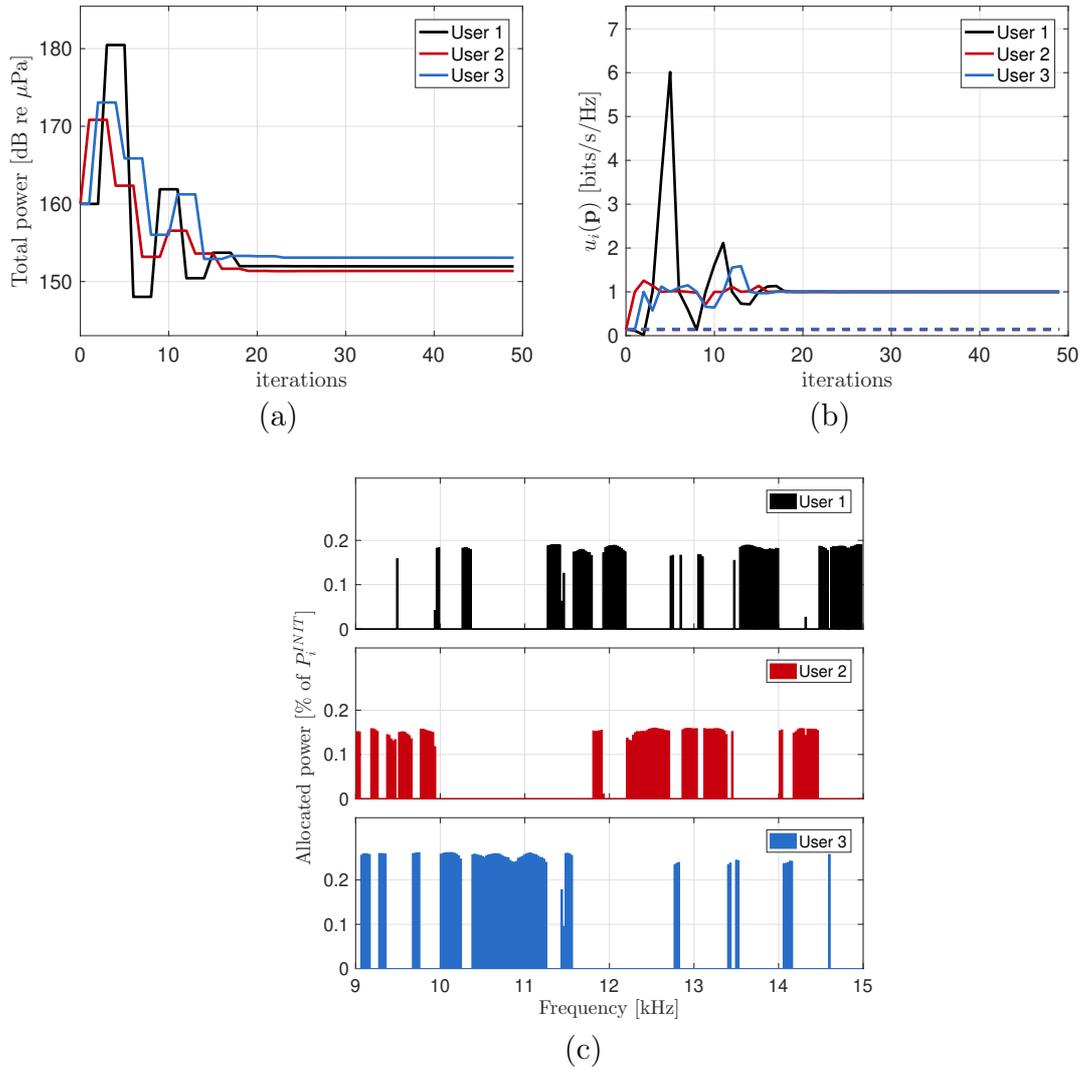


FIGURE 4.3 – (a) Puissance moyenne de transmission en fonctions des itérations. (b) Contrainte de QoS des joueurs. (c) Répartition des joueurs sur le spectre.  $d_{ii} = 1$  km,  $d_{ji} = 0.5$  km pour  $j \neq i$ .

fois plus d'interférence que de signal utile<sup>1</sup>. Les figures 4.3 illustrent les résultats. Les joueurs sont en équilibre après 25 itérations, au cours desquelles la QoS a considérablement été augmentée en comparaison de l'allocation initiale, et ce pour environ 7 à 9 dB de puissance transmise en moins. Ceci s'explique par le niveau élevé d'interférence qui amène les joueurs à opter pour un partage de spectre orthogonal en fréquence (FDMA). On peut voir, à partir de la figure 4.3-(a), que les émetteurs doivent augmenter considérablement la puissance transmise durant la phase de convergence. Chaque joueur, l'un après l'autre, augmente sa puissance transmise pour pouvoir satisfaire sa contrainte. Ils agissent de cette manière itération après itération jusqu'à ce que chaque joueur ait réussi à repousser les autres des sous-porteuses qu'il compte utiliser, atteignant ainsi l'allocation FDMA. À partir de cette étape, les itérations suivantes de l'algorithme conduisent alors les émetteurs à réduire la puissance de transmission étant donnée l'absence (dans ce scénario) d'interférence mutuelle. Cependant, l'augmentation de puissance transmise nécessaire pendant la phase de convergence peut devenir rapidement prohibitive pour la vaste majorité des systèmes de communications acoustiques sous-marines.

## 4.2 Approche par les jeux de satisfaction

La modélisation du problème de partage de canal sous contrainte de QoS par un jeu sous forme stratégique en vue de le résoudre par l'équilibre de Nash présente le problème majeur de l'existence non garantie d'une solution satisfaisant toutes les contraintes. De plus, et comme au chapitre 3, le calcul des meilleures réponses nécessite que les joueurs disposent de la connaissance des statistiques de leur canal. Celles-ci sont, en pratique, fournies à l'émetteur par un lien de retour depuis son récepteur, à travers lequel transitent plusieurs dizaines ou centaines de bits, selon l'encodage des statistiques du canal. Par exemple, dans les jeux précédents, si on suppose que les statistiques du canal sur chaque sous-porteuses sont encodées sur 16 bits, le lien de retour doit être capable de transmettre 4096 bits à l'émetteur. Habituellement, on préfère mettre en place des liens de retour bas débit, pour assurer plus de robustesse, ce qui pose alors la question du délai nécessaire pour transmettre toutes ces données.

### 4.2.1 Formulation du problème

Le problème de partage du canal ASM sous contrainte de QoS est présenté à travers le prisme des jeux de satisfaction, qui offrent des outils pour contourner les difficultés évoquées ci-dessus. Des jeux faisant intervenir des communications OFDM et à étalement de spectre par séquence directe (DSSS) sont étudiés. En plus de la flexibilité qu'elle offre pour l'allocation de puissance sur des sous-bandes fréquentielles, la modulation OFDM est utilisée habituellement pour les communications haut-débit [Zhou and Wang, 2014]. L'étalement de spectre DSSS offre quant à lui des gains d'étalement profitables à la robustesse des communications [Stojanovic et al., 1998]. Deux

---

1. En situation réelle, l'estimation du canal nécessaire à l'opérateur de water-filling peut être compromise par l'interférence.

types de lien de retour sont considérés pour l'étude de ces jeux : dans un premier temps, on suppose que le récepteur fournit l'état statistique du canal à l'émetteur, comme au chapitre précédent (on définira plus loin la forme prise par cette quantité dans le jeu DSSS). Puis, dans un second temps, ce retour est réduit à un signal binaire indiquant à l'émetteur l'état de satisfaction de la contrainte de QoS pour la stratégie utilisée. Dans la suite, on notera  $\mathcal{I} = \{1, \dots, I\}$  l'ensemble des joueurs constitué par les liens de communication, OFDM ou DSSS selon le jeu traité. Ces joueurs communiquent en même temps sur la même bande de fréquence  $B$ , et sont dotés de leurs ensembles privés de stratégies  $\mathcal{A}_i$  et de leurs fonctions d'utilité  $u_i$ ,  $i \in \mathcal{I}$ , modélisant respectivement leurs paramètres de transmission et leur performance. Selon le type de modulation utilisé, les stratégies dans  $\mathcal{A}_i$  correspondent à des puissances d'émissions et/ou des gains d'étalement, ces derniers correspondant à la taille de la séquence d'étalement. Chaque joueur est contraint de satisfaire une certaine qualité de service, modélisée par un seuil fixé sur son utilité *i.e.*,  $u_i(a_i, a_{-i}) \geq \Gamma_i$  où  $a_{-i} \in \mathcal{A}_{-i}$  sont les stratégies des adversaires de  $i$  et  $a_i \in \mathcal{A}_i$  est la stratégie du joueur  $i$ .

On peut formuler le jeu de satisfaction correspondant au scénario ci-dessus par le tuple

$$\mathcal{G} = (\mathcal{I}, \mathcal{A}, (u_i)_{i \in \mathcal{I}}, (\varphi_i)_{i \in \mathcal{I}}). \quad (4.11)$$

Des correspondances de satisfaction de la forme

$$\varphi_i(a_{-i}) = \{a_i \in \mathcal{A}_i : u_i(a_i, a_{-i}) \geq \Gamma_i\} \quad \forall a_{-i} \in \mathcal{A}_{-i} \quad (4.12)$$

sont considérées afin d'expliciter les contraintes de QoS. Ces correspondances donnent, pour chaque joueur, les stratégies de transmission qui satisfont la QoS requise face au profil de stratégies des autres joueurs. Le but des joueurs est d'apprendre, de manière décentralisée, un équilibre de satisfaction tel que défini au chapitre 2 par la définition 2.2.2. On conservera autant que possible les notations du paragraphe 2.2.1 du chapitre 2. Selon le type de modulation utilisée, les joueurs auront des espaces d'actions et des fonctions d'utilité différentes. Dans le cadre des modulations OFDM, la satisfaction sera évaluée en termes de débit théorique d'information pour une certaine allocation de puissances sur les sous porteuses, tandis que la satisfaction d'un système DSSS sera mesurée par son RSIB en réception, en fonction de la puissance émise et du gain d'étalement.

### 4.2.2 Algorithme d'apprentissage aveugle

Les équilibres de satisfaction du jeu (4.11) peuvent être appris au moyen de l'algorithme 1 exposé au paragraphe 2.2.2 du chapitre 2. Néanmoins, cela suppose que les émetteurs soient capables de déterminer toutes les actions satisfaisantes renvoyées par la correspondance  $\varphi_i(a_{-i})$  pour le profil de stratégies  $a_{-i}$  de leurs adversaires. Pour ce faire, une certaine connaissance de l'état du canal et des interférences est requise, et qui, comme mentionné au début de cette section, nécessite de transmettre plusieurs dizaines ou centaines de bits à travers le lien de retour depuis le récepteur. Dans un contexte de communications ASM, il est crucial de réduire l'activité du lien de retour afin de ne pas surcharger les bandes passantes disponibles déjà limitées. Une approche pragmatique consiste à ne transmettre qu'un signal binaire sur le lien de

retour, indiquant à l'émetteur s'il doit adopter une nouvelle stratégie de transmission ou non.

Un algorithme d'apprentissage aveugle des équilibres de satisfaction est donc proposé afin de permettre aux liens de communication ASM de se passer de l'observation de  $a_{-i}$  et de la connaissance des stratégies satisfaisantes de l'ensemble  $\varphi_i(a_{-i})$ . Posons  $a(t) \in \mathcal{A}$  le profil de stratégies joué au temps  $t \in \mathcal{T}$ , décomposable sous la forme  $a(t) = (a_i(t), a_{-i}(t))$  pour un  $i \in \mathcal{I}$  donné. Les joueurs du jeu de satisfaction ne sont donc pas capables d'évaluer à quel ensemble parmi  $\mathcal{I}_S(a(t))$ ,  $\mathcal{I}_N(a(t))$  ou  $\mathcal{I}_U(a(t))$ <sup>2</sup> ils appartiennent. Néanmoins, on remarque qu'un joueur insatisfait  $i \in \mathcal{I}$  n'a de déviation profitable que si  $\varphi_i(a_{-i}(t))$  est non-vide. En d'autres termes, et comme pour l'algorithme 1, seuls les joueurs appartenant à  $\mathcal{I}_N(a(t))$  joueront une nouvelle stratégie dans les itérations du processus d'apprentissage ultérieures à  $t$ . En vertu des hypothèses du paragraphe précédent sur les fonctions d'utilité et les espaces de stratégies, il suffit, pour un  $a_{-i}(t)$  donné, de tester si la condition  $u_i(\bar{a}_i, a_{-i}(t)) \geq \Gamma_i$  (avec  $\bar{a}_i$  le plus grand élément de  $\mathcal{A}_i$ ) est vérifiée pour savoir si  $\varphi_i(a_{-i}(t))$  est vide ou non. Ainsi, lorsque le joueur  $i$  n'est pas satisfait, son récepteur peut évaluer s'il existe une stratégie satisfaisante pour les prochaines itérations en testant si la stratégie la plus coûteuse est elle-même satisfaisante, et transmettre le signal de retour codé sur 1 bit dont l'état correspond au cas échéant. Le signal de retour du joueur  $i \in \mathcal{I}$  sera noté  $S_i$  dans la suite et prend ses valeurs dans  $\{0, 1\}$ , où  $S_i = 0$  correspond au cas où l'émetteur doit changer de stratégie - c'est à dire  $i \in \mathcal{I}_N(a(t))$  - tandis que  $S_i = 1$  signifie qu'il n'y a pas de déviation profitable -  $i \in \mathcal{I}_U(a(t))$  ou  $i \in \mathcal{I}_S(a(t))$ . Soit  $\mathcal{Y}_i(t)$  l'ensemble des stratégies strictement plus coûteuses que la stratégie  $a_i(t)$  jouée au temps  $t \in \text{set } \mathcal{T}$  par le joueur  $i$ ,

$$\mathcal{Y}_i(t) = \{y_i \in \mathcal{A}_i : a_i(t) \prec y_i \preceq \bar{a}_i\}. \quad (4.13)$$

Puisque l'émetteur ne connaît pas ce que contient l'ensemble  $\varphi_i(a_{-i}(t))$ , s'il reçoit le signal  $S_i = 0$ , il choisit aléatoirement une stratégie dans  $\mathcal{Y}_i(t)$ . On notera

$$\Phi_t^{(i)} : \Omega \rightarrow \mathcal{Y}_i(t) \quad (4.14)$$

la variable aléatoire modélisant ce choix. Dans le cas contraire, l'émetteur reçoit  $S_i = 1$  de son récepteur et conserve donc la même stratégie. L'algorithme 3 met cette procédure en forme, où on note  $\mathcal{T} = \{0, 1, 2, \dots\}$  des indices dans un ensemble de temps représentés par une séquence croissante de nombres réels, espacés d'un pas quelconque, et  $\mathcal{T}_i \subseteq \mathcal{T}$  l'ensemble des temps où le joueur  $i$  peut jouer une nouvelle stratégie

La proposition 4.2.1 établit une condition suffisante de convergence pour les algorithmes 1 et 3, pour autant que toutes les hypothèses du paragraphe 2.2.1 soient respectées. Cette proposition et sa preuve sont similaires à celles données dans [Goonewardena et al., 2017, Section III.A et III.B], mais sont complétées ici pour l'algorithme aveugle. Quelques définitions préliminaires doivent être introduites. Soit la relation  $<_{\mathcal{A}}$  définie sur l'ensemble  $\mathcal{A}$  des profils de stratégies par

$$\forall a, a' \in \mathcal{A}^2, \quad a <_{\mathcal{A}} a' \Leftrightarrow \forall i \quad a_i \preceq a'_i, \exists j \quad a_j \prec a'_j. \quad (4.16)$$

---

2. On rappelle que ces ensembles partitionnent l'ensemble des joueurs pour la stratégie  $a(t)$  et qu'ils correspondent respectivement aux joueurs satisfaits de leur stratégie actuelle, aux joueurs insatisfaits mais pour lesquels il existe une déviation unilatérale profitable, et aux joueurs insatisfaits n'ayant aucune stratégie satisfaisante si les autres joueurs ne changent pas de stratégie. Le lecteur est invité à se reporter aux équations (2.29), (2.30) et (2.31) du paragraphe 2.2.1 du chapitre 2 pour les définitions de ces ensembles.

---

**Algorithme 3** Dynamique de réponses de satisfaction aveugles
 

---

- 1:  $t = 0$
- 2:  $a(t) = (\underline{a}_1, \dots, \underline{a}_I)$
- 3: **tant que**  $\mathcal{I}_N(a(t)) \neq \emptyset$  **faire**  $\forall i \in \mathcal{I}$
- 4:     **si**  $t \in \mathcal{T}_i$  **alors**
- 5:

$$a_i(t+1) = \begin{cases} \Phi_t^{(i)}(\omega) & \text{si } S_i = 0 \\ a_i(t) & \text{si } S_i = 1 \end{cases} \quad (4.15)$$

- 6:     **fin si**
  - 7:      $t = t + 1$
  - 8: **fin tant que**
- 

La fonction  $\phi_i : \mathcal{A}_{-i} \rightarrow \mathcal{A}_i$  préserve l'ordre si, étant donnés  $a_{-i}, a'_{-i} \in \mathcal{A}_{-i}^2$  tels que  $a_{-i} <_{\mathcal{A}_{-i}} a'_{-i}$ , la relation suivante se vérifie :

$$\phi_i(a_{-i}) \preceq \phi_i(a'_{-i}). \quad (4.17)$$

**Proposition 4.2.1.** *Si dans le jeu (4.11) et  $\forall i \in \mathcal{I}$  les conditions suivantes sont vérifiées :*

1.  $\forall t \geq 0, \exists t' > t$  such that  $t' \in \mathcal{T}_i$ ,
2.  $\forall a_{-i} \in \mathcal{A}_{-i}, \varphi_i(a_{-i}) = \{a_i \in \mathcal{A}_i : \phi_i(a_{-i}) \leq a_i\}$ ,  
où  $\phi_i : \mathcal{A}_{-i} \rightarrow \mathcal{A}_i$  préserve l'ordre,
3.  $\forall a_i \in \mathcal{A}_i, \exists \bar{a}_i \in \mathcal{A}_i$  such that  $a_i \preceq \bar{a}_i$ ,

alors les algorithmes 1 et 3 convergent vers un équilibre de satisfaction en stratégies pures.

*Démonstration.* Tout d'abord il doit être noté que la condition 1) signifie simplement que tout joueur aura, à un moment ou à un autre, une opportunité de jouer une stratégie. Ceci n'est pas une hypothèse restrictive, mais est nécessaire dans la suite de la preuve.

Soit  $t = 0$ , alors  $a(0) = (\underline{a}_i)_{i \in \mathcal{I}}$ . Si  $\mathcal{I}_N(a(0)) = \emptyset$  alors  $a(0)$  est un équilibre et la procédure se termine. Dans le cas contraire, une déviation profitable de la part d'au moins un des joueurs existe. Soit  $t \geq 0$  et  $a(t) \in \mathcal{A}$ . Si  $\mathcal{I}_N(a(t)) \neq \emptyset$  alors  $\exists i \in \mathcal{I}$  tel que  $\varphi_i(a_{-i}(t)) \neq \emptyset$  et  $a_i(t) \notin \varphi_i(a_{-i}(t))$ , ce qui revient à dire que  $a_i(t) < \phi_i(a_{-i}(t))$ . Soit  $\mathcal{Y}_i(t)$  l'ensemble de stratégies défini par l'équation (4.13). En vertu de l'hypothèse 2) de la proposition, on a  $\varphi_i(a_{-i}(t)) \in \mathcal{P}(\mathcal{Y}_i(t))$ .

Lorsque  $\mathcal{I}_N(a(t)) \neq \emptyset$ , au moins un joueur  $i \in \mathcal{I}_N(a(t))$  devra jouer une réponse de satisfaction telle que  $\varphi_i(a_{-i}(t)) \in \mathcal{P}(\mathcal{Y}_i(t))$ . De l'hypothèse 1), on déduit que  $\forall t \geq 0$ ,

$$\exists t' > t \text{ tel que } a(t) <_{\mathcal{A}} a(t') \leq_{\mathcal{A}} \bar{a}. \quad (4.18)$$

Ainsi,  $a(t)$  augmente relativement à la relation d'ordre  $<_{\mathcal{A}}$ . De part l'hypothèse 3),  $a(t)$  ne peut excéder  $\bar{a}$ , qui est un équilibre de Nash du jeu sous forme normale  $\mathcal{G}_{NE}$  et aussi un équilibre de satisfaction (voir le paragraphe 2.2.1). Par conséquent, les algorithmes

1 et 3 convergent bien vers un équilibre qui est soit  $\bar{a}$  soit n'importe quel autre  $a^* \leq_{\mathcal{A}} \bar{a}$  such that  $\mathcal{I}_N(a(t)) = \emptyset$ .  $\square$

Les algorithmes 1 et 3 ont un nombre maximum d'itérations en  $\mathcal{O}(I \times \max_{i \in \mathcal{I}} |\mathcal{A}_i|)$  [Goonewardena et al., 2017] avant convergence. Le cas où le nombre d'itérations est exactement  $I \times \max_{i \in \mathcal{I}} |\mathcal{A}_i|$  apparaît lorsque que tous les joueurs sont initialement dans  $\mathcal{I}_N$ , jouent systématiquement la stratégie la moins coûteuse et se trouvent une nouvelle fois dans  $\mathcal{I}_N$  à l'itération suivante où ils doivent jouer. Concernant la complexité d'intégration dans un système réel, on s'attend à ce qu'elle soit négligeable au regard des quelques instructions simples nécessaires à la mise en œuvre algorithmes.

### 4.2.3 Modèle des transmissions

Des transmissions par blocs sont considérées en ce qu'elles généralisent les schémas de modulation OFDM et DSSS. Soit  $\mathbf{x}_i \in \mathbb{C}^{N_i}$  un bloc de  $N_i$  symboles i.i.d. et de moyenne nulle transmis par l'émetteur  $i$  à son récepteur. Selon le type de systèmes, la taille du bloc de  $N_i$  symboles transmis correspond à la taille de la séquence d'étalement ou au nombre de sous-porteuses. Pour tout  $i$  et quel que soit  $j \neq i$ , les symboles  $\mathbf{x}_j \in \mathbb{C}^{N_j}$  sont indépendants de  $\mathbf{x}_i$ . Chaque émetteur est sujet à une contrainte de puissance moyenne de transmission, de telle sorte que

$$P_i^{min} \leq P_i \leq P_i^{max} \quad (4.19)$$

où  $P_i = \|\mathbf{C}_i(a_i)\|_F = \text{Tr}[\mathbb{E}[\mathbf{x}_i \mathbf{x}_i^H]]$ . Du point de vue d'un joueur  $i$  particulier, la matrice de covariance des symboles, notée  $\mathbf{C}_i(a_i)$ , est une fonction de la stratégie  $a_i \in \mathcal{A}_i$  choisie par ce joueur au sens où la sélection d'une stratégie particulière influence certaines propriétés de cette matrice telles que sa taille, à travers la sélection d'une séquence d'étalement, ou sa norme, à travers le choix de la puissance d'émission.

En toute généralité, un bloc de symboles reçus par le récepteur  $i$  peut s'exprimer par

$$\mathbf{y}_i = \mathbf{H}_{ii} \mathbf{x}_i + \sum_{j \neq i} \mathbf{H}_{ji} \mathbf{x}_j + \mathbf{w}_i \quad (4.20)$$

$$= \mathbf{H}_{ii} \mathbf{x}_i + \mathbf{z}_i(\mathbf{x}_{-i}), \quad (4.21)$$

où  $\mathbf{w}_i \sim \mathcal{CN}(0, \sigma_i^2 \mathbf{I}_{N_i})$  est un bruit gaussien circulaire, symétrique et indépendant des  $\mathbf{x}_j, \forall j \neq i$ . Les matrices  $\mathbf{H}_{ji} \in \mathcal{M}_{N_j \times N_i}(\mathbb{C})$  représentent les canaux de communication entre le récepteur  $i$  et l'émetteur  $j$ , et dont les coefficients sont aléatoires et varient dans le temps. Le récepteur  $i$  n'a qu'une connaissance statistique de son canal direct  $\mathbf{H}_{ii}$ , considéré stationnaire au sens large sur une durée de  $L$  symboles, et de son terme d'interférence plus bruit  $\mathbf{z}_i(\mathbf{x}_{-i})$ . Rien n'est connu concernant  $\mathbf{H}_{ji}$  ou  $\mathbf{x}_j$  dès lors que  $j \neq i$ . On considère que  $L$  est assez grand de sorte que les moyennes empiriques convergent vers les espérances.

Les fonctions d'utilité  $(u_i)_{i \in \mathcal{I}}$ , déterminant le processus d'adaptation des stratégies de chaque joueur, doivent être choisis de sorte qu'elle ne fassent pas intervenir plus

de connaissance de la part des joueurs que la connaissance statistique disponible sur leurs canaux et l'interférence perçue. Afin de permettre l'adaptation des paramètres de transmission de l'émetteur  $i$ , le récepteur  $i$  renvoie sa connaissance de l'environnement par lien de retour, de la même manière qu'au chapitre précédent. Cette connaissance peut s'exprimer à travers un état statistique du canal, tel que celui donné par exemple à l'équation (3.10), ou par un signal binaire indiquant si l'émetteur doit changer de stratégies ou non. Dans ce dernier cas, l'état de ce signal est choisi par le récepteur en fonction des statistiques du canal et l'interférence. L'adaptation des stratégies est organisée de la même façon qu'au chapitre précédent : les émetteurs sont supposés s'adapter périodiquement, sans être synchronisés, et sur la base des informations fournies par leur récepteur respectif. En pratique, la période mise à jour de la stratégie de chaque émetteur correspond aussi au temps dont dispose le récepteur pour estimer les statistiques du canal et de l'interférence, avant de transmettre un signal sur le lien de retour. Selon le cas de figure et la perspective d'un joueur  $i \in \mathcal{I}$  particulier, la procédure est la suivante :

- Algorithme 1 : Au temps  $t_i \in \mathcal{T}_i$ , le récepteur  $i$  renvoie les statistiques du canal  $\mathbf{H}_{ii}$  et du terme d'interférence  $\mathbf{z}_i(\mathbf{x}_{-i})$  à l'émetteur à travers le lien de retour. L'émetteur est alors capable d'évaluer sa fonction d'utilité  $u_i(a_i, a_{-i})$  pour n'importe laquelle de ses stratégies  $a_i$  étant donnée l'interférence, dépendante de  $a_{-i}$  (à travers la covariance des symboles transmis). Il en déduit, si elles existent, les stratégies satisfaisantes de  $\varphi_i(a_{-i})$ , pour lesquelles  $u_i(a_i, a_{-i}) \geq \Gamma_i$ , et choisit la moins coûteuse pour la période suivante - dont la fin est marquée par le plus petit  $t'_i \in \mathcal{T}_i$  tel que  $t'_i > t_i$ .
- Algorithme 3 : Au temps  $t_i \in \mathcal{T}_i$ , le récepteur vérifie, sur la base de sa propre connaissance des statistiques de  $\mathbf{H}_{ii}$  et  $\mathbf{z}_i(\mathbf{x}_{-i})$ , si la contrainte  $u_i(a_i, a_{-i}) \geq \Gamma_i$  est satisfaite et si  $\varphi_i(a_{-i})$  est non-vide (tel que dans les descriptions du paragraphe 4.2.2). Il peut ainsi déterminer à quel ensemble, parmi  $\mathcal{I}_U(a)$ ,  $\mathcal{I}_S(a)$  ou  $\mathcal{I}_N(a)$ , le joueur qu'il forme avec son émetteur appartient. Si  $i$  appartient à  $\mathcal{I}_N(a)$ , le récepteur envoie  $S_i = 0$  à l'émetteur pour lui indiquer l'existence d'une déviation stratégique profitable. N'ayant pas la connaissance de  $\varphi_i(a_{-i})$ , ce dernier choisit aléatoirement une stratégie parmi celles qui sont plus coûteuses que  $a_i$  dans son espace de stratégies. Sinon, le joueur appartient soit à  $\mathcal{I}_U(a)$  soit à  $\mathcal{I}_S(a)$ , auquel cas le récepteur envoie le signal  $S_i = 1$  pour indiquer que la stratégie actuelle doit être conservée.

L'impact du degré de connaissance disponible à l'émetteur sera évalué par simulation au paragraphe 4.2.6 Le jeu de satisfaction est détaillé pour l'OFDM et le DSSS dans les deux paragraphes suivants.

#### 4.2.4 Étude de cas : Communications DSSS

En modulation DSSS, les vecteurs  $\mathbf{x}_i$  de l'équation (4.20) peuvent être écrits comme  $\mathbf{x}_i = S_i \times \mathbf{u}_i$  où  $\mathbf{u}_i \in \mathbb{C}^{N_i}$  est une séquence pseudo-aléatoire d'étalement de  $N_i$  symboles telle que  $\forall n \in \mathbb{Z}$ ,

$$R_{uu}[n] = \sum_{k=1}^N u_{i,k} u_{i,n-k}^* \approx N_i \delta_n, \quad (4.22)$$



Une fonction d'utilité basée sur le RSIB est choisie. Après dés-étalement au récepteur, on peut exprimer le RSIB moyen du joueur  $i$  par<sup>3</sup>

$$u_i(a_i, \mathbf{x}_{-i}) = \frac{\mathbb{E} [\|\mathbf{H}_{ii}\|_F^2] P_i^{(a_i)}}{\mathbb{E} [\|\mathbf{z}_i(\mathbf{x}_{-i})\|_2^2]} N_i^{(a_i)} \quad (4.27)$$

et sa correspondance de satisfaction par

$$\varphi_i(\mathbf{x}_{-i}) = \{a_i \in \{1, \dots, n_i\} : u_i(a_i, \mathbf{x}_{-i}) \geq \Gamma_i\}, \quad (4.28)$$

où  $\Gamma_i$  est le RSIB minimum satisfaisant la contrainte de QoS.

Bien que cela ne soit pas noté explicitement, les stratégies  $a_{-i}$  gouvernent évidemment la puissance du terme d'interférence  $\|\mathbf{z}_i(\mathbf{x}_{-i})\|_2^2$  perçu par le récepteur  $i$ . Il est également clair, à l'examen de l'équation (4.27), qu'il n'est pas nécessaire pour le joueur  $i$  d'observer et de dissocier les stratégies de ses adversaires. Seule une estimation de la puissance des interférences agrégées au bruit est requise.

#### 4.2.4.1 Jeu à connaissance parfaite des statistiques du canal

Sous cette hypothèse, le récepteur fournit à l'émetteur le RSIB de l'équation (4.27). Alors, l'émetteur peut chercher, si elle existe, la stratégie satisfaisante ayant le plus petit indice dans son ensemble de stratégies. Celle-ci correspond à la réponse de satisfaction de moindre coût du joueur, telle qu'exprimée par le problème (2.33), dans le cadre de la modulation DSSS.

#### 4.2.4.2 Jeu avec retour d'un signal binaire

Le récepteur calcule son RSIB donné à l'équation (4.27) à partir de ses estimations du canal et de l'interférence plus bruit. Il est également supposé que le récepteur est capable de calculer quelle serait son utilité  $u_i(\bar{a}_i, \mathbf{x}_{-i}(t))$  pour la stratégie  $\bar{a}_i$  où la plus grande puissance de transmission et le plus grand gain d'étalement sont utilisés, et pour la stratégie adverse  $\mathbf{x}_{-i}(t)$  au temps de la mise à jour  $t$ . Si la contrainte  $u_i(a_i(t), \mathbf{x}_{-i}(t)) \geq \Gamma_i$  est satisfaite ou si la stratégie la plus coûteuse ne satisfait pas la contrainte, *i.e.*  $u_i(\bar{a}_i, \mathbf{x}_{-i}(t)) < \Gamma_i$ , alors le récepteur transmet le signal  $S_i = 1$  pour signifier à l'émetteur que sa stratégie actuelle doit être conservée. Dans le cas contraire,  $u_i(\bar{a}_i, \mathbf{x}_{-i}(t)) \geq \Gamma_i$  et  $u_i(a_i(t), \mathbf{x}_{-i}(t)) < \Gamma_i$  ce qui signifie qu'il existe une déviation profitable parmi les stratégies strictement plus coûteuses que  $a_i(t)$ , que l'émetteur aura à tirer aléatoirement après réception du signal  $S_i = 0$  transmis sur le lien de retour.

Pour les deux types de retour, au moins un équilibre de satisfaction existe ; c'est l'équilibre de Nash  $\bar{a} = (n_1, n_2, \dots, n_I)$  où tous les joueurs font l'usage de leur plus grande puissance de transmission et de leur plus grand gain d'étalement. Il est possible

---

3. L'expression du RSIB pourrait être rendue plus réaliste en y intégrant de l'interférence entre symboles résiduelle. Pour des raisons de clarté de l'exposé, et afin de ne pas l'alourdir avec des détails qui ne relèvent pas du message qu'il entend transmettre au lecteur, il reste préférable de se restreindre à des expressions simples pour les fonctions d'utilité.

de prouver la convergence des algorithmes 1 et 3 à partir de la proposition 4.2.1 pour un jeu peuplé de joueurs utilisant la modulation DSSS tels que décrits ici. Ceci est reporté à l'annexe D.1.

### 4.2.5 Étude de cas : Communications OFDM

Dans le cas de joueurs utilisant la modulation OFDM, le vecteur  $\mathbf{x}_i$  de l'équation (4.20) est composé de  $N_i$  symboles i.i.d répartis sur autant de sous porteuses orthogonales, et  $\mathbf{y}_i$  est le vecteur de symboles reçus après transformation de Fourier discrète et suppression du préfixe cyclique. L'espacement entre sous-porteuses du joueur  $i$  est  $\Delta f_i = B/N_i$  et la durée du symbole OFDM est  $T_i = \Delta f_i^{-1} + T'_i$  où  $T'_i$  est la durée du préfixe cyclique choisie supérieure à l'étalement temporel du canal. Le symbole  $x_{i,n}$  du joueur  $i \in \mathcal{I}$  sur la sous-porteuse  $n \in \mathcal{N}_i = \{1, \dots, N_i\}$  est transmis à la puissance  $p_{i,n} = \mathbb{E}[|x_{i,n}|^2]$ , allouée par l'émetteur. La puissance allouée à la totalité de la bande de transmission  $B$  est  $P_i = \sum_{n \in \mathcal{N}_i} p_{i,n}$ . Sous l'hypothèse d'interférence entre sous-porteuses (ICI) négligeable<sup>4</sup>, la matrice de canal  $\mathbf{H}_{ii}$  est diagonale. Chacun de ses coefficients  $(h_{i,n})_{n=1}^{N_i}$  est une réalisation du gain du canal à la sous-porteuse  $n$ , considéré comme une variable aléatoire dont la distribution est inconnue du joueur  $i$ . Selon les mêmes arguments qu'au chapitre 3, la fonction d'utilité est définie à travers le débit moyen théorique d'information dans le cas de l'interférence la plus préjudiciable. On a donc

$$u_i(\mathbf{p}_i, \mathbf{p}_{-i}) = \alpha_i \sum_{n=1}^{N_i} \log(1 + \gamma_{i,n}(\mathbf{p}_{-i})p_{i,n}). \quad (4.29)$$

où  $\alpha_i = (N_i T_i \Delta f_i)^{-1}$  et  $\gamma_{i,n}(\mathbf{p}_{-i})$  l'état statistique du canal à la sous-porteuse  $n$  donné par l'équation (3.10), et dont la connaissance est acquise au récepteur selon les mêmes hypothèses qu'au précédent chapitre (voir 3.1). Le vecteur  $\mathbf{p}_i = [p_i(1), \dots, p_i(N)]^T \in \mathbb{R}_+^N$  est la stratégie d'allocation de puissance sur les  $N_i$  sous-porteuses du joueur  $i$ , et  $\mathbf{p}_{-i} = [\mathbf{p}_1^T, \dots, \mathbf{p}_{i-1}^T, \mathbf{p}_{i+1}^T, \dots, \mathbf{p}_M^T]^T$  sont les allocations des joueurs autres que  $i$ .

Les émetteurs sont soumis à une contrainte de puissance, de sorte que les allocations de puissance d'un joueur  $i$  quelconque appartiennent à l'ensemble suivant

$$\mathcal{P}_i \triangleq \{\mathbf{p}_i \in \mathbb{R}_+^{N_i} : P_i^{\min} \leq \|\mathbf{p}_i\|_1 \leq P_i^{\max}\}. \quad (4.30)$$

Cependant, les choix laissés aux émetteurs quant à leur allocation de puissance dépendent de la connaissance de l'environnement qui leur est fournie par l'émetteur. Dans le cas d'une connaissance telle qu'au chapitre 3 où l'émetteur obtient l'état statistique du canal  $\boldsymbol{\gamma}_{-i} = (\gamma_{i,n}(\mathbf{p}_{-i}))_{n \in \mathcal{N}_i}$ , il est possible de mettre en œuvre une stratégie basée sur le water-filling. En revanche, le retour d'un signal binaire est plus contraignant et impose d'autres stratégies d'adaptation. Le détail de ces difficultés est l'objet des deux paragraphes suivants.

---

4. De la même façon qu'au chapitre 3, considérer l'ICI ne ferait qu'alourdir l'expression de la fonction d'utilité par l'ajout d'un terme d'interférence au dénominateur des RSIB [Scutari et al., 2007].

#### 4.2.5.1 Jeu à connaissance parfaite des statistiques du canal

Ici les joueurs utilisent l'algorithme 1 et les émetteurs disposent de la connaissance de l'état statistique du canal  $\gamma_{-i}$  au moments de la mise à jour de leurs stratégies d'allocation de puissance. Il a été montré à la section 4.1 de ce chapitre que l'unique solution du problème

$$\begin{aligned} \min_{\mathbf{p}_i \in \mathcal{P}_i} \quad & \|\mathbf{p}_i\|_1 \\ \text{sous la contrainte} \quad & u_i(\mathbf{p}_i, \mathbf{p}_{-i}) \geq \Gamma_i \end{aligned} \quad (4.31)$$

est la solution de water-filling donnée par les équations (4.9) et (4.10) où la contrainte  $f_i(\mathbf{p}_i, \mathbf{p}_{-i}) \geq \Gamma_i$  est saturée. Si le problème 4.31 est faisable, il est facile de voir que sa solution est identifiable à l'unique réponse de satisfaction de moindre coût du joueur  $i$  au profil d'interférence  $\mathbf{p}_{-i}$ . De plus, pour une contrainte fixée sur la norme  $\ell_1$  de  $\mathbf{p}_i$ , la solution de water-filling est l'unique maximisateur de la fonction  $u_i(\mathbf{p}_i, \mathbf{p}_{-i})$ . Par conséquent, un joueur rationnel devrait toujours jouer un vecteur de water-filling calculé sur la base de  $\gamma_{-i}$  car tout autre choix de stratégie est soit plus coûteux en terme de puissance soit moins performant en terme de la fonction  $u_i$  - ce qui éloigne le joueur de la satisfaction de sa contrainte de QoS.

Soit

$$\begin{aligned} \mathbf{p}_i^* = \operatorname{argmax}_{\mathbf{p}_i \in \mathcal{P}_i} \quad & f_i(\mathbf{p}_i, \mathbf{p}_{-i}) \\ \text{sous la contrainte} \quad & \|\mathbf{p}_i\|_1 = P_i \end{aligned} \quad (4.32)$$

le vecteur d'allocation de puissance water-filling du joueur  $i$  maximisant sa fonction  $u_i$  lorsque la puissance disponible est  $P_i$ . Il est donc également vrai que

$$\begin{aligned} P_i = \min_{\mathbf{p}_i \in \mathcal{P}_i} \quad & \|\mathbf{p}_i\|_1 \\ \text{sous la contrainte} \quad & f_i(\mathbf{p}_i, \mathbf{p}_{-i}) = f_i(\mathbf{p}_i^*, \mathbf{p}_{-i}). \end{aligned} \quad (4.33)$$

Il est possible d'écrire  $\mathbf{p}_i^* = P_i \times \mathbf{q}_i$  où  $\mathbf{q}_i \in \mathbb{R}_+^{N_i}$  et  $\|\mathbf{q}_i\|_1 = 1$ . Les solutions des problèmes (4.31) et (4.32) coïncident lorsque la contrainte de norme  $P_i$  de (4.32) est choisie de telle sorte que le maximum atteignable de  $u_i$  avec cette norme soit égal à la contrainte  $\Gamma_i$  de (4.31) - ou, inversement, si la contrainte de QoS de (4.31) est choisie de telle sorte que sa solution produise un vecteur de norme égale à  $P_i$ . Considérant alors que les seuls vecteurs d'allocation de puissance rationnels sont ceux prenant la forme du water-filling, les ensembles de stratégies des joueurs OFDM ayant la connaissance de  $\gamma_{-i}$  peuvent se réduire au choix de la puissance allouée sur la totalité de la bande  $B$  donnée par la norme  $\ell_1$  de ces vecteurs. Ainsi,

$$\mathcal{A}_i = [P_i^{\min}, P_i^{\max}], \quad (4.34)$$

et on peut ré-écrire  $u_i$  comme

$$u_i(P_i, \mathbf{p}_{-i}) = \alpha_i \sum_{n=1}^{N_i} \log(1 + P_i \gamma_{i,n}(\mathbf{p}_{-i}) q_{i,n}), \quad (4.35)$$

où  $P_i \in \mathcal{A}_i$ . La correspondance de satisfaction du joueur  $i$  est alors

$$\varphi_i(\mathbf{p}_{-i}) = \{P_i \in \mathcal{A}_i : u_i(P_i, \mathbf{p}_{-i}) \geq \Gamma_i\}. \quad (4.36)$$

La réponse de satisfaction  $SR_i$  du joueur  $i$  devient donc la plus petite contrainte de puissance pour laquelle sa fonction d'utilité  $u_i$  excède  $\Gamma_i$ . D'un point de vue pratique, elle peut être évaluée à travers la résolution du problème (4.31) en vérifiant que la norme de la solution n'excède pas la puissance de transmission maximale  $P_i^{max}$ . Dans le cas contraire, cela signifie que la correspondance de satisfaction  $\varphi_i(\mathbf{p}_{-i})$  est vide car il n'existe pas de  $P_i \in \mathcal{A}_i$  satisfaisant  $u_i(P_i, \mathbf{p}_{-i}) \geq \Gamma_i$ .

Selon les arguments de la section 2.2.1, il existe au moins un équilibre de satisfaction dans le jeu peuplé de joueurs OFDM ayant la connaissance de  $\gamma_{-i}$ . Cet équilibre correspond à l'équilibre de Nash du jeu de water-filling où les joueurs sont contraints en puissance par  $(P_i^{max})_{i \in \mathcal{I}}$ , et dont l'existence est garantie en vertu du théorème de Debreu, Glicksberg et Fan (voir théorème 2.1.2). La convergence de l'algorithme 1 pour ce cas d'étude est discutée à l'annexe D.2.1.

#### 4.2.5.2 Jeu avec retour d'un signal binaire

Dans ce jeu, les émetteur ne disposent pas de la connaissance de  $\gamma_{-i}$  et reçoivent chacun un signal binaire  $S_i$  transmis par leur récepteur respectif à la place. N'étant donc pas capables de mettre en place un stratégies de type water-filling, on considérera que leur choix est limité à un nombre de sous-canaux sur lesquels est allouée une puissance non-nulle constante. Les ensembles de stratégies sont définis par

$$\mathcal{A}_i = \{1, \dots, N_i\}. \quad (4.37)$$

et les allocations de puissance possibles sont

$$\mathcal{P}_i(a_i) = \left\{ \mathbf{p}_i \in \{0, P_i\}^{N_i} : \|\mathbf{p}_i\|_0 = a_i \right\} \quad (4.38)$$

où  $P_i = P_i^{max}/N_i$  et  $a_i \in \mathcal{A}_i$ .

Chaque émetteur est initialisé à sa stratégie la moins coûteuse, en accord avec l'algorithme 3. Celle-ci consiste donc à n'avoir qu'un seul sous-canal, choisit aléatoirement, et sur lequel la puissance  $P_i$  est allouée. Lorsque le signal reçu par l'émetteur est  $S_i = 0$ , ce dernier doit choisir une stratégie plus coûteuse que la précédente. L'émetteur, ne connaissant pas le contenu de sa correspondance  $\varphi_i$ , choisit alors aléatoirement des sous canaux, parmi les sous-canaux non-utilisés, sur chacun desquels la puissance  $P_i$  est allouée. On peut formaliser la procédure comme suit. Pour un vecteur d'allocation de puissance  $\mathbf{p}_i(t) \in \mathcal{P}_i(a_i(t))$  choisi par le joueur  $i$  au temps  $t$ , on pose

$$\mathcal{N}_i(t) = \{n \in \{1, \dots, N_i\} : p_{i,n}(t) \neq 0\} \quad (4.39)$$

l'ensemble des sous-porteuses allouée avec une puissance  $P_i$  par le joueur  $i$ . Au temps  $t + 1$ , cet ensemble est étendu de telle façon que  $|\mathcal{N}_i(t + 1)| > |\mathcal{N}_i(t)|$  avec  $\mathcal{N}_i(t) \subset \mathcal{N}_i(t + 1)$ .

Comme pour le jeu DSSS, l'existence d'au moins un équilibre de satisfaction en stratégies pure est garantie, celui-ci étant l'équilibre de Nash  $\bar{a} = (N_1, \dots, N_I)$ . Une preuve de la convergence de l'algorithme 3 dans le jeu OFDM à retour de 1 bit est donnée en annexe D.2.2.

## 4.2.6 Résultats sur des canaux synthétiques

Cette section présente l'analyse de résultats de simulations pour les 4 jeux venant d'être présentés. Les résultats sont obtenus sur la base de canaux ASM synthétisés par le simulateur de [Socheleau et al., 2010b] décrit au chapitre 3. L'ensemble  $\mathcal{I}$  des joueurs des jeux étudiés est formés par des liens de communications ASM utilisant les modulations ODFM ou DSSS, et on étudie pour chaque cas l'impact du type de lien de retour : état statistique du canal ou signal binaire. Un scénario mettant en jeu les deux modulations en même temps sera étudié à la fin de cette section pour le lien de retour codé sur 1 bit. Les émetteurs transmettent en même temps et sur la même bande de fréquence  $B = 6$  kHz centrée sur  $f_c = 12$  kHz. Les résultats sont évalués en terme de probabilité de satisfaction à l'équilibre en fonction du nombre de joueurs impliqués. On évalue également la puissance de transmission utilisée en moyenne par les joueurs satisfaits, et la vitesse de convergence moyenne à un équilibre en nombre d'itérations.

Les terminaux sont aléatoirement immergés, entre 5 et 20 mètres de profondeur, dans une zone de 1.5 km de rayon où la distance minimale entre deux terminaux différents est contrainte à 500 mètres. La profondeur de la colonne d'eau est de 50 mètres. Un environnement petits fonds est donc supposé une nouvelle fois, cependant les méthodes décrites dans ce chapitre s'appliquent sans modifications aux canaux grands fonds. De la même façon qu'au chapitre 3, section 3.4.1, le simulateur est utilisé pour synthétiser des réponses impulsionnelles de canaux ASM variant dans le temps à partir desquelles on extrait les statistiques permettant de calculer l'état statistique du canal défini à l'équation (3.10) ou (4.27), selon la modulation considérée. Les joueurs peuvent ainsi vérifier leurs contraintes de QoS, à travers le calcul de la fonction d'utilité  $u_i$ . Sur cette base, les joueurs interagissent successivement, au cours d'un certain nombre d'itérations des algorithmes 1 ou 3 jusqu'à ce qu'un équilibre de satisfaction soit atteint. Les résultats de performance exposés dans cette section sont le fruit d'une moyenne de performances à l'équilibres pour 500 jeux lancés avec des immersions différentes (donc des canaux synthétiques différents).

Il est supposé que les canaux sont stationnaires au sens large sur la durée d'un jeu, ou au moins jusqu'à ce que l'équilibre soit atteint. Un rapport signal à bruit de référence, sans interférence, est fixé à 25 dB à 1 km pour la puissance maximale de transmission. La densité spectrale de puissance du bruit est modélisée par le modèle de Knudsen de l'équation (1.7).

### 4.2.6.1 Jeux en modulation DSSS

Le RSIB minimum pour satisfaire la contrainte de QoS est fixé à  $\Gamma_j^{dB})15$  dB pour tous les joueurs. Leur puissance minimum est fixée à  $P_i^{min,dB} = 170$  dB ref  $\mu\text{Pa}$  @ 1m et leurs stratégies sont données dans le tableau 4.1.

Le temps symbole est  $T_i = N_i \times T_c$  avec une durée chip constante de  $T_c \approx 1/B = 166\mu\text{s}$  et  $N_i$  la taille de la séquence d'étalement, qui change en fonction de la stratégie pour laquelle opte le joueur  $i$ .

TABLE 4.1 – Action space of the DSSS players

$a_i$	$P_i^{(a_i)}$	$N_i^{(a_i)}$
1	$1 \times P_i^{min}$	1
2	$1 \times P_i^{min}$	4
3	$4 \times P_i^{min}$	4
4	$4 \times P_i^{min}$	16
5	$16 \times P_i^{min}$	16
6	$16 \times P_i^{min}$	64
7	$64 \times P_i^{min}$	64

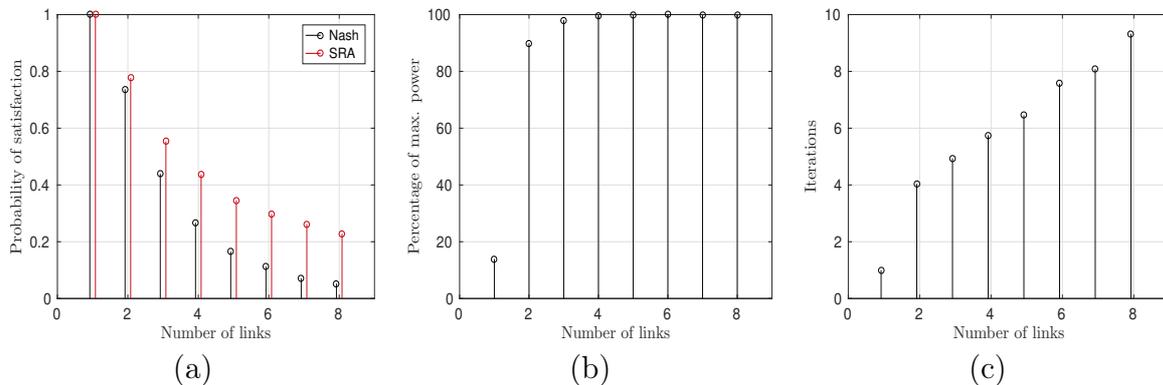


FIGURE 4.4 – Jeu DSSS avec connaissance du RSIB et algorithme 1,  $\Gamma_i^{(dB)} = 15 \text{ dB}$ ,  $\forall i \in I$ . (a) Probabilité de satisfaction, (b) Puissance moyenne utilisée par les joueurs satisfaits, (c) Nombre d'itérations pour le calcul de l'équilibre.

L'équilibre de Nash est  $\bar{a} = (\bar{a}_1 = 7, \dots, \bar{a}_I = 7)$ , où tous les joueurs utilisent la puissance maximale et le plus grand gain d'étalement au détriment à la fois de l'efficacité spectrale et de l'efficacité énergétique. Les résultats sont donnés aux figures 4.4 et 4.5

**Connaissance du RSIB** Les résultats du jeu avec connaissance du RSIB à l'émetteur sont donnés à la figure 4.4. La figure 4.4-(a) donne la probabilité de satisfaction moyenne par joueur à l'équilibre et la comparaison est faite avec la stratégie d'équilibre de Nash. Le comportement prédit par l'équilibre de satisfaction, en comparaison du comportement maximisateur de l'équilibre de Nash, offre de meilleures chances d'être satisfait. Ceci est d'autant plus vrai que le nombre de joueurs augmente. L'interférence est maximum à l'équilibre de Nash, tandis qu'à l'équilibre de satisfaction, les joueurs utilisent des niveaux de puissance plus bas. Lorsque les joueurs sont plus nombreux, l'équilibre particulier ou aucun des joueurs n'est satisfait apparaît plus souvent lorsqu'ils jouent l'équilibre de Nash que lorsqu'ils apprennent un équilibre de satisfaction. Pour un petit nombre de joueurs (1 à 3), la figure 4.4-(b) montre que la puissance utilisée par les joueurs satisfaits est moindre, comparée à l'équilibre de Nash et pour une probabilité d'être satisfait légèrement plus élevée. Ceci illustre le choix de la réponse de satisfaction définie par l'équation (2.33), et qui oblige les joueurs à toujours utiliser les

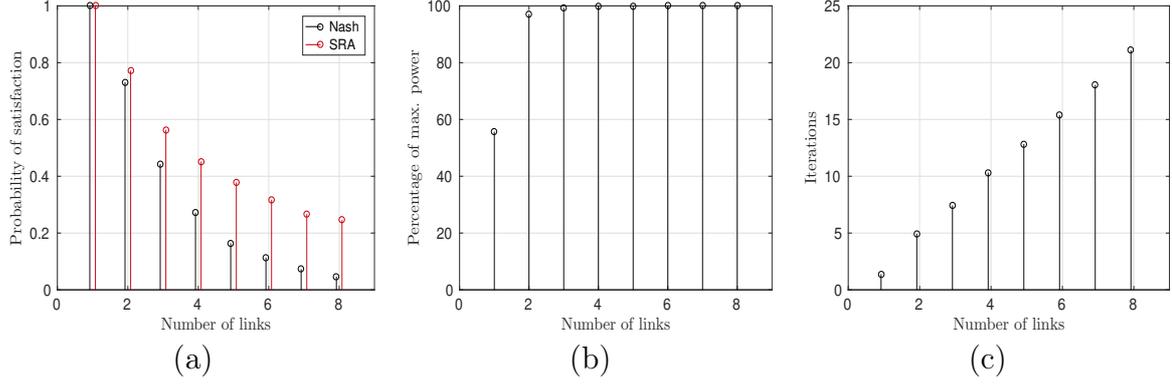


FIGURE 4.5 – Jeu DSSS avec signal binaire et algorithme 3,  $\Gamma_i^{(\text{dB})} = 15 \text{ dB}$ ,  $\forall i \in \mathcal{I}$ . (a) Probabilité de satisfaction, (b) Puissance moyenne utilisée par les joueurs satisfaits, (c) Nombre d’itérations pour le calcul de l’équilibre.

stratégies les moins coûteuses. Le nombre d’itérations avant la convergence de la figure 4.4-(c) ne se traduit en temps de convergence que lorsqu’il est mis en relation avec un schéma de mise à jour particulier à une période donnée. Par exemple, considérant un algorithme séquentiel où les joueurs ont chacun une période de mise à jour de  $T_{obs} = 20$  secondes et jouent l’un après l’autre, le temps de convergence, quelque soit le nombre de joueurs, est d’environ 40 secondes en moyenne. (Noter qu’il y a dans tous les cas 20 secondes s’écoulant entre deux itérations jouées par le même joueurs, tous les joueurs jouant l’un après l’autre). On notera qu’en moyenne les joueurs jouent entre 1 à 2 fois chacun seulement avant d’atteindre l’équilibre. Concernant le temps de convergence, la même analyse pourra être appliquée dans les jeux qui suivent.

**Connaissance sous la forme d’un signal binaire** Les résultats du jeu à signal binaire pour seule connaissance de l’environnement à l’émetteur sont donnés à la figure 4.5. À l’exception des jeux à 1 et 2 joueurs, peu de différences sont visibles en termes de performances moyennes par rapport au cas où l’émetteur connaît son RSIB. La principale différence réside dans la puissance consommée dans les jeux à 1 et 2 joueurs, visible à la figure 4.5-(b), et s’explique par le choix aléatoire d’une stratégie de transmission nécessairement supérieure ou égale à la réponse de satisfaction de moindre coût donnée par (2.33). Le choix aléatoire des stratégies a, de plus, pour effet de ralentir la convergence vers un équilibre, et ceci en raison du fait que les choix stratégiques des joueurs peuvent ne pas appartenir à la réponse de satisfaction qu’aurait retournée la connaissance du RSIB. Par exemple, on peut imaginer un cas où le joueur  $i$  choisit aléatoirement entre les stratégies  $a_i = 2$  et  $a_i = 7$  de la table 4.1, n’ayant pas été satisfait par la stratégie  $a_i(t) = 1$ , alors que sa correspondance de satisfaction aurait retourné  $\varphi_i(a_{-i}(t)) = 6, 7$ . Considérant, comme au précédent paragraphe, un schéma de mise à jour séquentiel avec une période de 20 secondes par joueur, le temps de convergence reste sous les 60 secondes.

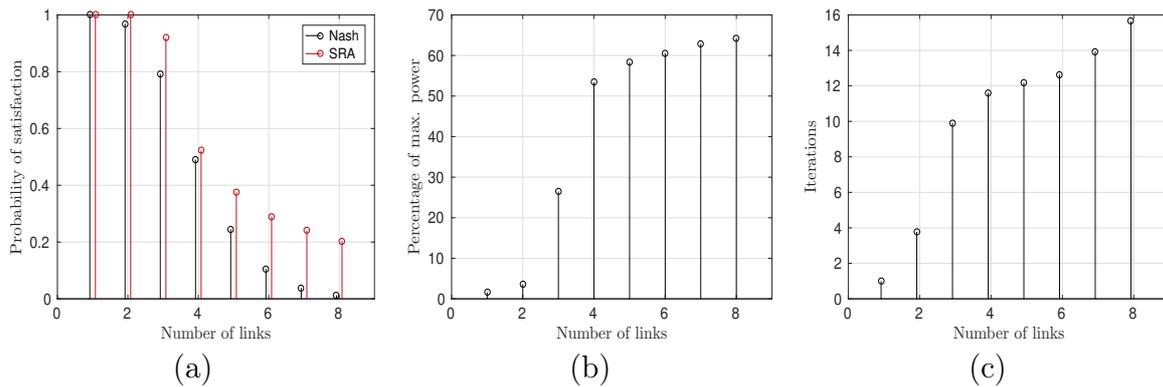


FIGURE 4.6 – Jeu OFDM avec connaissance de l’état statistique du canal et algorithme 1,  $\Gamma_i = 1$  bit/s/Hz,  $\forall i \in \mathcal{I}$ . (a) Probabilité de satisfaction, (b) Puissance moyenne utilisée par les joueurs satisfaits, (c) Nombre d’itérations pour le calcul de l’équilibre.

#### 4.2.6.2 Jeux en modulation OFDM

Pour les jeux OFDM, la bande  $B = 6$  kHz est divisée en  $N_i = N = 256$  sous-porteuses,  $\forall i \in \mathcal{I}$ . Ainsi  $B = N \times \Delta f$ ,  $\Delta f = 23.43$  Hz, et le temps symbole est  $T = T_g + \Delta f^{-1} = 57.7$  ms, où  $T_g = 15$  ms est la durée du préfixe cyclique. La contrainte de QoS est fixée pour tout joueur  $\Gamma_i = 1$  bit/s/Hz.

**Connaissance de l’état statistique du canal** Dans ce jeu, les stratégies sont les puissances auxquelles les joueurs contraignent leurs vecteurs d’allocation water-filling. Les puissances minimum et maximum sont fixées respectivement à  $P_i^{min} = 170$  et  $P_i^{max} = 190$  dB ref.  $\mu\text{Pa}$  @ 1m. Les résultats du jeu considérant la connaissance de l’état statistique du canal à l’émetteur sont affichés à la figure 4.6.

Les figures 4.6-(a) et 4.6-(b) montrent que les gains des performances entre la solution d’équilibre de Nash et l’équilibre de satisfaction s’expriment principalement en termes de puissance consommée. À moins de 4 joueurs, la probabilité de satisfaction pour l’équilibre de Nash reste élevée du fait du choix de la contrainte de la QoS et du rayon de la zone de transmission. La puissance consommée est maintenue en moyenne sous les 70% de  $P_i^{max}$  car l’algorithme 1 force les joueurs à minimiser la puissance allouée sur la bande de transmission. Le nombre d’itérations pour atteindre un équilibre est faible, avec 2 à 3 itérations par joueur.

**Connaissance sous la forme d’un signal binaire** Ici, il est considéré que chaque joueur possède un nombre différent de sous-canaux, choisis uniformément dans  $\{4, 8, 16, 32\}$  à chaque initialisation d’une réalisation d’un jeu. Ces sous-canaux prennent la forme d’une agrégation d’un certain nombre de sous-porteuses OFDM adjacentes, ces sous-porteuses étant au nombre de  $N = 256$  par ailleurs. La puissance allouée par un joueur  $i$  sur un sous-canal particulier est soit  $P_i$  soit 0, avec  $P_i = P_i^{max}/N = P_i^{min}$ . La puissance maximale est fixée à  $P_i^{max} = 190$  dB ref.  $\mu\text{Pa}$  @ 1m.

Les résultats sont affichés à la figure 4.7. Les probabilités de satisfaction sont

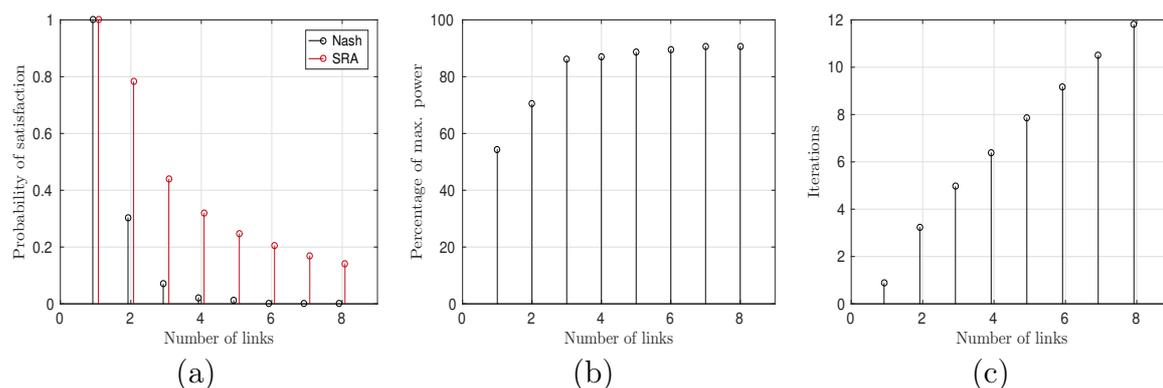


FIGURE 4.7 – Jeu OFDM avec signal binaire et algorithme 3,  $\Gamma_i = 1$  bit/s/Hz,  $\forall i \in \mathcal{I}$ . (a) Probabilité de satisfaction, (b) Puissance moyenne utilisée par les joueurs satisfaits, (c) Nombre d'itérations pour le calcul de l'équilibre.

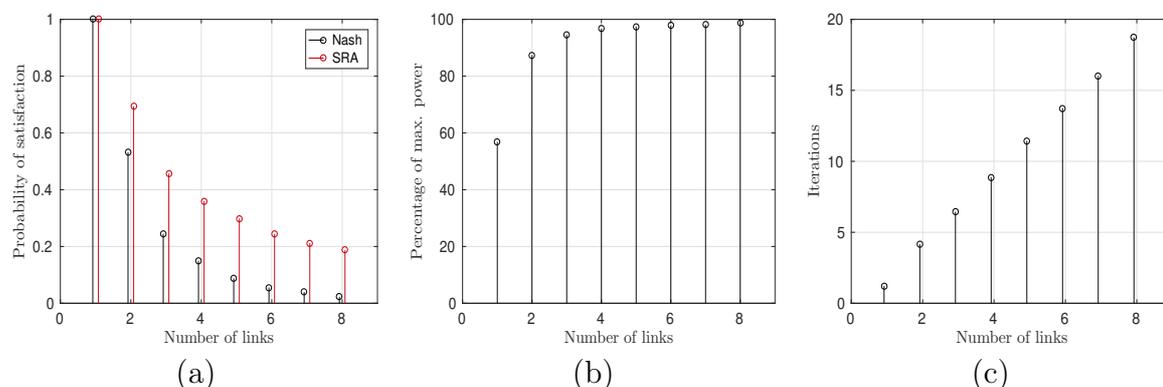


FIGURE 4.8 – Jeu mélangeant OFDM et DSSS avec signal binaire et algorithme 3,  $\Gamma_i = 1$  bit/s/Hz pour l'OFDM et  $\Gamma_i^{(\text{dB})} = 15$  dB pour le DSSS,  $\forall i \in \mathcal{I}$ . (a) Probabilité de satisfaction, (b) Puissance moyenne utilisée par les joueurs satisfaits, (c) Nombre d'itérations pour le calcul de l'équilibre.

considérablement accrues à l'équilibre de satisfaction en comparaison de l'équilibre de Nash, ce dernier consistant donc à transmettre selon une densité spectrale de puissance constante sur la bande et de puissance  $P_i^{max}$ . Le peu d'information sur l'environnement disponible à l'émetteur réduit néanmoins les performances comparées au cas du jeu à connaissance de l'état statistique du canal. À la figure 4.6-(a), le faible écart entre les deux équilibres s'explique par une meilleure corrélation des allocations de puissance water-filling à la fois à la sélectivité fréquentielle du canal de communication et aux niveaux d'interférences générées par les autres joueurs. Dans le jeu à signal de retour binaire, les allocations de puissances à l'équilibres de satisfaction ont été choisies aléatoirement, tandis qu'à l'équilibre de Nash aucune adaptation au canal ou à l'interférence n'est réalisée. Comme la puissance allouée aux sous-canaux est fixée, il est également logique que la puissance utilisée soit relativement élevée. En termes de convergence, 2 à 3 itérations par joueurs sont nécessaires.

### 4.2.6.3 Systèmes hétérogènes avec lien de retour binaire

Dans un dernier scénario sont évaluées les performances à l'équilibre de satisfaction atteint par un ensemble de joueurs hétérogènes, mélangeant systèmes OFDM et DSSS. On reprend les mêmes paramètres qu'aux paragraphes précédents traitant des jeux DSSS et OFDM dans le cas du lien de retour à signal binaire. À chaque réalisation d'un nouveau jeu, chaque joueur se voit attribué son type de modulation aléatoirement avec une probabilité  $1/2$  pour chacun des types. La figure 4.8 illustre les résultats obtenus. Comme attendu, l'équilibre de satisfaction améliore les performances en termes d'accès réussi au canal en comparaison de la stratégie agressive d'équilibre de Nash. Notons par ailleurs que les joueurs DSSS bénéficient du gain de traitement offert par les séquences d'étalement et génèrent aussi une interférence relativement difficile à combattre pour les systèmes OFDM, du fait de la répartition uniforme de leur puissance sur la bande de transmission. Comme pour les jeux OFDM et DSSS, les économies réalisées en puissance consommée sont plus significatives lorsque le nombre de systèmes en compétition dans la zone est petit. Le temps de convergence vers l'équilibre est du même ordre de grandeur que dans les autres jeux.

## 4.3 Conclusion

En réponse aux conclusions du chapitre 3 et aux problèmes qu'elles soulèvent, ce chapitre a proposé plusieurs méthodes de partage non-coopératif du canal ASM en posant, pour chaque système impliqué, une contrainte de satisfaction de qualité de service (QoS) exprimée par une métrique fonction du débit théorique d'information ou du RSIB. La formulation d'un jeu de partage de spectre où l'objectif est la minimisation de la puissance de transmission sous contrainte de débit conduit à la résolution d'un problème d'équilibre de Nash généralisé (selon la terminologie proposée par [Facchinei and Kanzow, 2007]), ce qui pose certaines difficultés, notamment vis-à-vis de l'existence non-garantie d'une solution d'équilibre et de la phase de convergence de l'algorithme d'apprentissage basé sur les meilleures réponses. Ceci a été vu à la section 4.1.

Ces difficultés ont été contournées grâce aux jeux de satisfaction à la section 4.2. Les équilibres de satisfaction généralisés [Goonewardena et al., 2017] offrent une solution plus flexible au problème en permettant que certains joueurs ne soient pas satisfaits à l'équilibre, tout en restant dans un cadre non-coopératif, décentralisé, et où les systèmes de communication ASM peuvent décider de leurs stratégies de transmission à partir d'une information locale sur leur environnement. Bien que cette solution soit un moyen simple de contourner les difficultés théoriques posées par l'équilibre de Nash généralisé, le risque existe, cependant, que le taux de satisfaction des joueurs soit bas. En d'autres termes, il est possible, selon les configurations et l'état des canaux, que peu de joueurs satisfassent leurs contraintes de QoS aux équilibres appris par les dynamiques de réponses de satisfaction. L'algorithme d'apprentissage des équilibres de satisfaction a été amélioré en considérant que l'information donnée aux joueurs sur leur environnement est réduite à un signal de retour codé sur 1 bit. Ceci présente un grand intérêt dans le contexte des communications ASM où les bandes passantes

sont limitées, et où, par conséquent, il est crucial d'éviter leur surcharge. Plusieurs exemples d'applications ont été donnés, en considérant des communications en OFDM et/ou DSSS. Les résultats montrent que les équilibres de satisfaction peuvent apporter des gains individuels en termes de probabilités de satisfaction de la contrainte de QoS, ainsi qu'en diminution de la puissance de transmission.



# Chapitre 5

## Réseaux de capteurs sous marins décentralisés et adaptatifs

Ce dernier chapitre propose une ouverture sur les réseaux de communications sous-marins adaptatifs et décentralisés, et l'intérêt va plus précisément se porter sur un exemple d'application aux réseaux de capteurs d'un algorithme d'apprentissage par renforcement.

Les réseaux de capteurs sous-marins ont attiré l'attention de la recherche en communications ASM depuis un certain temps, et de nombreux travaux ont été menés tant au niveau de la couche physique et d'accès multiple que sur des couches réseau plus élevées [Casari and Zorzi., 2011; Jornet et al., 2010; Kilfoyle and Baggeroer, 2000; Otnes et al., 2012; Sozer et al., 2000; Stojanovic et al., 2008]. Bien que quelques protocoles d'accès au médium aient été développés spécifiquement pour les réseaux sous-marins [Molins and Stojanovic, 2006; Syed et al., 2008], un grand nombre de travaux sur ce sujet adaptent plus ou moins bien des techniques issues des réseaux terrestres sans forcément prendre en compte toutes les contraintes posées par l'environnement marin. D'une manière générale, les protocoles d'accès au médium fixes tels que l'accès multiple par multiplexage en fréquences, en temps ou en codes (FDMA, TDMA et CDMA) présentent l'inconvénient d'être peu flexibles et n'offrent pas de possibilité de s'adapter aux variations du canal et aux changements d'états du réseau. De plus, une synchronisation entre émetteurs peut être requise (au moins pour TDMA), ce qui est difficile à réaliser et à maintenir dans un contexte où les délais de propagation sont de l'ordre du temps de cohérence du canal et sans signal GPS permettant de se baser sur une horloge commune précise. Les protocoles basés sur la contention et l'écoute du canal sont plus flexibles (par exemple, ALOHA, CSMA, etc...) mais sont également sensibles aux délais de propagation et doivent être optimisés en fonction de la répartition géographique des émetteurs. Les protocoles proposés par [Molins and Stojanovic, 2006; Syed et al., 2008] sont basés sur la contention et sont des variantes des protocoles CSMA. Enfin, les protocoles basés sur la découverte du réseau et les « handshakes » (par exemple, Request to Send / Clear to Send dans la norme IEEE 802.11) génèrent de la surcharge de bande passante due aux échanges de messages.

Dans le cadre du travail de cette thèse, l'objectif général serait de concevoir

des méthodes d'accès à la fois asynchrones, ne nécessitant pas ou peu de d'échanges de messages, n'étant pas basées sur l'écoute du canal, et pouvant conférer aux éléments du réseau certaines facultés d'adaptation de leur stratégie de transmission aux fluctuations de l'environnement. Les méthodes basées sur l'apprentissage par renforcement semblent pouvoir apporter ce genre de propriétés lorsqu'elles sont utilisées pour concevoir des protocoles d'accès au médium. L'algorithme  $Q$ -learning a notamment prouvé son efficacité dans les réseaux sans fils terrestres en apportant une certaine flexibilité aux schémas basés sur le TDMA [Chu et al., 2015; Konusalp et al., 2016], ainsi qu'à l'accès au spectre dans les réseaux LTE [Morozs et al., 2016]. Cette idée va être exploitée ici pour gérer un réseau de capteurs déployés au fond de la mer, dédiés à la collecte de mesures physiques (pression, température). Ces capteurs sont supposés devoir remonter leur données régulièrement à la surface en communiquant, par voie acoustique, avec des bouées de surface considérées comme récepteurs, ou points d'accès. Les capteurs sont munis de stratégies de transmission qui consistent en un choix de récepteur auquel envoyer les paquets de données. Ces paquets sont générés de manière asynchrone et peuvent entrer en collision à la réception. On considère qu'il est acceptable qu'une certaine proportion de paquets soit perdus ou reçus sans être décodés correctement. Cette hypothèse est réaliste dans le cadre d'une application dédiée à l'acquisition d'une grandeur physique telle que la pression ou la température, où il est possible, du fait de la forte corrélation spatiale du phénomène mesuré, de reconstruire les données manquantes ou erronées grâce à des techniques de traitement du signal [Fazel et al., 2011, 2013]. L'usage de l'algorithme  $Q$ -learning va permettre aux émetteurs de choisir de manière autonome quel récepteur maximise la probabilité de transmission réussie de leurs paquets. Il offre également une certaine capacité d'adaptation grâce à l'apprentissage se réalisant sur la base d'essais/erreurs, où une récompense encodée sur un bit est re-transmise par les récepteurs aux émetteurs ayant réussi leur transmission.

Le modèle du réseau et des canaux ASM correspondant à l'application est présenté dans un premier avant traiter plus en détail le problème de la sélection des bons récepteurs. La deuxième section présente succinctement le problème de l'apprentissage par renforcement, afin d'aboutir à l'algorithme  $Q$ -learning et d'illustrer son utilisation par l'exemple d'application considérée ici. Des résultats de simulations montrent l'intérêt que peut avoir de telles méthodes pour les réseaux sous-marins, et en particulier la flexibilité qui peut leur être apportée.

## 5.1 Modèle du réseau

Un ensemble  $\mathcal{I}$  de  $I$  émetteurs non-coopératifs transmet des paquets à un ensemble  $\mathcal{A}$  de  $A$  récepteurs. Les récepteurs sont supposés être des points d'accès, tels que des bouées de surface, des nœuds de relais ou de collecte des paquets transmis. Chaque émetteur  $i \in \mathcal{I}$  n'accède qu'à un récepteur  $a \in \mathcal{A}$  à la fois et interfère à la fois sur les autres récepteurs et, éventuellement, sur les paquets transmis par les autres émetteurs au même récepteur. On considérera que  $I > A$ , chaque récepteur peut alors être utilisé par plusieurs émetteurs à la fois. Les récepteurs décodent les paquets de chaque utilisateur séparément<sup>1</sup>, les autres paquets étant considérés comme de l'interfé-

---

1. En anglais, "single-user decoding".

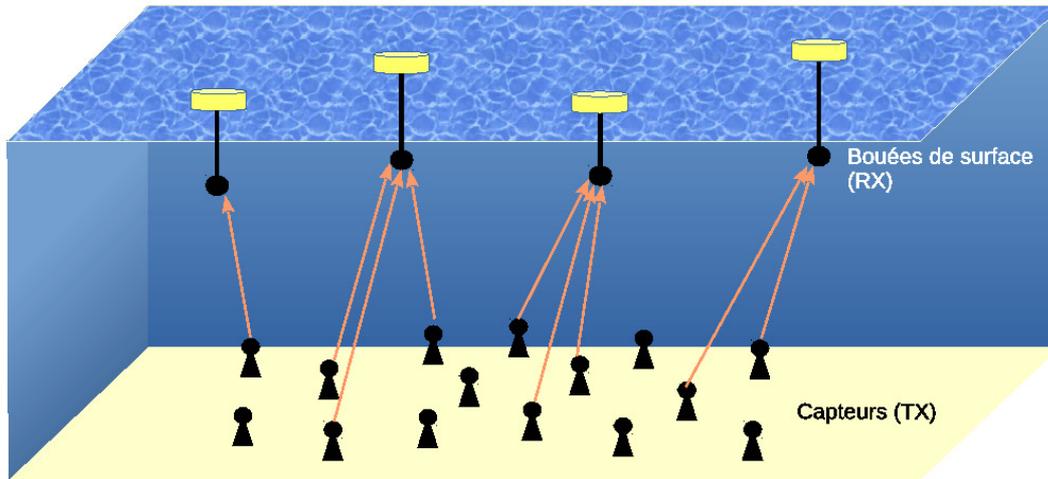


FIGURE 5.1 – Réseau de capteurs sous marins et de bouées de surface.

rence s'ajoutant au bruit. Les émetteurs ne sont pas synchronisés, ni entre eux ni avec les récepteurs. La figure 5.1 illustre le scénario envisagé.

Chaque émetteur génère aléatoirement des paquets indépendants de durée  $T_p$  à un rythme de  $\lambda$  paquets par seconde. Les temps d'arrivées des paquets sont distribués par la loi exponentielle  $\mathcal{E}(\lambda)$ . En moyenne,  $I \times \lambda$  paquets par seconde sont transmis au sein du réseau. Pour chaque paquet qu'il génère, l'émetteur doit choisir un récepteur qui maximise la probabilité que le paquet soit décodé sans erreur. On considère des transmissions en étalement de spectre par séquence (DSSS) afin de profiter des gains en rapport signal sur interférence plus bruit (RSIB) à la réception offerts par cette modulation. Un code d'étalement est associé à chaque récepteur. Les émetteurs connaissent les codes d'étalement de chaque récepteur, de sorte que le choix d'un récepteur pour le paquet généré se réduise au choix du code d'étalement utilisé pour le transmettre. On suppose que tous les codes ont la même longueur et offrent le même gain d'étalement noté  $G$ . Les récepteurs sont seulement capables de décoder les paquets encodés avec leur code d'étalement. Les transmissions de paquets sont asynchrones, ainsi les codes des différents récepteurs ne peuvent pas être considérés comme étant orthogonaux (ou alors avec une probabilité proche de 0). Par conséquent les émetteurs peuvent provoquer de l'interférence au niveau de tous les récepteurs.

Considérons un émetteur  $i$  particulier qui transmet un paquet de durée  $T_p$  au récepteur  $a$ . Le paquet est transmis à une puissance  $P_i$ , répartie uniformément sur la bande  $B$  centrée à la fréquence  $f_c$ , et arrive au récepteur au temps d'arrivée  $t_i$ . Ce paquet a subi une atténuation qui dépend de la fonction de transfert du canal entre  $i$  et  $a$ . Cette même fonction de transfert varie dans le temps, selon l'état de la mer et les déplacements des émetteurs et des récepteurs induits par le mouvement de l'eau. Comme il l'a été expliqué au chapitre 1, il est possible de décomposer les variations temporelles sur canal en deux composantes : une partie spéculaire, dont les évanouissements font varier la puissance reçue lentement en comparaison de la durée des paquets, et une partie diffusée à évanouissement rapides, dont les temps de cohérence sont petits devant  $T_p$ . Soit  $H_{i,a}(t, f)$  la fonction de transfert aléatoire et variant dans le

temps du canal entre l'émetteur  $i$  et le récepteur  $a$ . L'atténuation moyenne subie par le paquet arrivé au temps  $t_i$  peut être exprimée par

$$\bar{\rho}_{i,a}(t_i) = \frac{1}{T_p} \int_{t_i}^{t_i+T_p} \rho_{i,a}(t) dt \quad (5.1)$$

où

$$\rho_{i,a}(t) = \frac{1}{B} \int_{f_c-B/2}^{f_c+B/2} |H_{i,a}(f, t)|^2 df \quad (5.2)$$

est une réalisation d'un processus aléatoire sous-jacent à  $H_{i,a}(t, f)$ .

Durant la réception d'un paquet de l'émetteur  $i$ , des collisions peuvent survenir depuis d'autres émetteurs dès lors qu'il existe un  $j \neq i \in \mathcal{I}$  tel que le temps d'arrivée de son paquet  $t_j$  au même récepteur soit dans l'intervalle  $[t_i - T_p, t_i + T_p]$ . La puissance de l'interférence provoquée est alors définie par la variable aléatoire

$$I_{i,a}(t_i) = \frac{1}{T_p} \int_{t_i}^{t_i+T_p} \sum_{j \neq i} \rho_{j,a}(t) P_j \mathbb{1}_{[t_j, t_j+T_p]}(t) dt \quad (5.3)$$

qui dépend à la fois du processus d'évanouissement du canal ASM et du temps d'arrivée aléatoire des paquets.

Le succès de la transmission d'un paquet depuis l'émetteur  $i$  vers le récepteur  $a$  est évalué par le RSIB moyen de réception donné par

$$\gamma_{i,a}(t_i) = \frac{\bar{\rho}_{i,a}(t_i) P_i}{\sigma_a^2 + I_{i,a}(t_i)} \times G, \quad (5.4)$$

où  $\sigma_a^2$  est la puissance du bruit ambiant au récepteur  $a$ . On suppose que le paquet est décodé avec succès si le RSIB moyen satisfait une certaine contrainte  $\gamma_{i,a} \geq \Gamma$ .

En pratique, le RSIB devrait être estimé au récepteur grâce à des symboles pilotes insérés dans les paquets, ce qui réduit le débit d'information utile. Le choix d'un critère de satisfaction basé sur le RSIB moyen n'est qu'un exemple possible, et on pourra choisir d'autres critères selon les applications et les types de transmissions considérés. L'algorithme de  $Q$ -learning que l'on utilise par la suite se généralise à tout type de critère de performance. On pourrait choisir par exemple un critère de taux d'erreur binaire, ou vérifier le bon décodage du paquet par un code à redondance cyclique. Par conséquent la connaissance parfaite du RSIB n'est pas une hypothèse forte dans ce qui sera présenté ensuite.

Les émetteurs sont considérés comme des agents dans un problème d'apprentissage où chacun doit déterminer un choix optimal de point d'accès, selon l'état de l'environnement perçu à travers des signaux de retours fournis par les récepteurs. Ces signaux prennent la forme d'une variable binaire  $R = +1$  ou  $R = -1$  selon que le récepteur a décodé le paquet de l'émetteur  $i$  avec succès ou non. Un émetteur ayant appris quels sont les points d'accès optimaux pour le décodage de ses paquets devrait ainsi choisir les récepteurs qui maximisent sa probabilité d'obtenir un retour  $R = +1$ . Ceci peut être formulé comme un problème d'apprentissage par renforcement.

## 5.2 Sélection d'un point d'accès avec $Q$ -learning

Dans cette section sont tout d'abord rappelées les bases théoriques de l'apprentissage par renforcement dans le cadre des processus de décisions markoviens. L'algorithme  $Q$ -learning et les liens qui l'unissent à ces problèmes seront détaillés, de même que les aspects pratiques de sa mise en œuvre dans un problème tel que celui exposé au début de ce chapitre.

### 5.2.1 Apprentissage dans un processus de décision de Markov

Le problème de l'apprentissage par renforcement s'idéalise mathématiquement dans les processus de décision markoviens [Sutton and Barto, 1998], pour lesquels des descriptions et des résultats théoriques sont bien établis. Ne seront donnés ici que les concepts nécessaires à la compréhension du problème d'apprentissage par renforcement et à l'introduction de l'algorithme  $Q$ -learning à partir de ces concepts. Le lecteur pourra se reporter à l'ouvrage [Sutton and Barto, 1998] pour un traitement plus complet de ces sujets.

Les processus de décision markoviens formalisent toute situation où un agent a pour but d'apprendre à réaliser un objectif grâce à des interactions répétées avec son environnement. L'environnement modélise, de manière générale, toute chose utile au problème qui est extérieure à l'agent et non-contrôlable par lui. Pour réaliser ses objectifs, l'agent est muni d'actions  $a \in \mathcal{A}$ . L'environnement est décrit par un processus stochastique  $X = (X_t)_{t \geq 0}$  prenant ses valeurs dans un ensemble  $\mathcal{X}$ , et possédant la propriété de Markov

$$\mathbb{P}[X_{t+1} \mid X_t, X_{t-1}, \dots, X_0] = \mathbb{P}[X_{t+1} \mid X_t] \quad (5.5)$$

indiquant que l'état de l'environnement à un instant donné n'est déterminé que par l'état à l'instant immédiatement précédent. Pour chaque temps  $t = 0, 1, 2, \dots$ , l'agent reçoit une représentation de l'état de l'environnement  $x_t = X_t(\omega)$  et sélectionne une action  $a_t$  selon une règle de décision  $\pi_t(x_t)$  - aussi appelée *stratégie*. La sélection peut se faire de manière déterministe ou stochastique, dans ce second cas  $a_t = A_t(\omega)$  où la variable aléatoire  $A_t$  suit la distribution de probabilité  $\pi_t(x_t)$ . À chaque transition d'un état à un autre, l'agent reçoit un retour  $R_{t+1} = f(X_t, A_t, X_{t+1})$  fonction de l'action choisie et des états. La transition vers l'état  $X_{t+1}$  est conditionnée par l'état précédent  $X_t$  et, généralement, aussi par l'action  $A_t$ . On notera  $\forall(x, y) \in \mathcal{X}^2$  et  $\forall a \in \mathcal{A}$

$$p(y \mid x, a) = \mathbb{P}[X_{t+1} = y \mid X_t = x, A_t = a] \quad (5.6)$$

la probabilité de transition de l'état  $x$  vers l'état  $y$  pour l'action  $a$  sélectionnée à l'état  $x$ . Le retour espéré pour l'action  $a$  à l'état  $x$  est donné par

$$r(x, a) = \mathbb{E}[R_{t+1} \mid X_t = x, A_t = a]. \quad (5.7)$$

L'agent a pour but de déterminer une séquence de stratégies  $\pi = (\pi_t)_{t \geq 0}$ , appelée *politique*, qui maximise la somme des retours reçus dans le temps. En général dans

un problème d'apprentissage par renforcement, l'agent ne connaît ni ne contrôle parfaitement son environnement<sup>2</sup>, et ne connaît pas non plus les probabilités de transition entre état pour chacune de ses actions. La détermination des stratégies optimales se fait par alors par essais successifs et évaluation des retours reçus. L'agent est supposé apprendre grâce aux retours obtenus pour les actions qu'il a sélectionné dans les états qu'il a rencontré, et choisir les actions offrant le plus de chances de maximiser la somme cumulée de ses retours. Il est ainsi *renforcé* dans ses choix au fur et à mesure de ses interactions avec l'environnement.

La vaste majorité des algorithmes d'apprentissage par renforcement consiste à estimer une *fonction de valeur*, exprimant la somme escomptée des retours que l'agent espère obtenir à partir d'un état donné  $x$ , au temps  $t$ , en appliquant ensuite la politique  $\pi$  depuis cet état [Sutton and Barto, 1998] :

$$V_t^\pi(x) = \mathbb{E} \left[ \sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \mid X_t = x \right], \quad (5.8)$$

où l'espérance est prise conjointement par rapport aux distributions des actions et des états de l'environnement. Cette fonction est une image de la satisfaction qu'a l'agent de se trouver à l'état  $x$  sachant qu'il met en œuvre les stratégies données par la politique  $\pi$ . On définit également la fonction *action-valeur* exprimant la somme escomptée des retours que l'agent espère obtenir à partir d'un état donné  $x$  et d'une action  $a$  jouée au temps  $t$  en appliquant ensuite la politique  $\pi$  depuis cet état :

$$Q_t^\pi(x, a) = \mathbb{E} \left[ \sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \mid X_t = x, A_t = a \right]. \quad (5.9)$$

Cette fonction peut se développer comme

$$Q_t^\pi(x, a) = \mathbb{E} [R_{t+1} + \gamma V_t^\pi(X_{t+1}) \mid X_t = x, A_t = a] \quad (5.10)$$

$$= r(x, a) + \gamma \int V_t^\pi(y) dp(y \mid x, a) \quad (5.11)$$

à travers la récurrence impliquée par la somme escomptée et la propriété de Markov du processus stochastique modélisant l'environnement. Cette dernière équation est appelée *équation de Bellman pour  $Q_t^\pi$* . On peut trouver l'équation de Bellman pour la fonction  $V_t^\pi$  en intégrant le terme  $r(x, a)$  par rapport à la distribution  $\pi_t$  de sélection des actions au temps  $t$  :

$$V_t^\pi(x) = \mathbb{E} [R_{t+1} + \gamma V_{t+1}^\pi(X_{t+1}) \mid X_t = x]. \quad (5.12)$$

Pour une politique optimale  $\pi^*$ , la fonction de valeur est maximisée et vérifie donc

$$V_t^*(x) \triangleq \max_{\pi} V_t^\pi(x) \quad \forall x \in \mathcal{X}. \quad (5.13)$$

---

2. Noter que Sutton et Barto [Sutton and Barto, 1998] prennent l'exemple du puzzle comme problème d'apprentissage par renforcement dans lequel l'agent aurait une *connaissance complète et parfaite* de son environnement, mais devrait néanmoins procéder par essais successifs. Ainsi, l'apprentissage viendrait aussi combler un manque de "puissance de calcul".

De même, pour les fonctions action-valeur on aura  $\forall a \in \mathcal{A}$  et  $\forall x \in \mathcal{X}$

$$Q_t^*(x, a) \triangleq \max_{\pi} Q_t^{\pi}(x, a) \quad (5.14)$$

$$= \mathbb{E} [R_{t+1} + \gamma V_t^*(X_{t+1}) \mid X_t = x, A_t = a]. \quad (5.15)$$

La dernière équation montre que, pour une action  $a$  et un état de départ  $x$ , une politique optimale produit un retour égal au maximum espéré des fonctions de valeur, auquel s'ajoute la récompense espérée pour l'état  $x$  et l'action  $a$ , cette dernière étant ici quelconque. En toute logique, pour maximiser la fonction de valeur prise au temps  $t$ , il faudrait choisir l'action qui maximise la récompense immédiate à l'état  $x$  et suivre la politique optimale  $\pi^*$  à partir de cet état - puisqu'elle maximise la fonction de valeur espérée pour la suite des états suivants. Ce principe est connu en programmation dynamique comme le *Principe d'optimalité de Bellman* [Bellman, 1957]. On a donc

$$V_t^*(x) = \max_{a \in \mathcal{A}} Q_t^*(x, a). \quad (5.16)$$

Ce principe d'optimalité et les définitions précédentes vont nous servir à dériver l'algorithme  $Q$ -learning [Watkins and Dayan, 1992] qui permet, idéalement, de mettre en œuvre itérativement une politique optimale au regard de la somme escomptée des retours espérés.

## 5.2.2 Algorithme $Q$ -learning

Dans ce paragraphe on suppose que les espaces d'actions et d'états de l'environnement  $\mathcal{A}$  et  $\mathcal{X}$  sont finis et dénombrables. L'agent maintient un tableau où sont stockées les valeurs de la fonctions action-valeur  $Q_t^*(x, a)$  pour chaque couple action-état  $(x, a)$ . Ces valeurs ne sont pas connues à l'avance, et l'agent doit par conséquent les estimer. Supposons que l'agent observe au temps  $t$  la transition de l'état  $x_{t-1} = x$  vers l'état  $x_t = y$ , qu'il aie joué l'action  $a_{t-1} = a$  et reçu la récompense  $r_t$  pour cette transition. Un estimateur de  $Q_t^*(x, a)$  peut s'écrire

$$Q_t(x, a) = \frac{1}{t} \sum_{n=1}^t \left[ f(x, a, x_n) + \gamma \max_{b \in \mathcal{A}} Q_n(x_n, b) \right] \mathbb{1}_{[x_{n-1}=x, a_{n-1}=a]} \quad (5.17)$$

$$= \left(1 - \frac{1}{t}\right) Q_{t-1}(x, a) + \frac{1}{t} \left[ r_t + \gamma \max_b Q_t(y, b) \right] \quad (5.18)$$

où la dernière ligne est à la base de l'algorithme  $Q$ -learning .

L'agent acquiert ainsi une connaissance de la fonction action-valeur itérativement, au fur et à mesure de ses choix d'actions dans les différents états de l'environnement. En maintenant une table où sont stockées et mises à jour les estimations  $Q_t(x, a)$  pour chaque paire état-action, l'agent peut déduire quelles actions maximisent son retour espéré dans chaque état, et ainsi mettre en œuvre, asymptotiquement, une politique optimale. Lorsque l'agent exploite au mieux sa connaissance actuelle de la fonction action-valeur estimée, il sélectionne l'action à jouer dans l'état suivant, au temps  $t$ , selon la règle

$$a_t = \max_{a \in \mathcal{A}} Q_t(x_t, a). \quad (5.19)$$

Néanmoins, il est nécessaire d'essayer chaque action suffisamment souvent afin de garantir une estimation fiable de  $Q_t(x, a)$  pour chaque  $(x, a) \in \mathcal{X} \times \mathcal{A}$ . L'agent doit donc trouver un compromis entre l'*exploitation* de sa connaissance de l'environnement et l'*exploration* de nouvelles actions afin de ne pas rester sur une politique sous-optimale. La méthode d'exploration classique en  $Q$ -learning consiste à jouer la stratégie d'exploitation selon la règle (5.19) avec une probabilité  $\epsilon$  et d'explorer l'espace action-état avec une probabilité  $\epsilon$ , avec une action choisie selon la loi uniforme sur  $\mathcal{A}$ . Sous ces conditions, et dans un problème modélisable comme un processus décisionnel de Markov à espaces d'actions et d'états finis dénombrables, l'algorithme  $Q$ -learning basé sur la récursion (5.18) fait converger  $Q_t(x, a)$  vers l'unique fonction action-valeur optimale  $Q_t^*(x, a)$  pour tout couple  $(a, x) \in \mathcal{A} \times \mathcal{X}$  lorsque  $t \rightarrow \infty$  [Sutton and Barto, 1998; Watkins and Dayan, 1992].

---

**Algorithme 4**  $Q$ -learning

---

```

1: Paramètres :  $\epsilon, \alpha_t, \gamma$ 
2:  $\forall a \in \mathcal{A}, \forall x \in \mathcal{X} \quad Q(x, a) = 0$ 
3:  $x = X_0(\omega), \quad a = A_0(\omega)$  avec  $A_0 \sim \mathcal{U}(\mathcal{A})$ .
4: pour  $t = 1, 2, \dots$  faire
5:    $y = X_t(\omega)$ 
6:    $r = f(x, a, y)$ 
7:    $Q(x, a) = (1 - \alpha_t)Q(x, a) + \alpha_t \left[ r + \gamma \max_{b \in \mathcal{A}} Q(y, b) \right]$ 
8:    $x = y$ 
9:    $a = \begin{cases} \operatorname{argmax}_{a \in \mathcal{A}} Q(x, a) & \text{avec probabilité } 1 - \epsilon \\ A_t(\omega) \text{ avec } A_t \sim \mathcal{U}(\mathcal{A}) & \text{avec probabilité } \epsilon \end{cases}$ 
10: fin pour

```

---

L'algorithme 4 résume le principe du  $Q$ -learning . Noter qu'à la ligne 9, si plusieurs actions sont maximisatrices de  $Q(x, a)$ , alors le choix est réaliser au hasard parmi celles-ci. Le paramètre  $\alpha_t$  à la ligne 7 de l'algorithme 4 remplace le terme  $\frac{1}{t}$  de l'équation (5.18). Ce paramètre est appelé "taux d'apprentissage" et est utilisé pour pondérer différemment les nouvelles estimées des anciennes, ce qui permet la poursuite d'environnements non stationnaires. Ce paramètre peut être constant ou défini comme une suite de pondérations variables avec le temps. Le théorème donné par [Watkins and Dayan, 1992] établit la convergence des estimations  $Q(x, a)$  vers l'unique fonction action-valeur optimale d'un processus de décision markovien pour tout couple  $(x, a)$  pourvu que

1. La probabilité d'exploration  $\epsilon$  décroisse avec le temps
2. La suite  $\alpha_t$  soit telle que  $0 \leq \alpha < 1$ ,  $\sum_t \alpha_t = \infty$  et  $\sum_t \alpha_t^2 < \infty$ .

Dans de nombreux problèmes pratiques comme celui qui va être traité dans la suite de ce chapitre, le modèle de processus de décision markovien n'est pas directement applicable et l'environnement peut ne pas être stationnaire. La notion de convergence revêt alors moins d'intérêt et on préférera avoir la possibilité de poursuivre les variabilité de l'environnement. Dans ce cas, le paramètre de taux d'apprentissage est souvent réduit à une constante dont le rôle est de pondérer les estimations anciennes par rapport au nouvelles. Une plus grande importance est donnée aux anciennes estimations à mesure que  $\alpha$  tend vers 0 et vice versa.

### 5.2.3 Application à la sélection de point d'accès

L'algorithme  $Q$ -learning peut permettre aux émetteurs de choisir de manière autonome le point d'accès qui maximise leur probabilité de transmettre leurs paquets avec succès. Les émetteurs vont être considérés comme des agents indépendants ayant des actions à leur disposition pour maximiser la probabilité de transmettre leurs paquets avec succès. Dans le scénario du paragraphe 5.1, les actions des émetteurs sont identifiables à un choix de point d'accès parmi les récepteurs  $\mathcal{A} = \{1, \dots, A\}$ . À chaque récepteur est associé un code d'étalement DSSS, ainsi le choix d'un point d'accès revient à choisir un code. Tous les émetteurs connaissent les codes de tous les récepteurs. Pour chaque émetteur  $i$ , l'espace d'action est donc défini comme  $\mathcal{A}_i = \mathcal{A}$ . On fait abstraction du modèle d'états de l'environnement pour ce scénario. Il est possible que les gains des canaux représentent ces états, cependant leurs valeurs sont prises dans continuum, ce qui complique à la modélisation et la résolution du problème. De plus, les émetteurs ne sont pas supposés observer les canaux dans le problème traité ici. Notons cependant qu'il existe des méthodes d'apprentissage par renforcement permettant d'approximer les fonctions action-valeur dans le cas d'espaces action-états grands ou infini, qu'ils soient discrets ou continus [Sutton and Barto, 1998]. Cependant ces méthodes dépassent le travail réalisé durant cette thèse. Les fonctions action-valeur seront donc uniquement paramétrées par les actions des agents pour le problème traité ici. Pour la suite, on se place du point de vue d'un émetteur  $i$  en particulier.

Soit  $\mathcal{T}$  l'ensemble des temps de départ des paquets transmis par l'émetteur  $i$ . L'émetteur choisit un récepteur  $a \in \mathcal{A}$  pour le paquet émis au temps  $t \in \mathcal{T}$ . Ce paquet arrive au récepteur au temps  $t + \Delta_{i,a}$  où  $\Delta_{i,a}$  est le délai de propagation entre l'émetteur et le récepteur. Une collision se produit dès lors qu'un paquet d'un émetteur différent est reçu au même point d'accès dans un temps compris entre  $t + \Delta_{i,a} - T$  et  $t + \Delta_{i,a} + T$ . Le récepteur est supposé capable de calculer le RSIB moyen du paquet de l'émetteur  $i$  sur la base de l'équation (5.4). Au temps  $t + 2\Delta_{i,a}$ , l'émetteur reçoit un retour du récepteur défini pour le paquet émis au temps  $t$  par

$$R_t \triangleq \begin{cases} +1 & \text{if } \gamma_i(a) \geq \Gamma_i \\ -1 & \text{if } \gamma_i(a) < \Gamma_i \end{cases} \quad (5.20)$$

Ce retour peut prendre la forme d'un signal de 1 bit acquittant la bonne réception du paquet par le récepteur concerné. D'un point de vue pratique, il se peut que le récepteur ne soit pas en mesure de détecter la présence des paquets du capteur  $i$  lorsque le rapport signal à interférence plus bruit du lien de communication entre ces deux terminaux est trop faible. Dans ce cas, on peut supposer que si le capteur ne reçoit pas de signal de récompense après un certain temps prédéfini, il considère que la récompense est de  $-1$ .

L'émetteur ne connaît pas les gains des canaux de communication et les temps d'arrivés des paquets, et n'est pas supposé pouvoir les observer. De son point de vue, le retour  $R_t$  est donc une variable aléatoire de distribution inconnue. Comme dans le cas d'un processus de décision de Markov, l'agent va chercher la politique  $\pi$  qui maximise la somme escomptée de ses retours espérés, donnée au temps  $t$  pour l'action  $a$  par

$$Q_t^\pi(a) = \mathbb{E} \left[ \sum_k \gamma^k R_{t+k} \mid A_t = a \right]. \quad (5.21)$$

La fonction de valeur correspondante s'écrit

$$V_t^\pi = \mathbb{E} \left[ \sum_k \gamma^k R_{t+k} \right] \quad (5.22)$$

où l'espérance est prise par rapport aux distributions de probabilité sur l'espace d'action données par la politique  $\pi$ , et également par rapport à la distribution des retours  $R_t$ . La fonction de valeur est donc déterministe, pour un politique  $\pi$  donnée. La maximisation de cette fonction revient, selon le principe d'optimalité de Bellman, à trouver la politique optimale  $\pi^*$  telle que

$$V_t^* \triangleq \max_{a \in \mathcal{A}} Q_t^*(a) \quad (5.23)$$

$$= \max_{a \in \mathcal{A}} \mathbb{E} [R_{t+1} + \gamma V_{t+1}^* \mid A_t = a]. \quad (5.24)$$

Cette maximisation implique d'estimer la fonction action-valeur optimale  $Q_t^*(a)$  pour toute action  $a$  et à en déduire l'action maximisatrice à chaque étape du processus d'apprentissage. Ceci peut se réaliser à l'aide d'une version "sans états" du  $Q$ -learning initialement donné par l'algorithme 4.

### 5.3 Résultats numériques

Ce paragraphe donne des résultats numériques d'application de cette méthode à un scénario tiré du modèle du paragraphe 5.1. On se place dans le contexte applicatif d'un réseau de capteurs sous-marins dédié à la surveillance de l'environnement. Le réseau est déployé à grande échelle sur le fond marin. Les capteurs sont ancrés au fond de l'eau et ont pour vocation à réaliser des mesures d'une certaine grandeur physique (pression, température, etc) transmise par la suite à des bouées de surface. Dans le problème du paragraphe 5.1, les émetteurs sont les capteurs et les récepteurs (ou points d'accès) sont les bouées.

Un ensemble de  $I = 80$  émetteurs est considéré, avec  $A = 8$  récepteurs. Les communications prennent place sur une bande de  $B = 8$  kHz centrée sur  $f_c = 12$  kHz. Chaque capteur génère des paquets à un rythme de  $\lambda = 0.1$  paquet par seconde. En moyenne, 8 paquets par secondes circulent des capteurs aux bouées. Les symboles du paquet sont étalés avec un code de  $L = 128$  chips, offrant un gain de RSIB en réception de 21 dB. On approximera le temps chip à environ  $T_{chip} \approx 1/B = 0.125 \mu s$ . Le temps symbole est alors de  $L \times T_{chip} = 16 \mu s$ . La durée du paquet est fixée à  $T_p$  de sorte que 62 symboles soient transmis par paquet. Dans le cas pratique où les capteurs sont supposé mesurer et transmettre des mesures physiques, et en fonction de l'ordre de modulation, une telle longueur suffit pour contenir une ou plusieurs mesures, ainsi que des symboles supplémentaires pour l'estimation du canal, le codage correcteur d'erreurs, etc. Les terminaux sont supposés être déployés aléatoirement dans une zone de  $64 \text{ km}^2$ , avec une distance minimum séparant les émetteurs et les récepteurs de 100 m et 500 m, respectivement. La profondeur de la colonne d'eau est de 1 km. Les émetteurs sont immergés aléatoirement entre 1000 et 800 m de profondeur, tandis que les récepteurs sont entre 5 et 20 m.

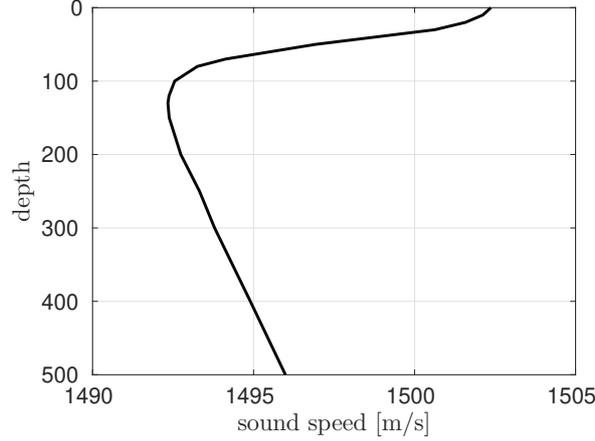


FIGURE 5.2 – Profil de célérité utilisé pour les simulations.

Les coefficients des gains des canaux  $\rho_{i,a}(t)$  sont simulés par des processus log-normaux auto-régressifs d'ordre 1 [Qarabaqi and Stojanovic, 2011b, 2013] tels que  $\forall i$  et  $\forall a$  (avec  $t$  devant être compris comme l'indice d'un temps discret dans ce qui suit) :

$$\rho_{i,a}(t) = 10^{\frac{1}{10}(g_{i,a}(t) + \bar{g}_{i,a})} \quad (5.25)$$

où

$$g_{i,a}(t) = \phi g_{i,a}(t-1) + \eta_{i,a}(t) \quad (5.26)$$

avec  $\eta_{i,a} \sim \mathcal{N}(0, \sigma_{dB}^2)$ . Ceci modélise les évanouissements lents du canal. La constante  $\phi < 1$  est calculée de manière à avoir un temps de cohérence de  $\tau_c = 60$  s pour le processus modélisant l'évanouissement lent. Cela qui signifie que, sur une durée  $\tau_c$  et conditionnellement à l'évanouissement lent, l'évanouissement rapide peut être considéré comme stationnaire au sens large de sorte que ses moyennes empiriques convergent vers ses espérances. L'étalement de puissance du processus d'évanouissement lent est fixé à  $\sigma_{dB}^2 = 10$  dB. Le processus d'évanouissement lent est centré autour d'une moyenne  $\bar{g}_{i,a}$  correspondant aux pertes de transmission renvoyées par le simulateur de tracé de rayons ASM Bellhop [Porter, 2011]. Ce simulateur trace des rayons acoustiques depuis une source vers un récepteur et en calcule les pertes de propagation subies étant donné en entrée une configuration géométrique (distances et profondeurs), une bande de fréquence, et un profil de célérité dans la colonne d'eau. Le profil de célérité utilisé ici a été enregistré en Atlantique Nord aux longitudes  $[-70^\circ, -60^\circ]$  et latitudes  $[22^\circ, 30^\circ]$ <sup>3</sup> et a été tronqué à 1 km de profondeur pour les besoins de la simulation. Ce profil de célérité est illustré à la figure 5.2. Une réponse impulsionnelle est calculée sur la base du tracé de rayons, et les pertes de propagation en sont déduites en intégrant la puissance reçue sur la bande de transmission considérée pour former le coefficient  $\bar{g}_{i,a}$ . La puissance émise est contrainte à 170 dB ref.  $\mu\text{Pa}$  et la puissance du bruit est fixée de manière à obtenir un rapport signal à bruit (RSB) de référence égal à 15 dB à une distance de 1 km sans interférences. La contrainte de RSIB pour la bonne réception d'un paquet est fixée à  $\Gamma_i = 10$  dB.

L'algorithme 4 est mis en œuvre sur chaque capteur du réseau et on s'intéresse au fonctionnement du réseau à long terme (plusieurs heures). Le taux d'apprentissage

3. voir aussi l'exemple fourni par [Dunshaw]

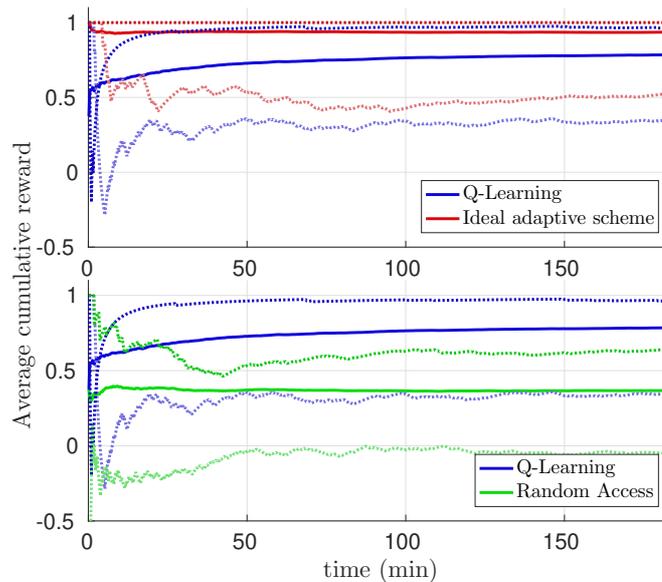


FIGURE 5.3 – Retours moyens cumulés dans le temps pour la méthode de sélection de point d'accès  $Q$ -learning et pour deux autres schémas d'accès : idéal et accès aléatoire. En traits pleins : performance moyenne sur l'ensemble des émetteurs. En pointillé : performances maximum et minimum.

est fixé à  $\alpha = 0.1$  et la probabilité d'exploration à  $\epsilon = 0.05$ . En pratique la probabilité d'exploration n'est pas nécessairement décréue avec le temps. Ceci s'explique par le fait qu'en situation réelle, il peut être bénéfique d'explorer à nouveau afin de mieux s'adapter aux non-stationnarités de l'environnement. Le paramètre  $\alpha$  a été fixé empiriquement ici. Notons que les processus d'évanouissement lents peuvent être considérés comme stationnaires sur la durée de la simulation car  $\phi < 1$  dans l'équation (5.26). Le gain d'un canal entre un émetteur  $i$  et un récepteur  $a$  fluctue donc toujours autour du même gain moyen  $\bar{g}_{i,a}$ . Les différentes simulations qui ont put être lancées avec d'autres valeurs de ce paramètre n'y ont pas montré de sensibilité au niveau des résultats finaux obtenus.

Les performances sont évaluées en termes de somme cumulées des retours. La comparaison est faite avec la méthode d'accès aléatoire (préconisée dans des réseaux similaires par [Fazel et al., 2011, 2013], en conjonction avec une méthode d'échantillonnage compressé au niveaux d'un nœud central pour la reconstruction de la mesure physique d'intérêt) et avec un schéma d'accès idéal où un génie donnerait aux émetteurs la connaissance du point d'accès avec le gain de canal le plus favorable au moment où un paquet doit être envoyé. Les résultats sont affichés sur la figure 5.3. On peut constater que le  $Q$ -learning atteint des performances assez proches du schéma idéal mais avec une connaissance de l'environnement beaucoup plus réduite - constituée seulement des retours encodés par un signal binaire. Asymptotiquement, les retours cumulés se traduisent par une probabilité d'accès réussi de 90.3%. Dans le cas du schéma idéal et de l'accès aléatoire, ces probabilités sont, respectivement, de 97% et 68.4%. Les écarts types à la performance moyennée sur l'ensemble des émetteurs sont de 8.5%, 14.6% et 13.2% respectivement pour le schéma idéal, le  $Q$ -learning, et l'accès aléatoire.

La figure 5.4 montre les fonctions action-valeur et les actions sélectionnées au cours du temps pour deux émetteurs dans le cas d'une panne simulée d'un point d'accès

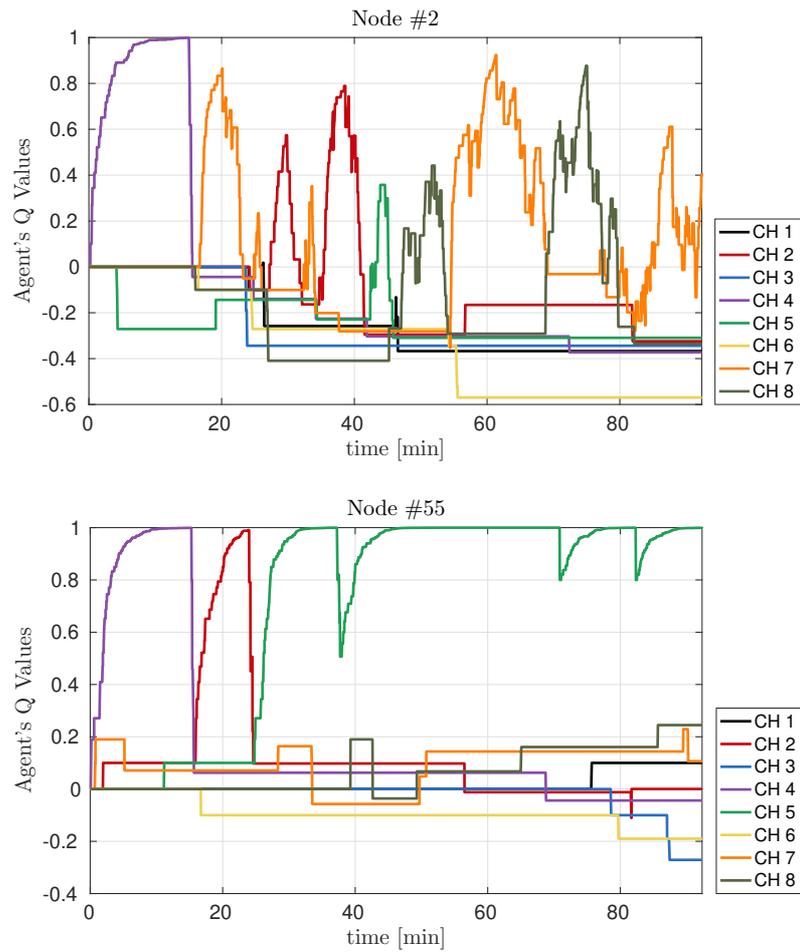


FIGURE 5.4 – Évolution des fonctions action-valeurs de deux émetteurs dans le cas de la disparition du récepteur 4 après 15 minutes. L'algorithme permet aux émetteurs d'adapter leurs stratégies d'accès.

qu'ils préféreraient initialement. Cette panne apparaît après 15 minutes d'opération. On voit nettement que les capteurs modifient leurs comportements après avoir reçu suffisamment de punitions. Le nombre de punitions nécessaires dépend de la valeur du paramètre  $\alpha$ . Le choix judicieux d'un autre récepteur dépend alors de l'exploration déjà réalisée. Le comportement de de chaque capteur peut également être assez différent. Le capteur 55 converge relativement rapidement vers le choix d'un unique récepteur tandis que le capteur 2 semble ne pas en être capable, tout en profitant de transmissions réussies durant certaines périodes. Ceci illustre la capacité d'adaptation pouvant être offerte aux réseaux ASM par les méthodes d'apprentissage par renforcement.

## 5.4 Conclusion

Les bénéfices pouvant être apportés aux réseaux sous-marins par des méthodes d'apprentissage par renforcement ont été illustrés à travers un exemple concret d'application aux réseaux de capteurs. La méthode utilisée, qui a été détaillée aux sections 5.2.2 et 5.2.1, a l'avantage d'être très générale et peut s'appliquer à d'autres scénario où les agents auraient d'autres stratégies de transmission et d'autres fonction de récompense. Les résultats de ce chapitre font écho à divers travaux très actuels sur ce même sujet [Anjangi and Chitre, 2018; Wang et al., 2018], ce qui constitue un résultat prometteur et une piste à explorer plus en détails.

L'algorithme  $Q$ -learning présente l'avantage de ne pas requérir de modèle de l'environnement pour pouvoir être mis en œuvre. Seuls des retours d'informations sur l'environnement sous la forme de récompenses/punitions attribuées en fonction des actions choisies par les agents sont nécessaires. L'inconvénient majeur de cet algorithme étant néanmoins sa limitation à des espaces d'actions relativement restreints, si l'on envisage un apprentissage *in situ*. Une des poursuites possibles de ce travail préliminaire et assez basique pourrait être d'étudier l'apport d'un pré-apprentissage piloté par des données, effectué en laboratoire, avant de mettre les systèmes en situation. Dans le scénario applicatif qui vient d'être étudié, la dépendance des récompenses obtenues par les agents vis-à-vis des actions choisies par d'autres agents n'est pas explicite. L'apprentissage multi-agents permet de prendre en compte cette dépendance dans les algorithmes d'apprentissage par renforcement. Ce cadre de travail est très proche de celui de la théorie des jeux et on pourra s'intéresser en particulier aux jeux stochastiques, qui sont une généralisation du problème d'apprentissage par renforcement présenté à la section 5.2.1 au cadre multi-agents [Hu and Wellman, 1998; Littman, 1994].

Enfin, pour le cas d'application traité ici où des capteurs doivent s'associer à des points d'accès, les jeux de congestion et de population peuvent s'avérer utiles. Ces jeux sont des jeux de forme normale où interviennent des joueurs se partageant des ressources et dont les fonctions d'utilité dépendent des ressources choisies ainsi que du nombre de joueurs choisissant la même ressource. Ce cadre a été employé dans certains travaux, tels que [Hasan et al., 2012] sur les réseaux cellulaires auto-optimisés. Il y est montré que des mobiles peuvent apprendre un schéma d'association à des stations de base qui soit un équilibre de Nash d'un jeu, où les débits des liens descendant sont individuellement maximisés.

# Conclusion générale et perspectives

## Synthèse des contributions

Le travail de recherche de cette thèse a pour vocation de fournir des méthodes d'allocation de ressources dans le canal ASM qui soient décentralisées et adaptatives. Dans un contexte où les bandes passantes sont limitées, où l'environnement est dynamique et où les systèmes de communications opèrent en dehors d'un cadre standardisé, il est important de permettre la re-configuration intelligente et automatique des paramètres de transmission des émetteurs afin d'aboutir à des partages efficaces de la ressource physique. Le problème a été présenté sous la perspective des systèmes multi-agents et les outils de la théorie des jeux et de l'apprentissage automatique ont permis de proposer des méthodes décentralisées, adaptatives, et ne nécessitant qu'une connaissance locale et imparfaite de l'environnement de la part de chaque émetteur. Celles-ci s'accordent avec la plupart des contraintes posées par l'environnement marin et qui n'existent pas forcément dans d'autres domaines d'application de la théorie des communications : absence de protocoles de communication standardisés, absence d'infrastructure, échelles de variabilités des évanouissements du canal, limitations en fréquence de transmission et en bande passante, et délais de propagation.

Les deux premiers chapitres ont présenté en détails les communications et le canal ASM puis les outils de la théorie des jeux qui permettent de modéliser les compétitions qui peuvent s'établir entre sources acoustiques dans le milieu marin. Les principales contributions de cette thèse au regard des objectifs fixés sont contenues dans les chapitres suivants.

Le troisième chapitre présente un problème de partage de spectre entre émetteurs OFDM dans le canal d'interférence ASM parallèle. La formulation du problème comme un jeu non-coopératif aboutit à une solution d'équilibre de Nash basée sur l'allocation de puissance water-filling, l'objectif individuel de chaque émetteur étant de maximiser son débit sous contrainte de puissance. Afin de tenir compte du fait que les différents couples émetteurs-récepteurs ne peuvent pas connaître les distributions de probabilité de l'interférence, le problème a été formulé comme une optimisation robuste face à l'interférence la plus préjudiciable. L'équilibre du jeu peut être appris itérativement par les émetteurs grâce à l'algorithme de dynamique de meilleures réponses calculées sur la base de la connaissance du rapport signal à interférence plus bruit. La variabilité temporelle du canal et les délais de propagation longs de plusieurs secondes contraignent les liens de retours fournissant l'information sur le canal aux émetteurs depuis leurs

récepteurs respectifs. Il a donc été choisi de fournir une information statistique afin de préserver sa cohérence lorsque celle-ci est disponible à l'émission, plusieurs secondes après avoir été estimée par la réception. L'impact des erreurs d'estimation du canal a été pris en compte de manière robuste face à une borne d'incertitude. Les résultats montrent que cette stratégie permet d'atteindre des partages de spectres plus proches de la frontière de Pareto que l'équilibre de Nash calculé avec une connaissance parfaite du gain du canal dans le cas d'une interférence suffisamment forte, sans mettre en place de coopération d'échanges d'informations entre les joueurs.

Le quatrième chapitre reprend certains résultats du chapitre 3 sur la base desquels le problème complémentaire de partage de spectre sous contrainte de débit est étudié, l'objectif étant cette fois la minimisation de la puissance de transmission. Un jeu non-coopératif sous forme normale est également formulé, cependant la recherche de son équilibre pose plusieurs difficultés. Ce problème d'optimisation multi-agents se situe dans la classe des problèmes d'équilibre de Nash généralisé, pour lesquels l'existence d'une solution n'est pas garantie. Néanmoins, lorsqu'une solution existe, celle-ci peut-être calculée à l'aide de l'algorithme de water-filling itératif de la même façon qu'au chapitre 3. Les jeux de satisfaction, dont les équilibres permettent que certains joueurs ne satisfassent pas leurs contraintes, offrent une formalisation plus flexible du problème. Plusieurs cas d'étude sont présentés avec des émetteurs en modulation DSSS et OFDM dans le canal ASM. Afin de réduire la surcharge sur le lien de retour fournissant l'information à l'émetteur sur l'état du canal, un algorithme ne nécessitant qu'une information encodée sur 1 bit est proposé. Les résultats montrent que les équilibres de satisfaction atteints offrent, en moyenne, de meilleures performances que les équilibres de Nash des jeux sous forme normale associés.

Le cinquième chapitre propose une ouverture sur les réseaux de capteurs sous-marins non centralisés et adaptatifs à l'environnement. L'intérêt s'est porté en particulier sur l'apprentissage par renforcement comme outil leur apportant de la flexibilité et des capacités d'adaptation à l'environnement. Les bénéfices pouvant être apportés par l'algorithme  $Q$ -learning sont illustrés à travers une application aux réseaux de capteurs. Les capteurs, déployés au fond de l'eau, sont dédiés à des mesures d'une quantité physique d'intérêt et doivent transmettre les données collectées à des récepteurs embarqués sur des bouées de surface. La configuration du réseau n'est pas pré-établie et on laisse les capteurs libres de choisir de manière autonome quel récepteur maximise la bonne réception des paquets de données, grâce à l'algorithme  $Q$ -learning. Cet algorithme ne nécessite pas de modèle, ce qui est intéressant dans un domaine des télécommunications où il n'existe pas de consensus sur la modélisation statistique du canal et où il est difficile de prédire à l'avance avec précision les conditions dans lesquelles les systèmes de communications vont opérer. L'adaptation des stratégies de transmission des capteurs peut se faire seulement sur la base d'un signal de retour contenant une valeur de récompense pour l'action choisie à un instant donné. Dans l'exemple traité, cette valeur est encodée sur 1 bit de manière à minimiser la surcharge du réseau. Enfin, l'algorithme  $Q$ -learning ne nécessite pas de synchroniser les éléments du réseau dans la manière dont il est mis en œuvre au chapitre 5.

## Perspectives

Les perspectives de poursuite des recherches exposées dans ce manuscrit sont les suivantes :

- Bien qu'étant couteux, et d'autant plus que le nombre de modems à mettre à l'eau peut être élevé dans le contexte de cette thèse, la validation par des essais en mer des solutions proposées leur apporterait une valeur ajoutée importante. De plus, et au meilleur de notre connaissance, aucun algorithme ne mettant en œuvre une adaptation de l'émetteur sur la base d'un lien de retour n'a été testé en mer à ce jour.
- Les travaux présentés ici se sont focalisés sur l'aspect non-coopératif des communications ASM. L'étude des mêmes problèmes dans un cadre coopératif, permettrait d'évaluer les bénéfices et les coûts d'introduire de la coopération en termes de performance gagnée contre la surcharge en échange de messages pour la mettre en place. La théorie des jeux offre également des outils pour traiter le cadre coopératif, avec des solutions d'équilibres basées sur la négociation [Nash, 1950] et des métriques comme le prix de l'anarchie permettant de comparer les équilibres coopératifs et non-coopératifs [Bacci et al., 2016].
- De manière générale, les équilibres non-coopératifs peuvent avoir le défaut d'être inefficaces en comparaison des mécanismes permettant la coopération. Une façon de forcer les joueurs à une forme de coopération peut être trouvée dans les jeux avec prix [Saraydar et al., 2002; Wang et al., 2008]. D'une certaine façon, l'allocation de puissance robuste aux erreurs d'estimation du chapitre 3 met déjà en œuvre une forme de mécanisme de prix en intégrant, dans la fonction d'utilité, une pénalité sur l'incertitude portée par les estimations du RSIB. D'autres manières de faire peuvent être explorées, par exemple, en pénalisant la puissance de transmission ou la bande passante occupée.
- L'apport des autres types de jeux que ceux qui sont utilisés ici devrait également être évalué. Selon les applications, les jeux répétés, stochastiques ou bayésiens peuvent apporter des raffinements dans les concepts d'équilibres qui peuvent être intéressants. Les jeux stochastiques et répétés sont, en particulier, très liés aux problèmes d'apprentissage par renforcement et peuvent permettre de mieux comprendre certains aspects de l'apprentissage automatique dans le contexte multi-agents.
- Enfin, les méthodes d'apprentissage de « bonnes » stratégies de partage de ressources proposées par ce manuscrit sont supposées être mise en place *in situ*, au sens où les systèmes découvrent au fur et à mesure leur environnement. Une amélioration possible peut consister à entraîner les systèmes en amont de leur mise en opération, de manière à leur fournir des *a priori* sur certains types de canaux ou de configurations pouvant être rencontrées afin, d'une part, de converger plus rapidement vers des bonnes stratégies lorsqu'elles ont été apprises par l'entraînement, et d'autre part laisser à l'apprentissage *in situ* la seule tâche d'apprendre d'éventuelles nouveautés.



# Publications

## Articles de journaux

1. A. Pottier, F-X. Socheleau, C. Laot, "Quality-of-Service Satisfaction Games for Noncooperative Underwater Acoustic Communications", *IEEE Access, Special Issue : Underwater Wireless Communications and Networking*, Vol. 6, Apr. 2018.
2. A. Pottier, F-X. Socheleau, C. Laot, "Robust Noncooperative Spectrum Sharing Game in Underwater Acoustic Interference Channels", *IEEE Journal of Oceanic Engineering*, Vol. 42, Num. 4, pp 1019-1034, October 2017.

## Articles de conférences

1. A. Pottier, P. D. Mitchell, F-X. Socheleau, C. Laot, "Q-Learning Based Adaptive Channel Selection for Underwater Sensor Networks", accepted at *Underwater Communications and Networking Conference (UComms) 2018, Lerici*, August 2018.
2. A. Pottier, F-X. Socheleau, C. Laot, " Power-Efficient Spectrum Sharing for Noncooperative Underwater Acoustic Communication Systems", *IEEE OCEANS 2016, Monterey, USA*, September 2016.
3. A. Pottier, F-X. Socheleau, C. Laot, " Adaptive Power Allocation for Noncooperative OFDM Systems in UWA Interference Channels", *Underwater Communications and Networking Conference (UComms) 2016, Lerici, Italy*, August 2016.
4. A. Pottier, F-X. Socheleau, C. Laot, "Distributed Power Allocation Strategy in Shallow Water Acoustic Interference Channels", *IEEE Signal Processing Advances for Wireless Comms. (SPAWC) 2016, Edinburgh, UK*, July 2016.
5. F-X. Socheleau, A. Pottier, C. Laot, "Stochastic Replay of SIMO Underwater Acoustic Communication Channels", *IEEE OCEANS 2015, Washington, USA*, October 2015.



# Annexe A

## Calcul des fonctions d'utilité pour le partage de spectre robuste

### A.1 Minimisation contre l'interférence la plus préjudiciable

Considérant que tous les émetteurs distribuent leurs symboles selon une loi normale, l'information mutuelle moyenne du joueur  $i$  est donnée par

$$R_i(\mathbf{p}_i, \mathbf{p}_{-i}) = \alpha \sum_{n=1}^N \mathbb{E} [\log (1 + \text{SINR}_{i,n})] \quad (\text{A.1})$$

$$= \alpha \sum_{n=1}^N \mathbb{E} \left[ \log \left( 1 + \frac{|H_{ii,n}|^2 p_{i,n}}{\sigma_{i,n}^2 + \sum_{j \neq i} |H_{ji,n}|^2 p_{j,n}} \right) \right] \quad (\text{A.2})$$

où l'espérance dans (A.1) est prise selon la distribution conjointe des  $H_{ji}$ ,  $i, j \in \mathcal{I}$ . On rappelle que  $p_{j,n} = \mathbb{E} [|X_{j,n}|^2]$ . Posons  $f_{\mathbf{H}_{ji}}$  la densité de probabilité de  $\mathbf{H}_{ji} = (H_{ji,n})_{n \in \mathcal{N}}$ ,  $\forall j \neq i$ . La fonction d'utilité proposée est

$$u_i(\mathbf{p}_i, \mathbf{p}_{-i}) = \min_{f_{\mathbf{H}_{ji}} \in \mathcal{L}^2(\mathbb{R})} R_i(\mathbf{p}_i, \mathbf{p}_{-i}). \quad (\text{A.3})$$

Pour résoudre ce problème, on peut, de manière équivalente, minimiser sur chaque sous-canal  $n \in \{1, \dots, N\}$ .

$$\mathbb{E} [\log (1 + \text{SINR}_{i,n})] = \mathbb{E}_{H_{ii,n}} \left[ \mathbb{E}_{H_{ji,n}} \left[ \log \left( 1 + \frac{|H_{ii,n}|^2 p_{i,n}}{\sigma_{i,n}^2 + \sum_{j \neq i} |H_{ji,n}|^2 p_{j,n}} \right) \mid H_{ii,n} \right] \right] \quad (\text{A.4})$$

$$\geq \mathbb{E}_{H_{ii,n}} \left[ \log \left( 1 + \frac{|H_{ii,n}|^2 p_{i,n}}{\sigma_{i,n}^2(n) + \sum_{j \neq i} \mathbb{E} [|H_{ji,n}|^2] p_{j,n}} \right) \right]. \quad (\text{A.5})$$

L'équation (A.4) provient de la loi de l'espérance itérée, l'inégalité (A.5) vient ensuite de l'inégalité de Jensen puisque la fonction sous l'espérance conditionnelle est convexe

en  $H_{ii,n}$ . La solution du problème (A.3) est donc donnée par le membre de droite de l'inégalité (A.5), ce minimum étant atteint pour une interférence

$$Z_{i,n}(\mathbf{p}_{-i}) = \sum_{j \neq i} H_{ji,n} X_{j,n} + \sigma_{i,n}^2 \quad (\text{A.6})$$

distribuée selon une loi normale.

Le jeu pourrait être défini avec (A.5) comme fonction d'utilité. Cependant, la solution de water-filling donnée par (3.17) n'aurait pas de solution analytique et les conditions d'unicité de l'équilibre de Nash et de convergence de l'algorithme de meilleures réponses seraient plus difficiles à exprimer. Ceci justifie le besoin que l'on a d'exprimer la fonction d'utilité comme au membre de droite de l'équation (3.9). L'utilité de l'équation (A.5) est tout d'abord ré-écrite comme

$$u_i(\mathbf{p}_i, \mathbf{p}_{-i}) = \alpha \sum_n \mathbb{E} \left[ \log \left( 1 + \frac{|H_{ii,n}|^2 p_{i,n}}{\zeta_{i,n}} \right) \right] \quad (\text{A.7})$$

$$= \alpha \sum_n \mathbb{E} \left[ \log \left( 1 + \frac{e^{\log |H_{ii,n}|^2} p_{i,n}}{\zeta_{i,n}} \right) \right] \quad (\text{A.8})$$

où

$$\zeta_{i,n} = \sigma_{i,n}^2(n) + \sum_{j \neq i} \mathbb{E} [|H_{ji,n}|^2] p_{j,n}. \quad (\text{A.9})$$

En notant que  $u_i(\mathbf{p}_i, \mathbf{p}_{-i})$  est alors convexe en  $\log |H_{ii,n}|^2$ , on a l'inégalité suivante

$$u_i(\mathbf{p}_i, \mathbf{p}_{-i}) \geq \alpha \sum_n \log \left( 1 + \frac{e^{\mathbb{E}[\log |H_{ii,n}|^2]} p_{i,n}}{\zeta_{i,n}} \right), \quad (\text{A.10})$$

qui termine la preuve.

On peut montrer numériquement que dans certains canaux à évanouissements de Rice ou sur les données réelles de la section 3.4.4 que (A.10) est une borne serrée de (A.7), ce qui justifie l'approximation (3.9).

## A.2 Développement de l'approximation (A.10) sous l'hypothèse d'évanouissements de Rice

En choisissant la distribution de Rice pour modèle des évanouissements du canal ASM, on a  $H_{ii,n} \sim \mathcal{CN}(\mu_{i,n}, \eta_{i,n}^2)$ ,  $n = 1, \dots, N$ , et les  $|H_{ii,n}|^2$  suivent une loi  $\mathcal{X}^2$  non centrée. Dans ce cas, l'inégalité (A.10) peut-être réécrite au moyen de l'espérance du logarithme d'une variable  $\mathcal{X}^2$  non centrée, dont on peut trouver l'expression dans [Lapidoth and Moser, 2003, équation (209)]. Soit  $H \sim \mathcal{CN}(0, 1)$ ,  $\mu \in \mathbb{C}$  et  $\eta \in \mathbb{R}_+ \setminus \{0\}$ . Alors,

$$\mathbb{E} [\log |\mu + \eta H|^2] = \log \eta^2 + \log \frac{|\mu|^2}{\eta^2} - \text{Ei} \left( -\frac{|\mu|^2}{\eta^2} \right) \quad (\text{A.11})$$

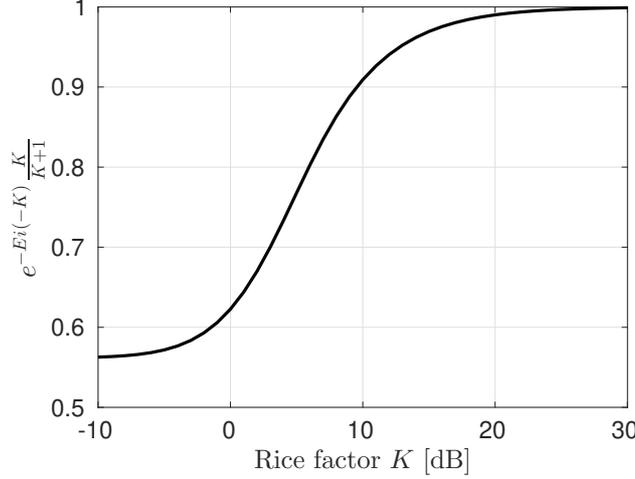


FIGURE A.1 – Coefficient  $g_{i,n}$  en fonction du facteur de Rice  $K_{i,n}$ .

où  $\text{Ei}(-x)$  est la fonction exponentielle intégrale définie pour tout  $x > 0$  par

$$\text{Ei}(-x) \triangleq - \int_x^{+\infty} \frac{e^{-t}}{t} dt. \quad (\text{A.12})$$

Il est possible de réécrire  $H_{ii,n} = \mu_{i,n} + \eta_{i,n}H$  avec  $H \sim \mathcal{CN}(0, 1)$ , et on définit le facteur de Rice  $K_{i,n}$  et le gain du canal  $H_{ii,n}$  par

$$K_{i,n} \triangleq \frac{|\mu_{i,n}|^2}{\eta_i} \quad (\text{A.13})$$

$$\mathbb{E} [|H_{ii,n}|^2] = |\mu_{i,n}|^2 + \eta_{i,n}^2. \quad (\text{A.14})$$

Ces deux définitions utilisées avec l'équation (A.11) amènent à l'expression suivante pour le logarithme espéré de  $|H_{ii,n}|^2$  :

$$\mathbb{E} [\log |H_{ii,n}|^2] = \log \eta_{i,n}^2 + \log K_{i,n} - \text{Ei}(-K_{i,n}) \quad (\text{A.15})$$

$$= \log \left( \frac{K_{i,n}}{K_{i,n} + 1} \mathbb{E} [|H_{ii,n}|^2] \right) - \text{Ei}(-K_{i,n}). \quad (\text{A.16})$$

L'inégalité (A.10) devient alors

$$u_i(\mathbf{p}_i, \mathbf{p}_{-i}) \geq \alpha \sum_n \log \left( 1 + \frac{g_{i,n} \mathbb{E} [|H_{ii,n}|^2] p_{i,n}}{\zeta_{i,n}} \right), \quad (\text{A.17})$$

où

$$g_{i,n} = \frac{K_{i,n}}{K_{i,n} + 1} e^{-\text{Ei}(-K_{i,n})}. \quad (\text{A.18})$$

Cette fonction est illustrée à la figure A.1 et tend vers 1 pour des facteurs de Rice suffisamment hauts. On peut remarquer que chaque terme de la somme dans l'expression (A.7) de  $u_i(\mathbf{p}_i, \mathbf{p}_{-i})$  peut être encadré de la façon suivante :

$$\log \left( 1 + g_{i,n} \frac{\mathbb{E} [|H_{ii,n}|^2] p_{i,n}}{\zeta_{i,n}} \right) \leq \mathbb{E} \left[ \log \left( 1 + \frac{|H_{ii,n}|^2 p_{i,n}}{\zeta_{i,n}} \right) \right] \leq \log \left( 1 + \frac{\mathbb{E} [|H_{ii,n}|^2] p_{i,n}}{\zeta_{i,n}} \right). \quad (\text{A.19})$$

Ceci montre que l'approximation (3.9), dans la forme donnée par l'équation (A.17), est meilleure à mesure que le facteur de Rice devient grand, les trois inégalités devenant des égalités lorsque  $g_{i,n} = 1$  - ce qui se vérifie asymptotiquement lorsque  $K_{i,n} \rightarrow +\infty$ . En termes physiques, cela signifie que l'approximation est valide dans des canaux de Rice dont la composante spéculaire a une puissance forte devant la composante diffusée. Des sondages en mer de canaux ASM petits fonds font état de facteurs de Rice élevés - plusieurs dizaines de décibels - pour les trajets les plus énergétiques [Socheleau et al., 2011].

# Annexe B

## Preuve du Théorème 3.2.1 : Conditions d'unicité de l'équilibre de Nash dans le jeu de partage de spectre

En général, seules des conditions suffisantes d'unicité des équilibres de Nash peuvent être données dans les jeux dont les utilités sont des fonctions concaves et continues sur des espaces d'actions compacts et convexes. Elles sont obtenues lorsque les fonctions d'utilité sont *diagonalement strictement concaves*, ce qui implique un équilibre unique. Cette notion de concavité et le théorème d'unicité qui en découle ont été introduits pour la première fois par [Rosen, 1965].

**Théorème B.0.1** ([Rosen, 1965]). *Soit un jeu  $\mathcal{G} = (\mathcal{I}, \mathcal{A}_1 \times \dots \times \mathcal{A}_I, u)$  dont les espaces d'actions  $\mathcal{A}_i \in \mathbb{R}^{n_i}$  sont compacts et convexes les fonctions d'utilités sont continues en  $\mathbf{a} \in \mathcal{A} = \mathcal{A}_1 \times \dots \times \mathcal{A}_I$  et concaves en  $\mathbf{a}_i \in \mathcal{A}_i, \forall i \in \mathcal{I}$ . Si  $\forall \mathbf{a}, \mathbf{a}' \in \mathcal{A}, \mathbf{a} \neq \mathbf{a}'$ , il existe un joueur  $i$  tel que la propriété suivante de concavité stricte et diagonale est vraie*

$$(\mathbf{a}'_i - \mathbf{a}_i)^T \nabla_{\mathbf{a}_i} u_i(\mathbf{a}) + (\mathbf{a}_i - \mathbf{a}'_i)^T \nabla_{\mathbf{a}_i} u_i(\mathbf{a}') > 0, \quad (\text{B.1})$$

alors  $\mathcal{G}$  possède un unique équilibre de Nash.

*Démonstration.* Voir [Rosen, 1965, Theorem 2]. □

Dans le jeu de partage de spectre du chapitre 3, l'idée principale amenant à des conditions pour un équilibre de Nash unique est de procéder par contradiction, en supposant l'existence de deux équilibres différents  $\mathbf{p}$  et  $\mathbf{p}'$ , et en déduire les conditions sur les statistiques des canaux  $\mathbf{H}_{ji}$  qui satisfont le théorème B.0.1.

Tout d'abord,  $u_i(\mathbf{p}_i, \mathbf{p}_{-i})$  est réécrite

$$u_i(\mathbf{p}_i, \mathbf{p}_{-i}) \propto \sum_{n=1}^N \log \left( 1 + \frac{G_{ii,n} p_{i,n}}{\sigma_{i,n}^2 + \sum_{j \neq i} G_{ji,n} p_{j,n}} \right) \quad (\text{B.2})$$

où

$$G_{ii,n} \triangleq \begin{cases} e^{\mathbb{E}[\log |H_{ii,n}|^2]} & \text{si } i = j, \\ \mathbb{E}[|H_{ji,n}|^2] & \text{sinon} \end{cases} \quad (\text{B.3})$$

On a  $\forall n \in \{1, \dots, N\}$ ,

$$[\nabla_{\mathbf{p}_i} u_i(\mathbf{p})]_n = \frac{G_{ii,n}}{\sigma_{i,n}^2 + \sum_{j=1}^I G_{ji,n} p_{j,n}}. \quad (\text{B.4})$$

Soit  $\mathbf{p}, \mathbf{p}' \in \mathcal{P} = \mathcal{P}_1 \times \dots \times \mathcal{P}_I$  et supposons  $\mathbf{p} \neq \mathbf{p}'$ . Les conditions de l'équation (B.1) ci-dessus conduisent à ce que  $\mathbf{p}$  et  $\mathbf{p}'$  ne peuvent pas être des équilibres de Nash distincts si  $\exists i \in \mathcal{I}$  tel que

$$\sum_{n=1}^N (p'_{i,n} - p_{i,n}) \frac{G_{ii,n}}{\xi_{i,n}} \times \sum_{j=1}^I G_{ji,n} (p'_{j,n} - p_{j,n}) > 0 \quad (\text{B.5})$$

où

$$\xi_{i,n} = \left( \sigma_{i,n}^2 + \sum_{j=1}^I G_{ji,n} p_{j,n} \right) \times \left( \sigma_{i,n}^2 + \sum_{j=1}^I G_{ji,n} p'_{j,n} \right) > 0. \quad (\text{B.6})$$

Il est possible de montrer [Scutari et al., 2008a, Appendix B] qu'une condition suffisante pour que (B.5) soit vérifiée est  $\exists i \in \mathcal{M}, \forall n \in \{1, \dots, N\}$

$$|p'_{i,n} - p_{i,n}| + \text{sign}(p'_{i,n} - p_{i,n}) \times \sum_{j \neq i} M_{ji,n} (p'_{j,n} - p_{j,n}) > 0 \quad (\text{B.7})$$

avec  $M_{ji}(n)$  définie pour  $j \neq i$  par

$$M_{ji,n} \triangleq \frac{G_{ji,n}}{G_{ii,n}}, \quad \forall n \in \{1, \dots, N\}. \quad (\text{B.8})$$

À partir de ce point, les mêmes étapes que la démonstration de [Scutari et al., 2008a, Appendix B] peuvent être suivies pour aboutir au théorème 1.

# Annexe C

## Preuve du théorème 3.2.2 : Conditions de convergence de l'algorithme de water-filling itératif

On va tout d'abord énoncer un théorème général de convergence des algorithmes asynchrones donné par [Bertsekas and Tsitsiklis, 1997], où on reprend les notations du chapitre 2 pour la théorie des jeux. L'idée est ensuite de montrer que lorsque les meilleures réponses des joueurs sont des applications contractantes au sens d'une certaine norme, le théorème de convergence s'applique. Il suffira ensuite de prouver que les meilleures réponses « water-filling » sont elles mêmes des applications contractantes pour garantir la convergence de l'algorithme de water-filling itératif (dans sa version asynchrone ou synchrone, la deuxième version étant un cas particulier de la première). Cette procédure est sensiblement la même que celle suivie dans [Scutari et al., 2008c].

### C.1 Convergence de la dynamique de meilleures réponse asynchrone

Rappelons pour commencer la dynamique de meilleures réponses asynchrones. Soient le jeu  $\mathcal{G} = (\mathcal{I}, \mathcal{A}, u)$  et la correspondance  $BR = (BR_i)_{i \in \mathcal{I}}$  où les  $BR_i$  sont les correspondances de meilleures réponses de chaque joueur. Soient  $\mathcal{T} = 1, 2, \dots$  les itérations de l'algorithme,  $\mathcal{T}_i \subseteq \mathcal{T}$  l'ensemble des itérations où le joueur  $i$  joue, et  $\tau_i(t)$  le dernier temps avant  $t$  où  $i$  a obtenu une connaissance de la stratégie  $a_{-i}$ . La dynamique s'écrit

$$a_i(t+1) = BR_i(a_{-i}(\tau_i(t))) \quad \forall t \in \mathcal{T}_i. \quad (\text{C.1})$$

Pour tout  $t \in \mathcal{T}$  on a donc

$$a(t+1) = BR(a(t)). \quad (\text{C.2})$$

**Théorème C.1.1** (Convergence asynchrone, [Bertsekas and Tsitsiklis, 1997]). *Si il existe une suite de sous-ensembles non-vides*

$$\dots \subset \mathcal{A}(t+1) \subset \mathcal{A}(t) \subset \dots \subset \mathcal{A}. \quad (\text{C.3})$$

tels que

1. (Condition de convergence synchrone). On a

$$BR(a) \in \mathcal{A}(t+1), \quad \forall k \text{ et } \forall a \in \mathcal{A}(t). \quad (\text{C.4})$$

De plus, si  $\{a(t)\}$  est une séquence telle que  $a(t) \in \mathcal{A}(t)$  pour tout  $t$ , alors elle converge vers un point-fixe de  $BR$ .

2. Pour tout  $t$ , il existe  $\mathcal{A}_i(t) \subset \mathcal{A}_i$  tels que

$$\mathcal{A}(t) = \mathcal{A}_1(t) \times \cdots \times \mathcal{A}_I(t), \quad (\text{C.5})$$

alors l'algorithme de meilleures réponses asynchrone de la récurrence (2.23) converge vers un point fixe qui est un équilibre de Nash en stratégies pure.

On pose  $\mathcal{A} \subseteq \mathbb{R}^n$  avec  $n = n_1 + \cdots + n_I$ , et  $\mathcal{A} = \mathcal{A}_1 \times \cdots \times \mathcal{A}_I$ . Le théorème précédent s'applique lorsque  $BR$  est contractante au sens de la norme maximum pondérée que l'on définit,  $\forall \mathbf{a} \in \mathcal{A}$ , par

$$\|\mathbf{a}\|_{\infty, \mathbf{w}} = \max_{i \in \mathcal{I}} \frac{\|\mathbf{a}_i\|_2}{w_i} \quad (\text{C.6})$$

où  $\mathbf{w} = (w_1, \dots, w_I)$  est un vecteur de poids positifs. La correspondance  $BR$  est alors dite *contractante par blocs*.

**Définition C.1.1** (Application contractante par blocs). Soit  $f : \mathcal{A} \rightarrow \mathcal{A}$ , où  $\mathcal{A} \subseteq \mathbb{R}^n$  et  $\mathcal{A} = \mathcal{A}_1 \times \cdots \times \mathcal{A}_I$ , avec  $\mathcal{A}_i \subseteq \mathbb{R}^{n_i}$  et  $n = n_1 + \cdots + n_I$ . Chaque élément de  $\mathcal{A}$  peut s'écrire  $\mathbf{a} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_I)$  et pour tout  $\mathbf{a} \in \mathcal{A}$  on a  $f(\mathbf{a}) = (f_1(\mathbf{a}), f_2(\mathbf{a}), \dots, f_I(\mathbf{a}))$  où chaque  $f_i : \mathcal{A} \rightarrow \mathcal{A}_i$  et le  $i^{\text{ème}}$  bloc de  $f(\mathbf{a})$ . L'application  $f$  est contractante par blocs si elle vérifie  $\forall \mathbf{a}, \mathbf{a}' \in \mathcal{A}$

$$\|f(\mathbf{a}) - f(\mathbf{a}')\|_{\infty, \mathbf{w}} \leq \beta \|\mathbf{a} - \mathbf{a}'\|_{\infty, \mathbf{w}} \quad (\text{C.7})$$

où  $\beta \in [0, 1[$  est le module de la contraction.

Le théorème suivant est une adaptation du théorème du point fixe de Banach au cas particulier des applications contractantes par blocs, qui est celui qui va nous intéresser pour la convergence de l'algorithme de water-filling du chapitre 3.

**Théorème C.1.2** (Point fixe pour les applications contractantes par blocs). Soient  $\mathcal{A} = \mathcal{A}_1 \times \cdots \times \mathcal{A}_I$ ,  $\mathcal{A} \subseteq \mathbb{R}^n$  où chaque  $\mathcal{A}_i \subseteq \mathbb{R}^{n_i}$  et une application  $f : \mathcal{A} \rightarrow \mathcal{A}$  contractante par blocs de module  $\beta$ . Cette application possède un unique point fixe  $\mathbf{a}^* = f(\mathbf{a}^*)$ , et toute suite d'éléments de  $\mathcal{A}$  issue de la récurrence

$$a(t+1) = f(a(t)) \quad (\text{C.8})$$

vérifie la majoration

$$\|a(t) - \mathbf{a}^*\|_{\infty, \mathbf{w}} \leq \beta^t \|a(0) - \mathbf{a}^*\|_{\infty, \mathbf{w}} \quad (\text{C.9})$$

et converge par conséquent vers  $\mathbf{a}^*$ .

Soient les ensembles définis pour tout  $t$  par

$$\mathcal{A}(t) = \left\{ a \in \mathcal{A} \mid \|a - a^*\|_{\infty, \mathbf{w}} \leq \beta^t \|a(0) - a^*\|_{\infty, \mathbf{w}} \right\} \quad (\text{C.10})$$

L'ensemble  $\mathcal{A}(t+1)$  est évidemment inclu strictement dans  $\mathcal{A}(t)$  et on vérifie donc l'hypothèse de départ du théorème de convergence asynchrone :

$$\dots \subset \mathcal{A}(t+1) \subset \mathcal{A}(t) \subset \dots \subset \mathcal{A}.$$

La condition 1 du théorème est vérifiée car  $BR$  est contractante. La récurrence définie par l'équation (C.2) converge donc bien vers un point fixe d'une manière telle que stipulée par la condition 1. De plus, on vérifie également la condition 2 car pour tout  $t$  on peut décomposer  $\mathcal{A}(t)$  comme  $\mathcal{A}(t) = \mathcal{A}_1(t) \times \dots \times \mathcal{A}_I(t)$  où

$$\mathcal{A}_i(t) = \left\{ a_i \in \mathcal{A}_i \mid \frac{\|a_i - a_i^*\|_2}{w_i} \leq \beta^t \|a(0) - a^*\|_{\infty, \mathbf{w}} \right\}. \quad (\text{C.11})$$

Ainsi, si la correspondance de meilleures réponses définie par  $BR$  est bien une contraction vis-à-vis de la norm maximum pondérée, les dynamiques de meilleures réponses asynchrones convergent géométriquement vers un l'équilibre de Nash en stratégies pures correspondant à l'unique point fixe de  $BR$ .

## C.2 Convergence du water-filling itératif

On reprend ici les notations du chapitre 3. L'opérateur de water-filling peut être vu comme une correspondance de  $\mathcal{P} = \mathcal{P}_1 \times \dots \times \mathcal{P}_M$  dans lui-même. Soit  $\mathbf{p} = [\mathbf{p}_1^T, \dots, \mathbf{p}_I^T]^T \in \mathcal{P}$  la concaténation de tous les vecteurs d'allocation de puissance  $\mathbf{p}_i \in \mathcal{P}_i$ . On peut définir la correspondance de water-filling par

$$WF[\mathbf{p}] = (WF_i[\mathbf{p}_{-i}])_{i=1}^I \quad (\text{C.12})$$

où  $\forall n \in \{1, \dots, N\}$

$$(WF_i[\mathbf{p}_{-i}])_n = \left[ \lambda_i - \frac{\sigma_{i,n}^2 + \sum_{j \neq i} \mathbb{E}[|h_{ji,n}|^2 p_{j,n}]}{e^{\mathbb{E}[\log |h_{ii,n}|^2]}} \right]^+. \quad (\text{C.13})$$

Un équilibre de Nash  $\mathbf{p}^*$  du jeu de partage de spectre  $\mathcal{G}$  défini par (3.5) est un point fixe de la correspondance de meilleures réponses (C.12), c'est-à-dire

$$\mathbf{p}^* = WF[\mathbf{p}^*]. \quad (\text{C.14})$$

L'algorithme 2 donne une version de la dynamique de meilleures réponses basée sur la correspondance définie par (C.12). Cet algorithme converge sous les mêmes conditions que celle données par le théorème C.1.1. On peut réécrire (C.13) comme

$$WF_i(\mathbf{p}_{-i}) = \left[ -\sigma_i - \sum_{j \neq i} \mathbf{H}_{ji} \mathbf{p}_j \right]_{\mathcal{P}_i} \quad (\text{C.15})$$

où

$$\boldsymbol{\sigma}_i \triangleq \left[ \frac{\sigma_{i,1}^2}{e^{\mathbb{E}[\log |h_{ii,1}|^2]}}, \dots, \frac{\sigma_{i,N}^2}{e^{\mathbb{E}[\log |h_{ii,N}|^2]}} \right]^T, \quad (\text{C.16})$$

$$\mathbf{H}_{ji} \triangleq \mathbf{D} \left( \left[ \frac{\mathbb{E}[|h_{ji,1}|^2]}{e^{\mathbb{E}[\log |h_{ii,1}|^2]}}, \dots, \frac{\mathbb{E}[|h_{ji,N}|^2]}{e^{\mathbb{E}[\log |h_{ii,N}|^2]}} \right]^T \right), \quad (\text{C.17})$$

et  $[\cdot]_{\mathcal{P}_i}$  désigne la projection euclidienne sur  $\mathcal{P}_i$ . On peut ainsi suivre les mêmes étapes que dans les preuves données par [Scutari et al., 2008b, Appendices B et C] et réutiliser leur interprétation de l'opérateur de water-filling comme une projection euclidienne sur un espace convexe pour trouver les conditions selon lesquelles il est bien une contraction vis-à-vis de la norme maximum pondérée, ce qui implique directement la convergence de l'algorithme 2 selon les mêmes arguments qu'à la section précédente. Ces conditions s'identifient aux conditions suffisantes du théorème 3.2.2.

# Annexe D

## Convergence des algorithmes 1 et 3 dans les jeux de satisfaction DSSS et OFDM

### D.1 Convergence pour les jeux de satisfaction en DSSS

Il suffit de vérifier que les conditions de la proposition 4.2.1 sont satisfaites. Les conditions 1) et 3) sont vérifiées par hypothèse : chaque joueur aura une opportunité de jouer et la définition des espaces d'actions s'accorde avec la condition 3). Pour prouver que la condition 2) est vérifiée, posons tout d'abord  $\phi_i : \mathcal{A}_{-i} \rightarrow \mathcal{A}_i$  telle que

$$\phi_i = \inf_{a_i \in \mathcal{A}_i} \{a_i \mid u_i(a_i, \mathbf{x}_{-i}) \geq \Gamma_i\} \quad (\text{D.1})$$

avec  $u_i(a_i, \mathbf{x}_{-i}) = \mathbb{E} [\|\mathbf{H}_{ii}\|_F^2] P_i^{(a_i)} N_i^{(a_i)} / \mathbb{E} [\|\mathbf{z}_i(\mathbf{x}_{-i})\|_2^2]$ . La correspondance de satisfaction  $\varphi_i$  satisfait donc  $\forall i$

$$\varphi_i(\mathbf{x}_{-i}) = \{a_i \in \mathcal{A}_i \mid \phi_i(\mathbf{x}_{-i}) \preceq a_i\}. \quad (\text{D.2})$$

Le vecteur  $\mathbf{z}_i(\mathbf{x}_{-i})$  représente l'interférence plus le bruit perçus au récepteur  $i$ . On peut ordonner l'ensemble  $\mathcal{A}_{-i}$  des stratégies des adversaires du joueur  $i$  par rapport à la puissance de l'interférence plus bruit perçue, c'est à dire

$$\forall \mathbf{x}_{-i}, \mathbf{x}'_{-i} \in \mathcal{A}_{-i}, \quad \mathbf{x}'_{-i} > \mathbf{x}_{-i} \Leftrightarrow \mathbb{E} [\|\mathbf{z}_i(\mathbf{x}'_{-i})\|_2^2] > \mathbb{E} [\|\mathbf{z}_i(\mathbf{x}_{-i})\|_2^2]. \quad (\text{D.3})$$

Pour deux  $\mathbf{x}'_{-i}$  et  $\mathbf{x}_{-i}$  tels que  $\mathbf{x}'_{-i} > \mathbf{x}_{-i}$ , on a donc  $u_i(a_i, \mathbf{x}_{-i}) > u_i(a_i, \mathbf{x}'_{-i})$  : le RSIB du joueur  $i$  diminue si la puissance de l'interférence générée par le bruit et les interférences des autres joueurs augmente. Ainsi, il est nécessaire pour ce joueur d'utiliser une puissance de transmission ou un gain d'étalement plus grands afin de satisfaire sa propre contrainte de RSIB. Ceci étant vérifié pour tous les joueurs du jeu en modulation DSSS. On conclut par conséquent que

$$\mathbf{x}'_{-i} > \mathbf{x}_{-i} \Rightarrow \phi_i(\mathbf{x}'_{-i}) \geq \phi_i(\mathbf{x}_{-i}) \quad (\text{D.4})$$

ce qui termine la preuve et montre que les algorithmes 1 et 3 convergent vers un équilibre de satisfaction dans le jeu en modulation DSSS.

## D.2 Convergence pour les jeux de satisfaction en OFDM

### D.2.1 Water-filling et retour de l'état du canal

En vertu des arguments donnés au paragraphe 4.2.5.1, les stratégies des joueurs sont les puissances de transmissions possibles  $P_i$  appartenant à

$$\mathcal{A}_i = [P_i^{min}, P_i^{max}], \quad (D.5)$$

et les seules allocations de puissance rationnelles sont celles données par la solution de water-filling au problème (4.32) où la contrainte est donnée par une puissance de transmission choisie dans  $\mathcal{A}_i$ . Ces solutions sont de la forme  $\mathbf{p}_i^* = P_i \times \mathbf{q}_i$  où  $\mathbf{q}_i \in \mathbb{R}_+^{N_i}$  et  $\|\mathbf{q}_i\|_1 = 1$ . Les correspondances de satisfaction des joueurs

$$\varphi_i(\mathbf{p}_{-i}) = \left\{ P_i \in \mathcal{A}_i : \alpha_i \sum_{n=1}^{N_i} \log(1 + P_i \gamma_{i,n}(\mathbf{p}_{-i}) q_{i,n}) \geq \Gamma_i \right\}, \quad (D.6)$$

où  $\gamma_{i,n}(\mathbf{p}_{-i})$  est défini à l'équation (3.10).

Selon la proposition 4.2.1, si  $\varphi_i$  peut s'exprimer par

$$\varphi_i(\mathbf{p}_{-i}) = \{a_i \in \mathcal{A}_i \mid \phi_i(\mathbf{p}_{-i}) \preceq a_i\} \quad (D.7)$$

avec  $\phi_i$  préservant l'ordre, alors l'algorithme 1 converge vers un équilibre de satisfaction en stratégies pures. Trouver une telle fonction ayant la propriété de préserver l'ordre pour toutes les configurations du jeu n'est pas possible. Supposons que nous ordonnions les stratégies  $\mathbf{p}_{-i} \in \mathcal{A}_{-i}$  par rapport à leur norme, c'est à dire par rapport à la puissance de transmission. Il est possible d'observer des cas où lorsqu'un joueur augmente sa puissance, l'interférence générée sur les autres joueurs oblige ceux-ci à répartir leur puissance sur les canaux où le joueur à forte puissance de transmission est absence. Ce faisant, la puissance de l'interférence perçue sur ces canaux est moindre est offre aux joueurs l'opportunité de réduire leur puissance de transmission. Ce genre de situation est visible par exemple aux figures 4.2 et 4.3. Dans ce genre de cas il est difficile de trouver une relation d'ordre qui permette d'exprimer les correspondances  $\varphi_i$  comme à l'équation (D.7). Une condition suffisante serait d'avoir

$$\phi(\mathbf{p}_{-i}) = \inf_{P_i \in \mathcal{A}_i} \left\{ P_i : \alpha_i \sum_{n=1}^{N_i} \log(1 + P_i \gamma_{i,n}(\mathbf{p}_{-i}) q_{i,n}) \geq \Gamma_i \right\} \quad (D.8)$$

avec une configuration des canaux permettant de vérifier que  $\gamma_{i,n}(\mathbf{p}'_{-i}) \leq \gamma_{i,n}(\mathbf{p}_{-i})$  pour tous  $\mathbf{p}_{-i}, \mathbf{p}'_{-i} \in \mathcal{P}_i$  tels que  $\|\mathbf{p}_{-i}\|_1 < \|\mathbf{p}'_{-i}\|_1$ . Dans ce cas, on aurait bien  $\phi_i$  qui préserve l'ordre.

Toutes les réalisations des simulations du jeu de satisfaction OFDM avec connaissance de l'état du canal et réponses de satisfaction basées sur le water-filling ont néanmoins convergé, ce qui semble indiquer que soit les conditions évoquées ci-dessus sont peu restrictives, soit il existe une relation d'ordre pour laquelle il est possible d'exprimer les  $\phi_i$  de la manière requise.

## D.2.2 Allocation ON/OFF et retour sur 1 bit

L'espace de stratégies des joueurs de ce jeu correspond au nombre de sous-canaux où une puissance  $P_i = P_i^{\max}/N_i$ , est allouée *i.e.*

$$\mathcal{A}_i = \{1, \dots, N_i\}. \quad (\text{D.9})$$

Les vecteurs d'allocation de puissance possibles sont

$$\mathcal{P}_i(a_i) = \left\{ \mathbf{p}_i \in \{0, P_i\}^{N_i} : \|\mathbf{p}_i\|_0 = a_i \right\}. \quad (\text{D.10})$$

Les stratégies choisies par  $i$  sont notées  $a_i \in \mathcal{A}_i$  et le vecteur d'allocation de puissance correspondant est  $\mathbf{p}_i(a_i)$ . De manière analogue,  $a_{-i} = (a_j)_{j \in \mathcal{I}, j \neq i} \in \mathcal{A}_{-i}$  et  $\mathbf{p}_{-i}(a_{-i})$  sont, respectivement, les stratégies et les vecteurs d'allocations de puissances associés des adversaires de  $i$ . La fonction d'utilité donnée initialement à l'équation (4.29) peut s'exprimer comme

$$u_i(a_i, a_{-i}) = \sum_{n=1}^{N_i} \log(1 + \gamma_{i,n}(\mathbf{p}_{-i}(a_{-i})) p_{i,n}(a_i)) \quad (\text{D.11})$$

avec

$$\gamma_{i,n}(\mathbf{p}_{-i}(a_{-i})) \triangleq \frac{e^{\mathbb{E}\{\log|h_{ii,n}|^2\}}}{\sigma_{w_i,n}^2 + \sum_{j \neq i} \mathbb{E}\{|h_{ji,n}|^2\} p_{j,n}(a_j)}. \quad (\text{D.12})$$

La correspondance de satisfaction  $\varphi_i$  satisfait

$$\varphi_i(a_{-i}) = \{a_i \in \mathcal{A}_i : \phi_i(a_{-i}) \preceq a_i\} \quad (\text{D.13})$$

avec  $\phi_i(a_{-i}) = \inf_{a_i \in \mathcal{A}_i} \{a_i : f_i(a_i, a_{-i}) \geq \Gamma_i\}$ .

Pour  $a_{-i} < a'_{-i}$ , on a  $\gamma_{i,n}(\mathbf{p}_{-i}(a'_{-i})) \leq \gamma_{i,n}(\mathbf{p}_{-i}(a_{-i}))$ , c'est à dire que lorsque les adversaires de  $i$  allouent de la puissance sur un plus grand nombre de canaux, le RSIB de  $i$  va soit décroître soit rester égal. Dans ce cas,  $u_i(a_i, a_{-i}) \geq u_i(a_i, a'_{-i})$  donc  $\phi_i$  préserve l'ordre. Par conséquent, les conditions de la proposition 4.2.1 sont satisfaites et l'algorithme 3 converge bien vers un équilibre de satisfaction.



## Bibliographie

- E. Altman and Z. Altman. S-Modular Games and Power Control in Wireless Networks. *IEEE Trans. on Automatic Control*, 48, 2003.
- A. J. G. Anandkumar, A. Anandkumar, S. Lambbotharan, and J. A. Chambers. Robust Rate Maximization Game Under Bounded Channel Uncertainty. *IEEE Trans. Vehicular Technologies*, 60(9) :4471–4486, Nov 2011.
- P. Anjani and M. Chitre. Model-based Data-driven Learning Algorithm for Tuning an Underwater Acoustic Link. In *Proceedings of the 4th Underwater Communications and Networking Conference (Ucomms), Lerici*, 2018.
- R. J. Aumann. Correlated Equilibrium as an Expression of Bayesian Rationality. *Econometrica*, 55(1) :1–18, Janvier 1986.
- R. J. D. Aumann and M. B. Maschler. *Repeated Games with Incomplete Information*. MIT Press, 1995.
- Y. Aval, S. K. Wilson, and M. Stojanovic. On the Achievable Rate of a Class of Acoustic Channels and Practical Power Allocation Strategies for OFDM Systems. *IEEE J. of Oceanic Eng.*, 40(4) :785–795, Oct 2015.
- G. Bacci, S. Lasaulce, W. Saad, and L. Sanguinetti. Game Theory for Networks : A tutorial on game-theoretic tools for emerging signal processing applications. *IEEE Signal Processing Magazine*, 33(1) :94–119, January 2016.
- A. Baggeroer. Acoustic telemetry - An overview. *IEEE Journal of Oceanic Engineering*, 9(4) :229–235, October 1984.
- N. Baldo, P. Casari, and M. Zorzi. Cognitive Spectrum Access for Underwater Acoustic Communications. in *Proc. MTS/IEEE OCEANS 2008, Quebec City*, September 2008.
- E. Bejjani and J. C. Belfiore. Multicarrier coherent communications for the underwater acoustic channel. In *OCEANS 96 MTS/IEEE Conference Proceedings. The Coastal Ocean - Prospects for the 21st Century*, volume 3, pages 1125–1130 vol.3, Sep 1996.
- R. E. Bellman. *Dynamic Programming*. Princeton University Press, 1957.
- P. Bello. Characterization of Randomly Time-Variant Linear Channels. *IEEE Transactions on Communications Systems*, 11(4) :360–393, December 1963.
- E.V. Belmega, S. Lasaulce, and M. Debbah. Power allocation games for MIMO multiple access channels with coordination. *IEEE Transactions on Wireless Communications*, 8(6) :3182–3192, June 2009.
- C. R. Berger, S. Zhou, J. C. Preisig, and P. Willett. Sparse Channel Estimation for Multicarrier Underwater Acoustic Communication : From Subspace Methods to Compressed Sensing. *IEEE Transactions on Signal Processing*, 58(3) :1708–1721, March 2010.

- R. A. Berry and D. Tse. Shannon Meets Nash on the Interference Channel. *IEEE Transactions on Information Theory*, 57(5) :2881–2835, May 2011.
- D.P. Bertsekas and J.N. Tsitsiklis. *Parallel and Distributed Computation : Numerical Methods*. Belmont, MA : Athena Scientific, 1997.
- M. Biagi, A. Petroni, S. Colonnese, R. Cusani, and G. Scarano. On Rethinking Cognitive Access for Underwater Acoustic Communications. *IEEE Journal of Oceanic Engineering*, 41(4) :1045–1060, Oct 2016.
- E. Biglieri, J. Proakis, and S. Shamai. Fading channels : information-theoretic and communications aspects. *IEEE Transactions on Information Theory*, 44(6) :2619–2692, Oct 1998.
- P. Casari and M. Zorzi. Protocol Design Issues in Underwater Acoustic Networks. *Computer Communications*, 34 :2013–2025, 2011.
- S. E. Cho, H. C. Song, and W. S. Hodgkiss. Successive Interference Cancellation for Underwater Acoustic Communications. *IEEE J. of Oceanic Eng.*, 36(4) :490–501, Oct. 2011.
- S. E. Cho, H. C. Song, and W. S. Hodgkiss. Asynchronous multiuser underwater acoustic communications. *The Journal of the Acoustical Society of America*, 132(1) : 5–8, Jul. 2012.
- Y. Chu, S. Konusalp, P.D. Mitchell, D. Grace, and T. Clarke. Application of reinforcement learning to medium access control for wireless sensor networks. *Eng. Applications of Artificial Intelligence*, 46 :23 – 32, 2015.
- C. Clay and H. Medwin. *Fundamentals of Acoustical Oceanography*. Academic Press, San Diego, CA, USA, 1998.
- R.F.W Coates. *Noise and Reverberation*. Macmillan New Electronics Series. Palgrave, London, 1990.
- A. A. Cournot. *Recherches sur les principes de la théorie des richesses*. 1838.
- T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley & Sons, 2006.
- B. Dunshaw. Worldwide Sound Speed, Temperature, Salinity, and Buoyancy from the NOAA World Ocean Atlas. <http://staff.washington.edu/dushaw/WOA/>.
- Francisco Facchinei and Christian Kanzow. Generalized Nash Equilibrium Problems. *Springer-Verlag 4OR*, 5(3) :173–210, September 2007.
- F. Fazel, M. Fazel, and M. Stojanovic. Random Access Compressed Sensing for Energy-Efficient Underwater Sensor Networks. *IEEE J. on Selected Areas in Communications*, 29(8) :1660–1670, September 2011.
- F. Fazel, M. Fazel, and M. Stojanovic. Random Access Compressed Sensing over Fading and Noisy Communication Channel. *IEEE Trans. on Wireless Communications*, 12 (5) :2114–2125, May 2013.

- N. Feng, S.-C. Mau, and N. B. Mandayam. Pricing and power control for joint network-centric and user-centric radio resource management. *IEEE Transactions on Communications*, 52(9) :1547–1557, Sept 2004.
- D. Fudenberg and D. K. Levine. *The Theory of Learning in Games*. Cambridge MA MIT Press, 1998.
- D. Fudenberg and J. Tirole. *Game Theory*. Cambridge MA MIT Press, 1991.
- R. G. Gallager. *Principles of Digital Communications*. Cambridge University Press, 2008.
- Benoit Geller, Vittorio Capellano, and Geneviève Jourdain. Equalizer for real time high rate transmission in underwater communications. In *1995 International Conference on Acoustics, Speech, and Signal Processing, ICASSP '95, Detroit, Michigan, USA, May 08-12, 1995*, pages 3179–3182, 1995.
- M. Goonewardena, S. M. Perlaza, A. Yadav, and W. Ajib. Generalized Satisfaction Equilibrium for Service-Level Provisioning in Wireless Networks. *IEEE Trans. on Communications*, 65(6) :2427–2437, June 2017.
- Z. Han, D. Niyato, W. Saad, T. Basar, and A. Hjørungnes. *Game Theory in Wireless and Communication Networks*. Cambridge University Press, 2012.
- C. Hasan, E. Altman, J-M. Gorce, and M. Haddad. Non-Cooperative Association Of Mobiles To Access Points Revisited. In *WiOpt'12 : Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks, Paderborn, Germany, May 2012*.
- S. Hayashi and Z.Q Luo. Spectrum Management for Interference-Limited Multiuser Communication Systems. *IEEE Trans. on Information Theory*, 55(3) :1153–1175, March 2009.
- G. He, M. Debbah, and S. Lasaulce. K-player Bayesian waterfilling game for fading multiple access channels. In *2009 3rd IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pages 17–20, Dec 2009.
- G. He, M. Debbah, and E. Altman. A Bayesian Game-Theoretic Approach for Distributed Resource Allocation in Fading Multiple Access Channels. *EURASIP Journal on Wireless Communications and Networking, SpringerOpen*, (1), 2010.
- Junling Hu and Michael P. Wellman. Multiagent reinforcement learning : Theoretical framework and an algorithm. In *Proceedings of the Fifteenth International Conference on Machine Learning, ICML '98*, pages 242–250, 1998.
- J. Huang, S. Zhou, and Z. Wang. Performance Results of Two Iterative Receivers for Disitributed MIMO OFDM With Large Doppler Deviations. *IEEE Journal of Oceanic Engineering*, 38(2) :347–357, Apr. 2013.
- J. M. Jornet, M. Stojanovic, and M. Zorzi. On Joint Frequency and Power Allocation in a Cross-Layer Protocol for Underwater Acoustic Networks. *IEEE Journal of Oceanic Engineering*, 35(4) :936–947, Oct 2010.

- D. Kilfoyle and A. Baggeroer. The State of the Art in Underwater Acoustic Telemetry. *IEEE J. of Oceaninc Eng.*, 25(1) :4–27, Jan 2000.
- V. O. Knudsen, R. S. Alford, and J. W. Emling. Underwater Ambient Noise. *Journal of Maritime Research*, 1948.
- S. Konusalp, Y. Chu, P.D. Mitchell, D. Grace, and T. Clarke. Use of Q-learning approaches for practical medium access control in wireless sensor networks. *Eng. Applications of Artificial Intelligence*, 55 :146 – 154, 2016.
- C. Laot and R. LeBidan. Adaptive MMSE turbo equalization with high-order modulations and spatial diversity applied to underwater acoustic communications. In *17th European Wireless 2011 - Sustainable Wireless Technologies*, pages 1–6, April 2011.
- A. Lapidoth and S.M. Moser. Capacity Bounds via Duality with Applications to Multiple-Antenna Systems on Flat-Fading Channels. *IEEE Trans. Information Theory*, 49(10) :2426–2467, October 2003.
- S. Lasaulce and H. Tembine. *Game Theory and Learning for Wireless Networks*. Elsevier Academic Press, 2011.
- Joel Z. Leibo, Vinícius Flores Zambaldi, Marc Lanctot, Janusz Marecki, and Thore Graepel. Multi-agent Reinforcement Learning in Sequential Social Dilemmas. *CoRR*, abs/1702.03037, 2017.
- A. Lesherm and E. Zehavi. Game Theory and the Frequency Selective Interference Channel. *IEEE Signal Processing Mag.*, 26 :28–40, September 2009.
- K. Leyton-Brown and Y. Shoham. *Multiagent Systems : Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge University Press, 2009.
- B. Li, J. Huang, S. Zhou, K. Ball, M. Stojanovic, L. Freitag, and P. Willett. MIMO-OFDM for High-Rate Underwater Acoustic Communications. *IEEE Journal of Oceanic Engineering*, 34(4) :634–644, Oct 2009.
- Michael L. Littman. Markov games as a framework for multi-agent reinforcement learning. In *Proceedings of the Eleventh International Conference on International Conference on Machine Learning*, ICML’94, pages 157–163, 1994.
- G. Loubet, J Lacoume, B. Faure, and F. Vial. Effet de la dispersion sur la propagation acoustique sous marine. *Colloque de Physique*, 51 (C2)(2) :945–948, Février 1990.
- X. Lurton. *Acoustique sous-marine : présentation et applications*. Ifremer, Bibliothèque de Brest, 1998.
- X. Lurton. *An Introduction to Underwater Acoustics : Principles and Applications*. Springer, 2011.
- A. Mahmood, M. Chitre, and M. A. Armand. PSK Communication with Passband Additive Symmetric alpha-Stable Noise. *IEEE Transactions on Communications*, 60 (10) :2990–3000, October 2012.
- J. Maynard-Smith and G. R. Price. The Logic of Animal Conflict. *Nature*, 246 :15–19, November 1973.

- E. McCarthy. *International Regulation Of Underwater Sound : Establishing Rules and Standards to Address Ocean Noise Pollution*. Kluwer Academic Publishers, 2004.
- J. McGee, J. Catipovic, S. Schoenecker, and P. Swaszek. Leveraging spatial diversity to mitigate partial band interference in undersea networks through waveform reconstruction. In *Proc. MTS/IEEE OCEANS 2014 - St John's*, pages 1–8, Sept 2014.
- P. Mertikopoulos, E.V. Belmega, A.L. Moustakas, and S. Lasaulce. Distributed Learning Policies for Power Allocation in Multiple Access Channels. *IEEE J. Selected Areas on Communications*, 30(1) :96–106, January 2012.
- J. Mitola. *Cognitive Radio : an integrated agent architecture for software defined radio*. PhD Dissertation, Royal Institute of Technology (KTH), 2000.
- M. Molins and M. Stojanovic. Slotted FAMA : a MAC protocol for underwater acoustic networks. In *OCEANS 2006 - Asia Pacific*, pages 1–7, May 2006.
- Dov Monderer and Lloyd S. Shapley. Potential games. *Games and Economic Behavior*, 14(1) :124 – 143, 1996.
- N. Morozs, T. Clarke, and D. Grace. Cognitive spectrum management in dynamic cellular environments : A case-based Q-learning approach. *Eng. Applications of Artificial Intelligence*, 55 :239 – 249, 2016.
- A. S. Motahari and A. K. Khandani. Capacity Bounds for the Gaussian Interference Channel. *IEEE Transactions on Information Theory*, 55(2) :620–643, February 2009.
- F. Mériaux, S. M. Perlaza, S. Lasaulce, Z. Han, and V. Poor. Achievability of Efficient Satisfaction Equilibria in Self Configuring Networks. In *Int. Conf. on Game Theory for Networks (GameNets) 2012, Vancouver, Canada*, pages 1–15, May 2012.
- National Academy of Sciences NAS. *Present and Future Civil Uses of Underwater Sound*. Washington DC : National Academy of Sciences, 1970.
- J. Nash. The Bargaining Problem. *Econometrica*, 18(2) :155–162, April 1950.
- J. F. Nash. Non Cooperative Games. *The Annals of Mathematics*, 54 :286–295, September 1951.
- Centre for Maritime Research & Experimentation NATO. JANUS, the CMRE underwater communications protocol becomes a NATO Standard. April 2017. available at <http://www.cmre.nato.int/news-room/news-room/>.
- Y. Noam, A. Leshem, and H. Messer. A Rayleigh Fading Interference Game with Incomplete Information. In *2010 IEEE Sensor Array and Multichannel Signal Processing Workshop*, pages 49–52, Oct 2010.
- M. J. Osborne and A. Rubinstein. *A Course in Game Theory*. Cambridge MA MIT Press, 1994.
- R. Otnes, A. Asterjadhi, P. Casari, M. Goetz, T. Husoy, I. Nissen, K. Rimstad, P. van Walree, and M. Zorzi. *Underwater Acoustic Networking Techniques*. Springer Briefs in Electrical and Computer Eng., 2012.

- J. S. Pang, G. Scutari, F. Facchinei, and C. Wang. Distributed Power Allocation With Rate Constraints in Gaussian Parallel Interference Channels. *IEEE Trans. on Information Theory*, 54(8) :3471–3489, Aug 2008.
- S. M. Perlaza, H. Tembine, S. Lasaulce, and M. Debbah. Quality-of-Service Provisioning in Decentralized Networks : A Satisfaction Equilibrium Approach. *IEEE J. of Selected Topics in Signal Processing*, pages 104–116, 2012.
- M. B. Porter. The BELLHOP Manual and Users Guide : Preliminary draft. *Heat, Light and Sound Research, Inc., La Jolla, CA, USA, Tech Report*, 2011.
- J. Potter, J. Alves, D. Green, G. Zappa, I. Nissen, and K. McCoy. The JANUS underwater communication standard. *Proc. 1st Underwater Communications and Networking Conf. (Ucomms)*, pages 1–4, Sept. 2014.
- A. Pottier, F.-X. Socheleau, and C. Laot. Power-Efficient Spectrum Sharing for Noncooperative Underwater Acoustic Communication Systems. In *MTS/IEEE OCEANS'16 Monterey*, Sept. 2016a.
- A. Pottier, F.-X. Socheleau, and C. Laot. Distributed Power Allocation Strategy in Shallow Water Acoustic Interference Channels. In *IEEE Int. workshop on Signal Processing Advances in Wireless Communications (SPAWC) 2016*, Jul. 2016b.
- A. Pottier, F.-X. Socheleau, and C. Laot. Adaptive Power Allocation for Noncooperative OFDM Systems in UWA Interference Channels. In *Proc. 3rd Underwater Acoustic Communications and Networking Conf. (UComms)*, Sept. 2016c.
- A. Pottier, F.-X. Socheleau, and C. Laot. Robust Noncooperative Spectrum Sharing Games in Underwater Acoustic Interference Channels. *IEEE J. of Oceanic Eng.*, 42(4) :1019 – 1034, October 2017.
- A. Pottier, P. D. Mitchell, F.-X. Socheleau, and C. Laot. Q-Learning based Adaptive Channel Selection for Underwater Sensor Networks. In *Proc. 4th Underwater Acoustic Communications and Networking Conf. (UComms)*, August. 2018a.
- A. Pottier, F.-X. Socheleau, and C. Laot. Quality-of-Service Satisfaction Games for Noncooperative Underwater Acoustic Communications. *IEEE Access*, 6, April 2018b.
- J. G. Proakis. *Digital Communications*. McGraw-Hill, 2001.
- P. Qarabaqi and M. Stojanovic. Adaptive power control for underwater acoustic communications. In *OCEANS 2011 IEEE - Spain*, pages 1–7, June 2011a.
- P. Qarabaqi and M. Stojanovic. Modeling the large scale transmission loss in underwater acoustic channels. In *2011 49th Annual Allerton Conf. on Communication, Control, and Computing (Allerton)*, pages 445–452, Sept 2011b.
- P. Qarabaqi and M. Stojanovic. Statistical Characterization and Computationally Efficient Modeling of a Class of Underwater Acoustic Communication Channels. *IEEE J. Ocean. Eng.*, 38(4) :701–717, October 2013.
- A. Radosevic, T. M. Duman, J. G. Proakis, and M. Stojanovic. Channel prediction for adaptive modulation in underwater acoustic communications. In *OCEANS 2011 IEEE - Spain*, pages 1–5, June 2011.

- A. Radosevic, R. Ahmed, T.M. Duman, J. G. Proakis, and M. Stojanovic. Adaptive OFDM Modulation for Underwater Acoustic Communications : Design Considerations and Experimental Results. *IEEE J. of Oceaninc Eng.*, 39(2) :357–370, Apr 2014.
- J. B. Rosen. Existence and Uniqueness of Equilibrium Points for Concave N-Persons Games. *Econometrica*, 33(3) :521–534, July 1965.
- S. Ross and B. Chaib-draa. Satisfaction Equilibrium : Achieving Cooperation in Incomplete Information Games. In *Proc. Canadian Conf. on AI, Quebec City*, pages 61–72, June 2006.
- C. U. Saraydar, N. B. Mandayam, and D. J. Goodman. Efficient power control via pricing in wireless data networks. *IEEE Transactions on Communications*, 50(2) : 291–303, Feb 2002.
- D. A. Schmidt, C. Shi, R. A. Berry, M. L. Honig, and W.S Utschick. Distributed resource allocation schemes : Pricing algorithms for power control and beamformer design in interference networks. *IEEE Signal Processing Magazine*, 26(5) :53–63, January 2009.
- G. Scutari, D.P. Palomar, and S. Barbarossa. Distributed totally asynchronous iterative waterfilling for wideband interference channel with time/frequency offset. In *IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP) 2007*, volume 4, pages IV–1325–IV–1328, April 2007.
- G. Scutari, D. P. Palomar, and S. Barbarossa. Optimal Linear Precoding Strategies for Wideband Non-Cooperative Systems Based on Game Theory-Part I : Nash Equilibria. *IEEE Trans. on Signal Processing*, 56(3) :1230–1249, March 2008a.
- G. Scutari, D. P. Palomar, and S. Barbarossa. Optimal Linear Precoding Strategies for Wideband Non-Cooperative Systems Based on Game Theory-Part II : Algorithms. *IEEE Trans. on Signal Processing*, 56(3) :1250–1267, March 2008b.
- G. Scutari, D.P. Palomar, and S. Barbarossa. Asynchronous Iterative Waterfilling for Gaussian Frequency-Selective Interference Channels : A Unified Framework. *Transactions on Information Theory*, 54(7) :2868–2879, July 2008c.
- G. Scutari, F. Facchinei, J-S. Pang, and D. P. Palomar. Real and Complex Monotone Communication Games. *IEEE Trans. on Inform. Theory*, 60(7) :4197–4231, July 2014.
- B. S. Sharif, J. Neasham, O. R. Hinton, and A. E. Adams. A computationally efficient Doppler compensation system for underwater acoustic communications. *IEEE Journal of Oceanic Engineering*, 25(1) :52–61, Jan 2000.
- A. C. Singer, J. K. Nelson, and S. S. Kozat. Signal Processing for Underwater Acoustic Communications. *IEEE Communications Mag.*, 2009.
- F-X. Socheleau, C. Laot, and Passerieux J-M. A Maximum Entropy Framework for Statistical Modeling of Underwater Acoustic Communication Channels. In *MTS/IEEE Oceans 2010, Sydney*, pages 1–7, June 2010a.

- F.-X. Socheleau, C. Laot, and J.-M. Passerieux. Concise Derivation of Scattering Function from Channel Entropy Maximization. *IEEE Trans. on Communications*, 58 (11) :3098–3103, November 2010b.
- F.-X. Socheleau, C. Laot, and J.-M. Passerieux. Stochastic Replay of non-WSSUS Underwater Acoustic Communication Channels Recorded at Sea. *IEEE Trans. Signal Process.*, 59(10) :4838–4849, June 2011.
- F.-X. Socheleau, A. Pottier, and C. Laot. Stochastic Replay of SIMO Underwater Acoustic Communication Channels. *OCEANS 2015*, pages 1–6, October 2015.
- E.M. Sozer, M. Stojanovic, and J.G. Proakis. Underwater Acoustic Networks. *IEEE J. of Oceanic Eng.*, 25(1) :72–83, January 2000.
- M. Stojanovic and J. Preisig. Underwater Acoustic Communication Channels : Propagation Models and Statistical Characterization. *IEEE Communications Mag.*, 2009.
- M. Stojanovic and Z. Zvonar. Multichannel processing of broad-band multiuser communication signals in shallow water acoustic channels. *IEEE Journal of Oceanic Engineering*, 21(2) :156–166, Apr 1996.
- M. Stojanovic, J. Catipovic, and J. G. Proakis. Adaptive multichannel combining and equalization for underwater acoustic communications. *The Journal of the Acoustical Society of America*, 94(3) :1621–1631, 1993.
- M. Stojanovic, J. A. Catipovic, and J. G. Proakis. Phase-coherent digital communications for underwater acoustic channels. *IEEE Journal of Oceanic Engineering*, 19 (1) :100–111, Jan 1994.
- M. Stojanovic, J. G. Proarkis, J. A. Rice, and M. D. Green. Spread spectrum underwater acoustic telemetry. In *OCEANS '98 Conference Proceedings*, volume 2, pages 650–654 vol.2, Sep 1998.
- M. Stojanovic, S. Shahabudeen, and M. Chitre. Underwater Acoustic Communications and Networking : Recent advances and future challenges. *Marine Technology Society J.*, 42(1), 2008.
- R. Sutton and A. Barto. *Reinforcement Learning : An Introduction*. MIT Press, Cambridge, MA, 1998.
- A. A. Syed, W. Ye, and J. Heidemann. T-lohi : A new class of mac protocols for underwater acoustic sensor networks. In *IEEE INFOCOM 2008 - The 27th Conference on Computer Communications*, April 2008.
- P. L. Tyack. Population biology, social behavior and communication in whales and dolphins. *Trends in Ecology and Evolution*, 1(6) :144 – 150, 1986.
- R. J. Urick. *Principles of Underwater Sound, 3rd Edition*. McGraw-Hill, 1983.
- R. J. Urick. *Ambient Noise in the Sea*. Undersea technology Warfare Office, Naval Sea Systems Command, Department of the Navy, Washington D.C. 20362, 1984.

- P. van Walree. *Channel sounding for acoustic communications : techniques and shallow-water examples*. Norwegian Defence Research Establishment (FFI), Kjeller, Norway, 2011.
- P. van Walree and R. Otnes. Ultrawideband Underwater Acoustic Communication Channels. *IEEE J. of Oceanic Eng.*, 38(4) :678–688, October 2013.
- P. van Walree, P. Jenserud, and M. Smedsrud. A Discrete-Time Channel Simulator Driven by Measured Scattering Functions. *IEEE J. Selected Areas in Communications*, 26(9) :1628–1637, 2008a.
- P. van Walree, E. Sangfelt, and G. Leus. Multicarrier spread spectrum for covert acoustic communications. In *OCEANS 2008*, pages 1–8, Sept 2008b.
- P. A. van Walree and G. Leus. Robust Underwater Telemetry With Adaptive Turbo Multiband Equalization. *IEEE Journal of Oceanic Engineering*, 34(4) :645–655, Oct 2009.
- T. L. Vincent and J. S. Brown. *Evolutionary Game Theory, Natural Selection, and Darwinian Dynamics*. Cambridge University Press, 2005.
- J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behaviors*. Princeton University Press, 1944.
- C. Wang, Z. Wang, W. Sun, and D. R. Fuhrman. Reinforcement Learning based Adaptive Transmission in Time-Varying Underwater Acoustic Channels. *IEEE Access*, 6 : 2541–2558, 2018.
- F. Wang, M. Krunz, and S. Cui. Price-based spectrum management in cognitive radio networks. *IEEE Journal of Selected Topics in Signal Processing*, 2(1) :74–87, Feb 2008.
- Z. Wang, S. Zhou, J. Catipovic, and P. Willett. Parameterized Cancellation of Partial-Band Partial-Block-Duration Interference for Underwater Acoustic OFDM. *IEEE Trans. Signal Processing*, 60(4) :1782–1795, Apr. 2012.
- Z. Wang, S. Zhou, J. Catipovic, and P. Willett. Asynchronous Multiuser Reception for OFDM in Underwater Acoustic Communications. *IEEE Transactions on Wireless Communications*, 12(3) :1050–1061, March 2013.
- Christopher J. C. H. Watkins and Peter Dayan. Q-learning. *Machine Learning*, 8(3) : 279–292, May 1992.
- G. M. Wenz. Review of Underwater Acoustics Research : Noise. *Journal of the Acoustical Society of America*, 1972.
- T. C. Yang and Wen-Bin Yang. Performance analysis of direct-sequence spread-spectrum underwater acoustic communications with low signal-to-noise-ratio input signals. *The Journal of the Acoustical Society of America*, 123(2) :842–855, 2008.
- W. B. Yang and T. C. Yang. High-Frequency Channel Characterization for M-ary Frequency Shift Keying Underwater Acoustic Communications. *J. of the Acoustical Society of America*, 120(5) :2615–2626, August 2006.

- Y. Yang, G. Scutari, D. P. Palomar, and M. Pesavento. A Parallel Decomposition Method for Nonconvex Stochastic Multi-Agent Optimization Problems. *IEEE Transactions on Signal Processing*, 64(11) :2949–2964, Juin 2016.
- Y. Zhao and G. J. Pottie. Optimal Spectrum Management in Multiuser Interference Channels. *IEEE Transactions on Information Theory*, 59(8) :4961–4976, August 2013.
- S. Zhou and Z. Wang. *OFDM for underwater acoustic communications*. John Wiley & Sons, 2014.



---

**Titre : Méthodes décentralisées d'allocation des ressources dans le canal d'interférence acoustique sous-marin**

**Mots clés :** communications numériques, théorie des jeux, acoustique sous-marine

**Résumé :** L'onde acoustique est utilisée par de nombreux systèmes et organismes marins pour communiquer, naviguer ou inférer de l'information sur l'environnement. Le développement des activités humaines liées au monde de la mer induit une augmentation du nombre de sources acoustiques en activité simultanée dans l'océan. L'environnement acoustique sous-marin (ASM) est donc partagé par de nombreuses sources hétérogènes (sonars, modems, mammifères marins, ...) entrant involontairement en compétition pour l'utilisation d'une ressource offerte par le canal de transmission.

L'objectif de ce travail de thèse est de proposer des solutions permettant aux systèmes de communications ASM d'adapter leurs paramètres de transmission de manière intelligente, autonome et décentralisée selon l'environnement acoustique dans lequel ils évoluent. A plusieurs égards, les problématiques de ce sujet de thèse sont proches de celles ayant motivé le développement des recherches sur la radio cognitive. Cependant, les spécificités du milieu acoustique, les sources d'interférences de diverses natures, et l'absence de standards de communications posent de nouvelles difficultés.

---

**Title : Decentralized methods for resource allocation in the underwater acoustic interference channel**

**Keywords :** digital communications, game theory, underwater acoustics

**Abstract :** Underwater acoustic waves are used by many systems and biologic organisms to communicate, navigate or infer information about the environment. Future developments of human maritime activities imply an increase of the number of active acoustic sources in the oceans. The underwater environment is therefore shared by many heterogeneous sources (sonars, modems, marine mammals, ...) competing involuntarily for using the physical resources offered by the communication channel.

The goal of this thesis is to provide solutions allowing autonomous and decentralized adaptation of the transmission strategies of underwater acoustic communication systems, according to the environment. To some extent, this work deals with topics that are closely related to what has motivated the first researches on cognitive radio systems. However, the specific properties of the underwater environment, the heterogeneity of interfering acoustic sources, and the absence of communication standards rise new difficulties.