



HAL
open science

Stratégies d'allocation de ressources basées sur la théorie des jeux et l'apprentissage par renforcement pour les réseaux acoustiques sous-marins

Khouloud Gharsalli

► To cite this version:

Khouloud Gharsalli. Stratégies d'allocation de ressources basées sur la théorie des jeux et l'apprentissage par renforcement pour les réseaux acoustiques sous-marins. Acoustique [physics.class-ph]. Université de Brest; Université de Carthage (Tunisie), 2025. Français. ⟨NNT : 2025BRUN0105⟩. ⟨tel-05578933⟩

HAL Id: tel-05578933

<https://theses.hal.science/tel-05578933v1>

Submitted on 3 Apr 2026

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.



HAL Authorization

THESE DE DOCTORAT EN COTUTELLE DE

L'UNIVERSITE
DE BRETAGNE OCCIDENTALE

ECOLE DOCTORALE N° 644

Mathématiques et Sciences et Technologies

de l'Information et de la Communication en Bretagne Océane

Spécialité : Télécommunications – Signal, Image, Vision, Son

AVEC

ECOLE DOCTORALE EN TIC (EDTIC)

Spécialité : Technologie de l'Information et de la Communication

Par

Khouloud GHARSALLI

**Stratégies d'allocation de ressources basées sur la théorie des jeux et
l'apprentissage par renforcement pour les réseaux acoustiques
sous-marins**

Thèse présentée et soutenue à SUP'COM, Tunis, le 12 Décembre 2025

Unités de recherche : *L@bISEN & COSIM*

Rapporteurs avant soutenance :

Benoît PARREIN Professeur des universités, Université de Rennes
Zouhair BEN JEMAA Maître de conférences habilité, Faculté des sciences de Bizerte

Composition du Jury :

Présidente : Fatma ABDELKEFI Professeure, SUP'COM
Rapporteurs : Benoît PARREIN Professeur des universités, Université de Rennes
Zouhair BEN JEMAA Maître de conférences habilité, Faculté des sciences de Bizerte
Examineur : Alexandre PAUCHET Maître de conférences HDR, INSA Rouen
Dir. de thèse : Pierre Jean BOUVET Enseignant chercheur/HDR, ISEN Ouest
Co-dir. de thèse : Hichem BESBES Professeur, SUP'COM

Encadrants

Encadrante : Sameh NAJEH Maitresse de conférences, SUP'COM
Co-encadrant : Thierry LE PORS Enseignant-chercheur, ISEN Ouest

Invités

Leila NAJJAR Professeure, SUP'COM

*À ma chère famille,
Pour votre amour, votre patience et votre soutien indéfectible tout au
long de ce parcours.
Vous êtes ma plus grande source de force et d'inspiration.*

Remerciements

Je souhaite exprimer avant tout ma profonde reconnaissance à mes directeurs de thèse, le Professeur Hichem BESBES et le Dr/HDR Pierre-Jean BOUVET, pour leur encadrement bienveillant, la pertinence de leurs conseils et la confiance qu'ils m'ont témoignée tout au long de ce travail. Leurs compétences complémentaires ainsi que leur disponibilité ont largement contribué à l'aboutissement de cette recherche.

J'adresse également mes sincères remerciements à Thierry LE PORS, Sameh NAJEH et Leila NAJJAR, dont les précieuses suggestions et l'accompagnement attentif ont enrichi et orienté de manière significative mes recherches.

Je tiens à exprimer toute ma gratitude aux membres du jury pour l'honneur qu'ils me font en acceptant d'évaluer ce travail. Je remercie Mme Fatma ABDELKEFI, présidente du jury, ainsi que M. Zouhaier BEN JEMAA et M. Benoît PARREIN, rapporteurs, pour le temps consacré à l'examen attentif de ce manuscrit et pour la richesse de leurs remarques. J'adresse également mes remerciements à M. Alexandre PAUCHET, examinateur, pour son intérêt porté à mes travaux.

Ma gratitude s'étend aussi aux équipes du laboratoire COSIM à SUP'COM et du Lab@ISEN à l'ISEN de Brest, qui m'ont accueilli avec générosité, offert un environnement de recherche stimulant et permis de travailler dans d'excellentes conditions.

Je tiens également à remercier mes collègues et amis des deux institutions, avec qui j'ai partagé non seulement des échanges scientifiques et des discussions constructives, mais aussi de nombreux moments de convivialité.

Enfin, j'exprime toute ma gratitude à ma famille et à mes proches pour leur soutien indéfectible, leur patience et leurs encouragements constants. Leur présence et leur confiance ont constitué une source essentielle de motivation, sans laquelle ce travail n'aurait pas pu aboutir.

Résumé

Le milieu sous-marin constitue un environnement acoustique partagé entre une grande diversité d'acteurs, naturels (mammifères marins, faune benthique) et artificiels (modems de communication, sonars, capteurs, systèmes de télémétrie), contrairement aux radio-communications terrestres, où le spectre est strictement encadré par des régulations internationales, l'environnement Acoustique sous-marin(e) (*ASM*) demeure non coordonné, non normalisé, et soumis à une pression croissante du fait des activités anthropiques (océanographie, défense, prospection, exploitation des ressources marines, etc.).

Dans ce contexte, la cohabitation non réglementée de nombreuses sources concurrentes peut conduire à un masquage mutuel des signaux, compromettant sévèrement la fiabilité des communications. Cette situation met en lumière une problématique majeure : le partage efficace et équitable des ressources acoustiques limitées (temps, fréquence, puissance) entre les différents utilisateurs du canal. En effet, sans mécanismes adaptés pour gérer l'accès multiple au canal, les interférences et la concurrence entre les utilisateurs peuvent rapidement dégrader les performances globales du système.

Cette thèse vise à résoudre le problème du partage des ressources dans les environnements acoustiques sous-marins en proposant des stratégies d'accès multiple à la fois robustes, adaptatives et décentralisées. Les solutions envisagées offrent la capacité aux systèmes de communication d'adapter automatiquement et intelligemment leurs paramètres de transmission selon l'état du canal et l'activité des utilisateurs, tout en s'affranchissant de toute coordination centralisée.

Pour atteindre cet objectif, la thèse adopte une approche interdisciplinaire s'appuyant sur la théorie des jeux coopératifs et l'apprentissage par renforcement. Plutôt que de considérer les utilisateurs comme des compétiteurs opportunistes, elle explore des modèles de coopération entre agents acoustiques, où chaque nœud vise à maximiser la performance globale du réseau et à réduire les interférences grâce à des décisions locales mais coordonnées. Ce paradigme, inspiré des réseaux cognitifs terrestres, s'avère particulièrement adapté au contexte sous-marin, caractérisé par un environnement dynamique, bruité et à faible bande passante, où la gestion centralisée demeure difficile à mettre en œuvre.

Mots clés : communications acoustiques sous-marines, allocation de ressources, théorie des jeux non coopératifs, théorie des jeux coopératifs, apprentissage par renforcement, Q-learning.

Abstract

The underwater environment is an acoustic medium shared by a wide range of actors, both natural (e.g., marine mammals, benthic fauna) and artificial (e.g., communication modems, sonars, sensors, telemetry systems), unlike terrestrial radio communications, where spectrum usage is strictly regulated by international standards, the Underwater acoustic (UWA) environment remains largely uncoordinated and unstandardized. It is also facing increasing pressure due to the expansion of human activities such as oceanographic research, military operations, and resource exploitation.

In this context, the unregulated coexistence of numerous competing sources can result in mutual signal masking, severely degrading the reliability of underwater communications. This situation highlights a major challenge : the efficient and fair sharing of limited acoustic resources (time, frequency, power) among multiple users of the channel so without appropriate mechanisms to handle multiple access, the interference and the competition between users can degrade the overall performance of the system.

This thesis aims to address the resource allocation problem in underwater acoustic environments through the development of robust, adaptive, and decentralized multiple-access strategies. The proposed approaches allow communication systems to intelligently and autonomously optimize their transmission parameters according to channel dynamics and network activity, eliminating the need for centralized control.

To achieve this goal, the thesis adopts an interdisciplinary approach based on cooperative game theory and reinforcement learning. Rather than considering users as opportunistic competitors, she explores cooperation models between acoustic agents, where each node aims to maximize the overall performance of the network and reduce interference through local but coordinated decisions. This paradigm, inspired by terrestrial cognitive networks, is particularly adapted to the underwater context, characterized by a dynamic, noisy and low-bandwidth environment, where centralized management remains difficult to implement.

Keywords : Underwater acoustic communications, resource allocation, non-cooperative game theory, cooperative game theory, reinforcement learning, Q-Learning.

Table des matières

Remerciements	3
Résumé	4
Abstract	5
Table des matières	9
Table des figures	10
Table des figures	11
Liste des tableaux	12
Liste des tableaux	12
Liste des algorithmes	13
Acronymes	14
Notations	17
Introduction générale	18
Introduction générale	18
1 Communications acoustiques sous-marines (CASM)	22
1.1 Introduction	22
1.2 Communications sans fil	23
1.3 Canal acoustique sous-marin : caractéristiques et défis	25
1.3.1 Propriétés physiques du milieu de propagation	25
1.3.2 Contraintes pour les communications ASM	32
1.4 Partage des ressources : Problématique de l'accès multiple	33
1.4.1 État de l'art des protocoles d'accès multiple pour les réseaux acoustiques sous-marins	33
1.4.2 Technique de modulation Orthogonal Frequency Division Multiplexing (OFDM)	38
1.4.3 Position du problème d'allocation de ressources en réseau multi-utilisateurs	40
1.4.4 Solutions classiques et modernes : état de l'art	40
1.5 Conclusion	45

2	Théorie des jeux dans les réseaux acoustiques sous marins	46
2.1	Théorie des jeux pour l'allocation de ressources	46
2.1.1	Définitions et types de jeux	47
2.1.2	Équilibre de Nash (Nash Equilibrium)	49
2.2	Application à la gestion des ressources en ASM	50
2.2.1	Jeux non coopératifs pour l'allocation des ressources	51
2.2.2	Modélisation du problème par un jeu non coopératif	51
2.2.3	Formulation du problème de partage des ressources	52
2.2.4	Choix de la fonction d'utilité	54
2.2.5	Condition de Nash Equilibrium (NE)	57
2.2.6	Meilleures Réponses, Résolution du Jeu et Algorithme d'Allocation de Puissance	57
2.2.7	Unicité et convergence vers NE	59
2.3	Résultats de simulation et analyse	60
2.3.1	Paramètres et hypothèses de simulation	60
2.3.2	Jeu à deux joueurs : impact de l'interférence	60
2.3.3	Impact de la distance sur la fonction d'utilité et la région de capacité	63
2.4	Conclusion	65
3	Stratégies d'allocation coopérative basées sur la négociation dans les CASM	67
3.1	Introduction	67
3.2	Modélisation coopérative basée sur la négociation de Nash	69
3.2.1	Définition, hypothèses et contraintes	69
3.2.2	Modèle mathématique et axiomes de Nash	71
3.2.3	Application à l'allocation de ressources par un jeu coopératif	72
3.2.4	Hypothèse simplificatrice	73
3.2.5	Théorie de la négociation de Nash : cas de deux utilisateurs	73
3.2.6	Conclusion	77
3.3	Apprentissage par renforcement dans les réseaux ASM	77
3.3.1	Introduction à l'apprentissage par renforcement	77
3.3.2	Théorie de la négociation de Nash avec Q-Learning pour un scénario multi-utilisateurs	81
3.3.3	Algorithme proposé de Nash Bargaining Multi-utilisateurs avec Q-Learning	83
3.4	Résultats de simulation et analyse	84
3.4.1	Cas à 2 utilisateurs	84
3.4.2	Cas multi-utilisateurs	89
3.5	Conclusion	93
4	Jeux coopératifs à base de coalitions avec apprentissage distribué	94
4.1	Introduction	94
4.2	Outils de formation de coalitions	96
4.2.1	Jeux en forme caractéristique	96
4.2.2	Algorithmes de formation de coalitions	97
4.2.3	Apprentissage adaptatif pour la formation de coalitions	99
4.3	État de l'art sur la formation de coalitions dans les réseaux sans fil	101
4.3.1	Communications Device-to-Device (D2D)	101
4.3.2	Réseaux Femtocellulaires	102

4.3.3	Réseaux à base de drones et réseaux radar	103
4.3.4	Limites des approches classiques	103
4.3.5	Intelligence Artificielle dans les Jeux de Coalition	104
4.3.6	Réseaux Acoustiques Sous-Marins et Jeux de Coalition	105
4.4	Jeux coopératifs de coalition pour l'allocation des ressources (Connaissance parfaite de Channel State Information (CSI))	107
4.4.1	Introduction aux jeux de coalition et hypothèses appliquées	107
4.4.2	Formulation mathématique du problème	108
4.4.3	Représentation schématique	111
4.4.4	Représentation mathématique	112
4.5	Solution proposée et intégration du Q-Learning pour l'optimisation des coalitions	114
4.5.1	Allocation et formation de coalition initiale	115
4.5.2	Résolution des conflits et allocation de puissance basée sur la valeur de Shapley	117
4.5.3	Apprentissage par renforcement pour optimiser la répartition des utilisateurs au sein des coalitions finales	118
4.6	Allocation des ressources avec variation temporelle (Connaissance partielle de CSI)	122
4.6.1	Modélisation du jeu de coalition sous information partielle sur l'état du canal	123
4.6.2	Modélisation de l'effet Doppler	123
4.7	Conclusion	124
5	Évaluation des contributions apportées au problème d'allocation de ressources en ASM : étude comparative	126
5.1	Introduction	126
5.2	Résultats de simulation et analyse	127
5.2.1	Analyse des résultats : Cas de CSI parfait	128
5.2.2	Étude de la fonction d'utilité globale du système	129
5.2.3	Étude de la fonction d'utilité individuelle	131
5.2.4	Analyse d'équité	132
5.2.5	Analyse de satisfaction	133
5.2.6	Comparaison des algorithmes selon différents niveaux d'interférence	135
5.2.7	Scénario sans K-Means vs Scénario sans Q-learning	137
5.2.8	Analyse de la robustesse avec CSI partielle et augmentation du nombre d'utilisateurs	138
5.3	Discussion sur la complexité de l'algorithme proposé	141
5.4	Conclusion	142
	Conclusion générale et perspectives	144
5.5	Conclusion générale	144
5.6	Perspectives	145
	Liste des publications	148
	Annexes	149
	Bibliographie	155

Bibliographie

155

Index

165

Table des figures

1.1	Profil de célérité en océan profond [Chauvaud et al., 2018].	26
1.2	Coefficient d'atténuation en fonction de la fréquence.	28
1.3	Représentation schématique du Time Division Multiple Access (TDMA), Frequency Division Multiple Access (FDMA), Code Division Multiple Access (CDMA) et Additive Links On-line Hawaii Area (ALOHA).	35
1.4	Représentation schématique du protocole Carrier Sense Multiple Access (CSMA).	35
1.5	Différence entre l'allocation des ressources en OFDM et en Orthogonal Frequency Division Multiple Access (OFDMA)	38
2.1	Canal acoustique à interférences avec K paires émetteur-récepteur	52
2.2	Canal acoustique à interférences avec 2 paires émetteur-récepteur	53
2.3	Réponse fréquentielle des canaux directs et d'interférences simulés selon le modèle de [Qarabaqi and Stojanovic, 2013].	61
2.4	Répartition de puissance sur les sous-porteuses pour 2 utilisateurs obtenue par l'algorithme itératif water-filling dans un jeu non coopératif.	62
2.5	Évolution de la fonction d'utilité de chaque joueur au cours des 30 itéra- tions, selon la distance entre les paires.	63
2.6	Régions de capacité et points de NE en fonction du niveau d'interférence.	65
3.1	Évolution de la fonction d'utilité des utilisateurs (Coopératif vs Non co- opératif).	86
3.2	Évolution de l'utilité en fonction du niveau d'interférence (Coopératif vs Non coopératif).	87
3.3	Évolution de la fonction d'utilité moyenne et globale (Coopératif vs Non coopératif).	88
3.4	Comparaison de la fonction d'utilité individuelle (Jeu coopératif vs Jeu non-coopératif).	90
3.5	Convergence de la fonction d'utilité globale.	91
3.6	Matrice d'allocation des sous-porteuses pour 12 utilisateurs (en haut) et 8 utilisateurs (en bas).	92
4.1	Exemple de formation de coalitions dans un réseau multi-utilisateurs sous- marin.	111
4.2	Processus de la solution proposée.	115
4.3	Organigramme de la solution proposée.	120
5.1	Évolution de la fonction d'utilité globale selon les différentes approches.	130
5.2	Évolution de la fonction d'utilité individuelle selon les différentes approches.	131

5.3	Comparaison de l'indice d'équité des utilités individuelles.	133
5.4	Histogramme des coalitions Pareto-optimales.	134
5.5	Satisfaction des utilisateurs au sein des coalitions.	135
5.6	Évaluation des performances en fonction du nombre d'utilisateurs (CSI Partielle).	139

Liste des tableaux

1.1	Comparaison des technologies de communication sans-fil	25
1.2	Modèles empiriques des densités spectrales de bruit sous-marin.	29
1.3	Comparaison des schémas classiques d'accès multiple dans le contexte ASM	36
2.1	Comparaison des jeux bayésien et robuste	56
2.2	Paramètres	60
3.1	Gains de canal directs moyens pour chaque sous-porteuse	74
3.2	Allocation finale des sous-porteuses	75
3.3	Comparaison des méthodes d'apprentissage par renforcement pour les réseaux ASM	80
3.4	Paramètres.	85
5.1	Paramètres de simulation.	128
5.2	Comparaison de l'indice d'équité moyen : approche Cooperative Coalition Optimization based on Reinforcement Learning for Underwater network (U-CoCRL) vs méthodes classiques.	132
5.3	Comparaison de l'utilité globale (bits/s/Hz) et de l'indice de performance selon différents niveaux d'interférence.	135
5.4	Résumé des performances selon les scénarios d'ablation	138
5.5	Comparaison de Performance Index (PI) sous haute interférence (CSI Partielle).	140
5.6	Comparaison des performances (14 utilisateurs, CSI Partielle, Interférence élevée).	140
5.7	Comparaison de la complexité et de l'efficacité des algorithmes de formation de coalitions.	141

Liste des algorithmes

1	Nash Bargaining Algorithm – Two-User Case : Inspiré de la solution de Nash, ce mécanisme itératif permet à chaque utilisateur d’ajuster progressivement sa répartition de puissance sur les sous-porteuses afin d’améliorer l’utilité individuelle et les performances globales du système.	149
2	Nash Q-learning-based Power Allocation – Multi-Users Case	150
3	Initial coalition formation and resource allocation	151
4	Intra-Coalition Evaluation	152
5	Decentralized Coalition Optimization via Q-Learning	154

Acronymes

CASM Communication acoustique sous-marine

ASM Acoustique sous-marin(e)

AUV véhicules autonomes sous-marins

OTAN Organisation du Traité de l'Atlantique Nord

QL Q-Learning

OFDM Orthogonal Frequency Division Multiplexing

TDMA Time Division Multiple Access

FDMA Frequency Division Multiple Access

CDMA Code Division Multiple Access

ALOHA Additive Links On-line Hawaii Area

CSMA Carrier Sense Multiple Access

MARL Apprentissage par renforcement multi-agents

ISI Inter Symbol Interference

ICI Inter Carrier Interference

SNR rapport signal sur bruit

SINR Rapport signal sur interférence plus bruit

NE Nash Equilibrium

IWF Iterative water filling

NB Nash Bargaining

CSI Channel State Information

DQN Deep Q-Network

SARSA State-Action-Reward-State-Action

RL Reinforcement Learning

DBSCAN Density-Based Spatial Clustering of Applications with Noise

CFSS Coalition Formation with Subset Stability

D2D Device-to-Device

OCF (Overlapping Coalition Formation

RSSI Received Signal Strength Indicator

FUE Femtocell User Equipment

U-CoCRL Cooperative Coalition Optimization based on Reinforcement Learning for Underwater network

IS Individual satisfaction

CS Collective satisfaction

PI Performance Index

AUVs autonomous underwater vehicle

BER bit error rate

PSO Particle Swarm Optimization

NIA Nature Inspired Algorithm

BR Best Response

UWA Underwater acoustic

OFDMA Orthogonal Frequency Division Multiple Access

MDP Markov Decision Process

MAMDP Multi-Agent Markov Decision Process

VHF Very High Frequency

UHF Ultra High Frequency

Notations

\mathbb{C} Ensemble des nombres complexes

\mathbb{R} Ensemble des nombres réels

\mathcal{K} Ensemble de joueurs de 1 à K

$!$ Factorielle

$|\cdot|$ Cardinal

\mathcal{A}_i Ensemble des stratégies disponible au joueur i dans un jeu

\mathcal{A}_{-i} Ensemble des stratégies de tous les joueurs excepté i

\mathbf{a}_i Vecteur de stratégie pour l'utilisateur i

\mathbf{a}_{-i} Vecteur de stratégie pour tous les utilisateurs autres que i

\mathbf{a}_i^* Vecteur optimal de stratégie pour l'utilisateur i

u_i Fonction d'utilité du joueur i dans le jeu

Introduction générale

Les communications sans fil sous-marines revêtent une importance capitale pour un large éventail d'applications, allant de la recherche scientifique aux usages industriels et militaires. Étant donné que les ondes électromagnétiques ne se propagent pas efficacement dans l'eau, les ondes acoustiques sont le principal moyen de transmission sur de longues distances. Alors que l'eau de mer absorbe les ondes radio et lumineuses en quelques mètres, le son se propage sur des kilomètres, ce qui fait de la communication acoustique l'unique technologie fiable pour transmettre des informations sur de longues distances en milieu acoustique. Grâce à cette technologie, il devient possible de relier entre eux des capteurs, des véhicules autonomes ou des stations de contrôle en surface.

Dans le domaine scientifique, cette technologie est particulièrement précieuse. Elle s'avère essentielle pour le suivi en temps réel de paramètres clés tels que la température, la salinité, la pression et les courants marins. De plus, elle permet de piloter à distance des véhicules autonomes sous-marins (AUV) qui explorent les fonds océaniques, cartographient les reliefs sous-marins, identifient des zones d'intérêt écologique ou contribuent à la modélisation du climat. Concernant le domaine industriel, la communication acoustique est au cœur des opérations menées en mer : elle permet d'ouvrir ou fermer des vannes sous-marines sur des puits de pétrole, de surveiller l'état des pipelines ou de contrôler l'intégrité des infrastructures d'énergies renouvelables en mer. Elle est également essentielle pour les opérations d'inspection, de maintenance et de réparation, menées par des robots spécialisés. Pour le domaine de la défense et de la sécurité maritime, cette technologie joue un rôle stratégique. Elle permet d'assurer des communications discrètes entre sous-marins et de surveiller des zones sensibles.

Bien que la communication sous-marine soit devenue indispensable dans de nombreux domaines, le canal ASM a des caractéristiques physiques complexes, telles qu'une atténuation croissante avec la fréquence, un étalement dans le temps lié aux réflexions et aux réfractions, ainsi qu'une grande variabilité spatiale et temporelle. De plus, la faible vitesse de propagation des ondes acoustiques (1500 m/s) qui induit des délais de propagation longs de plusieurs secondes, voire dizaines de seconde et la mobilité des émetteurs/récepteurs induisent des effets Doppler significatifs.

Ces contraintes limitent fortement la bande passante disponible et dégradent les performances des systèmes de communication. Le spectre acoustique sous-marin est également soumis à un bruit important, d'origine anthropique (activités humaines) et biologique (faune marine), ce qui accentue les interférences et le masquage des signaux. Contrairement aux communications électromagnétiques terrestres, le spectre acoustique sous-marin est peu réglementé et il n'existe pas de structures de réseaux préétablies permettant d'organiser la cohabitation et le partage du canal de communication entre systèmes hétéro-

gènes, ce qui complique la coexistence de différentes sources acoustiques.

Dans ce contexte, marqué par une multiplication rapide de dispositifs autonomes sous-marins (sondes, modems, robots, capteurs, etc.) souvent déployés sans coordination, l'accès équitable au canal acoustique devient un enjeu majeur. L'absence de cadre global de régulation entraîne la concurrence entre systèmes, provoquant collisions, pertes d'information et gaspillage de ressources.

Le protocole JANUS, développé par l'Organisation du Traité de l'Atlantique Nord (OTAN), constitue une première tentative notable de normalisation. Fonctionnant autour de 11,5 kHz avec une bande passante de 900 Hz et un débit limité (80 bits/s), il permet l'échange de messages simples entre plateformes hétérogènes pour assurer interopérabilité, signalisation et coordination initiale [Potter et al., 2014, Petroccia and Alves, 2024]. Cependant, ses capacités limitées et l'absence de mécanismes dynamiques de gestion des ressources le rendent insuffisant pour résoudre les problèmes d'interférences et de partage adaptatif dans des réseaux denses ou dynamiques.

Face aux contraintes de l'environnement sous-marin, il est crucial de concevoir des mécanismes intelligents de partage des ressources. Cette thèse aborde donc la problématique sous l'angle de la Théorie des Jeux, en se concentrant sur les jeux coopératifs comme cadre formel pour modéliser et optimiser les interactions entre les différents systèmes autonomes partageant le même milieu. En effet, contrairement aux approches purement compétitives, les jeux coopératifs permettent de concevoir des mécanismes incitatifs où les agents ont intérêt à coopérer pour atteindre des performances collectives optimales. Combinée à des techniques issues de l'apprentissage multi-agents, cette approche permet aux dispositifs ASM de s'auto-configurer de manière décentralisée et intelligente, en apprenant à adapter leurs paramètres de transmission à la fois aux caractéristiques du canal et au comportement des autres utilisateurs. L'objectif est ainsi de proposer des stratégies robustes, autonomes et adaptatives pour le partage efficace du spectre acoustique dans des réseaux sous-marins dynamiques.

Ce cadre théorique ouvre la voie à des solutions réalistes et adaptées aux environnements sous-marins, où la coordination directe entre joueurs est souvent limitée par des contraintes physiques telles que la latence, le bruit ambiant ou la mobilité. Dans ce contexte, une stratégie de coopération peut néanmoins favoriser l'intérêt collectif. L'objectif est donc de développer des algorithmes et des mécanismes d'allocation des ressources (fréquences, puissance) qui soient à la fois performants, équitables et adaptés aux spécificités dynamiques du milieu sous-marin.

Plan du manuscrit et contributions :

Ce manuscrit s'organise autour de cinq chapitres. Le premier chapitre présente une vue d'ensemble sur la communication ASM, une technologie clé pour les échanges d'informations en milieu sous-marin, où les communications radio sont souvent inefficaces. Il met en lumière les défis spécifiques liés à l'allocation des ressources, tels que la gestion de la bande passante et l'optimisation de la puissance, dans un environnement caractérisé par une bande étroite, une forte latence et un canal instable.

Le deuxième chapitre introduit la théorie des jeux comme une approche particulièrement prometteuse pour modéliser et résoudre, de manière décentralisée et dynamique,

les problèmes d'allocation de ressources dans les systèmes de communication acoustique sous-marine. Ce cadre théorique permet l'analyse des interactions entre entités autonomes telles que les modems ou les capteurs sous-marins qui doivent prendre des décisions localement, souvent sans coordination explicite avec d'autres.

L'accent dans ce chapitre est sur les outils de la théorie des jeux, en particulier les jeux non coopératifs, où chaque joueur cherche à maximiser sa propre utilité dans un environnement conflictuel. Ces jeux permettent notamment de représenter des situations de compétition pour l'accès au canal acoustique, en tenant compte des contraintes spécifiques du milieu sous-marin. Différentes notions d'équilibre, telles que l'équilibre de Nash, sont explorées pour identifier des stratégies stables dans un contexte distribué.

Ce cadre servira de base à l'étude ultérieure, dans les chapitres suivants, des jeux coopératifs, l'axe principal de cette thèse, où la collaboration entre les utilisateurs est envisagée pour améliorer la performance globale du système. Ce chapitre couvre l'article suivant :

- K. Gharsalli, P. -J. Bouvet, S. Najeh, H. Besbes, T. L. Pors and H. Gazzah, "Non-cooperative Resource Allocation Game-Theoretic Approach in Underwater Acoustics," 2023 International Wireless Communications and Mobile Computing (IWCMC), Marrakesh, Morocco, 2023, pp. 632-637, doi : 10.1109/IWCMC58020.2023.10182682.

Le troisième chapitre, s'intéresse à l'application des jeux coopératifs et plus précisément les jeux de négociation coopératifs pour résoudre le problème d'allocation de ressources dans les réseaux [ASM](#) avec une comparaison aux approches non coopératives où chaque utilisateur agit pour optimiser son propre intérêt. Deux scénarios sont étudiés : le premier implique deux utilisateurs et permet d'analyser en détail les mécanismes de négociation entre eux ainsi que les équilibres possibles, tandis que le second étend le modèle à un contexte multi-utilisateur, plus proche des conditions réelles, où la coordination devient plus complexe pour améliorer la capacité d'adaptation à un environnement dynamique, une approche hybride est proposée en combinant les jeux coopératifs avec l'apprentissage par renforcement, notamment le Q-Learning ([QL](#)). Ce couplage permet aux agents d'apprendre à négocier de manière efficace sur le long terme, sans connaissance *a priori* complète du système, tout en conservant les avantages de la coopération. Ce chapitre comprend les articles :

- K. Gharsalli, T. L. Pors, S. Najeh, P. -J. Bouvet and H. Besbes, "Nash Bargaining Game for Cooperative Resource Allocation in Underwater Acoustic Communication," 2024 International Wireless Communications and Mobile Computing (IWCMC), Ayia Napa, Cyprus, 2024, pp. 1016-1021, doi : 10.1109/IWCMC61514.2024.10592489.
- K. Gharsalli, T. L. Pors, S. Najeh, L. NAJJAR, P. -J. Bouvet and H. Besbes, "Optimizing Underwater Acoustic Communication with Game Theory and Machine Learning," OCEANS 2025 conference, BREST, FRANCE.

Dans le quatrième chapitre, nous approfondissons l'approche coopérative en élargissant notre étude des jeux de négociation aux jeux de coalition. Cette évolution vise à mieux

répondre aux défis de scalabilité et d'équité rencontrés dans les systèmes [ASM](#). Bien que les jeux de négociation offrent des solutions efficaces pour coordonner de petits groupes de joueurs, leur complexité computationnelle et la charge de coordination croissante limitent leur applicabilité aux systèmes avec un grand nombre d'utilisateurs. Pour surmonter ces limitations, nous adoptons des jeux de coalition, qui permettent la formation de groupes stables et organisés d'agents coopérants. Ce cadre favorise une répartition équitable des ressources et une gestion collective du canal acoustique, tout en assurant une meilleure robustesse des stratégies à grande échelle.

Afin de permettre une prise de décision autonome et adaptative dans ce cadre coopératif, nous avons intégré des techniques d'apprentissage par renforcement, en particulier le [QL](#). Cette approche permet aux agents d'apprendre, par l'interaction avec l'environnement, les politiques optimales pour la formation de coalitions et le partage de ressources, sans connaissance a priori du modèle de canal ou des comportements des autres agents. [QL](#) facilite ainsi l'adaptation dynamique des stratégies dans un environnement acoustique sous-marin marqué par l'incertitude, la variabilité temporelle et le bruit.

Ce chapitre présente ainsi les fondements théoriques des jeux de coalition, leur formalisation dans le contexte des réseaux acoustiques sous-marins, ainsi que leur mise en œuvre algorithmique à travers l'apprentissage par renforcement. Il s'appuie sur les travaux publiés dans l'article suivant :

- K. Gharsalli, S. Najeh, T. L. Pors, L. Najjar, P. -J. Bouvet and H. Besbes, "Coalition Game-Based Reinforcement Learning in Underwater Acoustic Networks," in *IEEE Access*, 2025, doi : 10.1109/ACCESS.2025.3582918.

Le chapitre 5 présente les résultats expérimentaux obtenus à partir de l'approche basée sur les jeux de coalition développée dans le chapitre 4. Il vise à évaluer l'efficacité de cette méthode dans différents scénarios de communication acoustique sous-marine, en la comparant aux autres approches abordées au même chapitre.

Les performances observées sont ensuite comparées à celles des jeux non-coopératifs et aux jeux de négociation étudiés dans le chapitre 3, afin de mettre en évidence les avantages et les limites de chaque méthode. Enfin, une analyse de la complexité de l'algorithme proposé est réalisée, en vue d'évaluer sa scalabilité et son applicabilité dans des contextes à plus grande échelle.

Ce manuscrit se termine par une conclusion générale, qui résume les principales contributions de ce travail, en particulier l'application de la théorie des jeux et de l'apprentissage par renforcement au partage des ressources dans les communications acoustiques sous-marines. Elle ouvre également plusieurs perspectives.

Chapitre 1

Communications acoustiques sous-marines (CASM)

1.1 Introduction

L'objectif d'un système de communication numérique est de transmettre de l'information depuis une source (l'émetteur) vers une destination (le récepteur), sous forme de symboles ou de bits, à un certain débit, généralement exprimé en symboles par seconde ou en bits par seconde. Cette transmission s'effectue à travers un milieu physique qui impose des lois de propagation spécifiques aux ondes utilisées pour transporter l'information. Ces ondes, ou signaux de communication, véhiculent les symboles d'information sur une bande de fréquence donnée, centrée autour d'une ou plusieurs fréquences porteuses.

Ce milieu de transmission constitue ce que l'on appelle un canal de communication, chaque canal possède une capacité maximale, qui représente le nombre de symboles (ou de bits) pouvant être transmis par seconde avec un taux d'erreur arbitrairement faible [Shannon, 1949], cette capacité formalise le compromis entre débit et fiabilité, deux critères fondamentaux de performance d'un système de communication et ce compromis est lié aux propriétés physiques du canal de transmission, lesquelles varient fortement en fonction de l'environnement. Alors que les communications terrestres reposent sur des canaux électromagnétiques bien maîtrisés, le contexte sous-marin impose des contraintes bien différentes. En raison de la forte absorption des ondes électromagnétiques par l'eau, celles-ci deviennent inefficaces au-delà de courtes distances. Cela rend nécessaire le recours aux ondes acoustiques, seul moyen efficace de propagation sur de longues distances dans le milieu marin.

Les océans couvrent plus de 70% de la surface de la Terre, ce qui en fait une ressource cruciale pour l'environnement et l'économie mondiale [Stojanovic and Beaujean, 2016]. Dans ce contexte, l'exploration, l'observation et l'exploitation des milieux sous-marins sont devenues des enjeux stratégiques majeurs dans les domaines scientifique, économique et militaire et la capacité à établir des communications fiables et efficaces sous l'eau est ainsi devenue essentielle. Les systèmes de communication ASM permettent d'assurer la transmission d'informations entre dispositifs immergés, ou entre engins sous-marins et stations de surface, ils jouent un rôle central dans de nombreuses applications, telles que les AUV, les capteurs environnementaux, les réseaux de capteurs marins et les systèmes de sécurité maritime [Wynn et al., 2014]–[Grindheim, 2018].

Toutefois, ce mode de transmission est soumis à de nombreuses contraintes physiques : une vitesse de propagation lente (environ 1500 m/s), des phénomènes de multi-trajets (dus aux réfractions et aux réflexions), une variabilité spatio-temporelle du canal, ainsi qu'une bande passante limitée, souvent inférieure à quelques dizaines de kilohertz [Akyildiz et al., 2004]. À ces défis techniques s'ajoute un bruit ambiant élevé, provenant à la fois d'activités humaines (navigation, extraction offshore, infrastructures industrielles) et de sources biologiques (mammifères marins, organismes benthiques, etc.).

Ce bruit ambiant est aujourd'hui reconnu comme une forme de pollution acoustique : il interfère non seulement avec les communications, mais aussi avec les écosystèmes marins. L'Union Européenne a ainsi intégré le bruit sous-marin comme descripteur environnemental (D11) dans sa Directive-Cadre Stratégie pour le Milieu Marin. Malgré une décennie de recherches intensives, l'établissement de normes quantitatives permettant de définir un bon état écologique reste un défi scientifique et politique non résolu, illustrant l'absence de standards harmonisés dans ce domaine [Catipovic, 2002, Merchant et al., 2022].

Cette absence de cadre normatif accentue le caractère instable et conflictuel du spectre acoustique marin, qui apparaît désormais comme une ressource rare, partagée entre usages technologiques et exigences environnementales. Contrairement aux communications hertziennes, régies par des standards internationaux bien établis (3G, 4G, 5G, Wi-Fi, etc.), les communications acoustiques sous-marines évoluent dans un cadre encore immature. Cela renforce l'importance d'analyser leurs spécificités techniques et réglementaires afin de mieux comprendre les enjeux liés à leur conception et à leur déploiement.

Afin de mieux comprendre les spécificités des communications acoustiques sous-marines, il est pertinent de les comparer aux systèmes de communication sans fil classiques comme présenté dans la section suivante.

1.2 Communications sans fil

La communication sans fil désigne tout échange d'informations réalisé sans support physique, comme des câbles, en s'appuyant sur la transmission d'ondes pour transférer les données d'un point à un autre. Selon les exigences du contexte telles que la portée, le débit, la consommation énergétique ou les conditions du canal, différents types de technologies sans fil peuvent être mis en œuvre, notamment les communications radio, optiques ou acoustiques. Chacune repose sur une nature d'onde spécifique et présente des propriétés distinctes en matière de performance et d'adaptabilité à l'environnement.

1. **Communication radio** : Ce type de communication repose sur l'utilisation d'ondes électromagnétiques dans les bandes Very High Frequency (**VHF**) (30–300 MHz), Ultra High Frequency (**UHF**) (300 MHz–3 GHz) et micro-ondes (>3 GHz) pour transmettre des données. Elle est largement exploitée dans les réseaux terrestres et aériens tels que la téléphonie mobile (3G, 4G, 5G), le Wi-Fi, les systèmes IoT, ou encore les communications aéronautiques, en raison de sa large bande passante, de sa faible latence, de la maturité de ses équipements, et de sa bonne portée en

environnement dégagé. Toutefois, en milieu sous-marin, cette technologie devient inopérante : les ondes électromagnétiques sont rapidement absorbées par l'eau, notamment salée, où leur pénétration est limitée à quelques centimètres à haute fréquence, et nécessite l'utilisation de très basses fréquences (inférieures à 30 kHz) pour une portée de quelques mètres seulement, au prix d'un débit extrêmement faible [Che et al., 2010]. Bien que la communication radio soit généralement limitée dans les environnements sous-marins en raison de l'absorption ionique et des pertes d'énergie importantes qui réduisent considérablement sa portée, elle n'est pas totalement inadaptée. En effet, elle peut s'avérer pertinente dans certains cas spécifiques, notamment pour des communications à haut débit sans contact sur de très courtes distances (quelques centimètres), ou pour des transmissions à très bas débit sur de longues distances.

2. **Communication optique sans fil** : Repose sur l'utilisation de la lumière visible, infrarouge ou laser pour transmettre des données à très haut débit, avec une latence faible et une sécurité accrue grâce à la directivité du faisceau lumineux. Elle est utilisée dans divers contextes, tels que les liaisons point-à-point terrestres, les réseaux Li-Fi, ou encore les communications sous-marines à courte portée [Schirripa Spagnolo et al., 2020]. Toutefois, cette technologie présente plusieurs limitations majeures, elle est particulièrement sensible à l'alignement précis entre l'émetteur et le récepteur, ainsi qu'aux obstacles physiques, aux conditions du milieu de transmission et à la lumière ambiante, qui peut perturber la réception des signaux optiques. En milieu marin, des facteurs tels que la turbidité de l'eau, la diffusion, l'absorption lumineuse et la présence de particules en suspension réduisent significativement la portée utile, qui généralement limitée à quelques mètres, mais peut s'étendre jusqu'à quelques dizaines de mètres dans des conditions exceptionnellement favorables, tout en maintenant un débit supérieur à celui des communications acoustiques. Néanmoins, malgré ses performances élevées en termes de débit et de sécurité, la communication optique reste difficile à déployer dans des environnements dynamiques, non maîtrisés ou fortement perturbés, notamment en milieu sous-marin [Shherbakov et al., 2020].
3. **Communication acoustique** : Ce type utilise les ondes sonores, particulièrement adaptée aux milieux où les ondes électromagnétiques ou optiques sont atténuées, comme le milieu sous-marin. Elle permet d'atteindre des distances de communication allant de plusieurs centaines de mètres à quelques kilomètres, selon la fréquence utilisée. Toutefois, elle souffre de limitations importantes : faible bande passante (donc débit réduit), latence de propagation élevée (environ 1500 m/s dans l'eau), interférences dues aux réflexions, au bruit ambiant ou aux effets de canal variables [Yang, 2011]-[Zhou et al., 2019]. Cela rend la conception des protocoles de communication plus complexe et nécessite une optimisation fine des ressources.

Une comparaison de ces technologies de communication sans-fil est présentée par le tableau suivant :

Critère	Radio	Optique	Acoustique
Milieu idéal	Air, vide	Air, eau claire (courte portée)	Eau
Portée typique	10 m – plusieurs centaines de Kms	1 – 100 m	100 m – plusieurs cen- taines de Kms
Débit	Moyen à élevé	Très élevé	Faible à moyen
Latence	Faible	Très faible	Élevée
Sensibilité aux obs- tacles	Moyenne	Élevée	Moyenne
Robustesse environne- mentale	Bonne	Faible	Variable
Consommation éner- gétique	Moyenne	Élevée	Moyenne
Applications typiques	Réseaux IoT, Wi-Fi, 5G	VLC, Li-Fi, communication inter-véhicules	Réseaux sous-marins, sonar, capteurs océa- nographiques

TABLE 1.1 – Comparaison des technologies de communication sans-fil

En résumé, le choix de la technologie de communication sans-fil dépend fortement de l’environnement physique, de la distance de transmission ainsi que des contraintes liées au débit et à la consommation énergétique [Ali and Miraz, 2020]. Dans un environnement aérien, les communications radio et optiques sont prédominantes, tandis que dans les milieux aquatiques, seule la communication acoustique permet une couverture raisonnable à moyenne ou longue distance [Mendez and James, 2015]. Dans ce contexte, afin de mieux comprendre les particularités de la communication acoustique en milieu sous-marin, il est essentiel d’analyser en détail les caractéristiques physiques du canal acoustique ainsi que les défis qu’il impose aux systèmes de transmission.

1.3 Canal acoustique sous-marin : caractéristiques et défis

Le milieu marin présente des particularités physiques qui influencent fortement la propagation des ondes acoustiques et déterminent en grande partie les performances des communications sous-marines [Lurton, 1998]. Contrairement à l’air ou au vide, l’environnement marin est dense, hétérogène et variable, où la propagation du son est soumise à plusieurs phénomènes complexes.

1.3.1 Propriétés physiques du milieu de propagation

1. **Variabilité spatio-temporelle** : La vitesse du son dans l’eau dépend fortement de paramètres environnementaux tels que la température T en degré Celsius, S est la salinité en ‰ et z est la profondeur en mètres. Ces paramètres varient dans le

temps (saisons, marées, cycles jour/nuit) et dans l'espace (profondeur, localisation géographique), ce qui entraîne une variabilité continue des conditions de propagation.

Une approximation empirique couramment utilisée pour modéliser la vitesse du son en fonction de ces paramètres est donnée par [Medwin, 1975] :

$$c(T, S, z) = 449.2 + 4.6T - 0.055T^2 + 0.00029T^3 + (1.34 - 0.01T)(S - 35) + 0.016z. \quad (1.1)$$

Cette équation montre la non-linéarité du phénomène et la forte sensibilité du canal acoustique à de faibles variations environnementales. En général, la célérité de l'onde acoustique dans le milieu sous-marin est comprise entre 1440 et 1540 m/s et elle varie avec la profondeur selon un profil influencé par la température, la salinité et la pression.

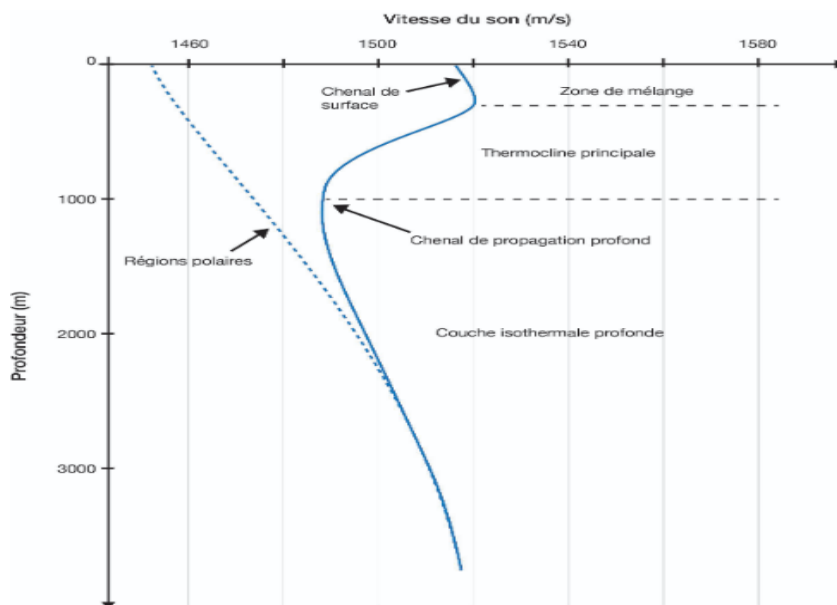


FIGURE 1.1 – Profil de célérité en océan profond [Chauvaud et al., 2018].

Par la figure 1.1 on peut distinguer qu'en surface, une zone de mélange est présente, brassée par le vent et les vagues, où la température reste relativement constante, entraînant une légère augmentation de la vitesse du son avec la profondeur. En période chaude, un chenal de surface peut apparaître, causé par le réchauffement solaire, qui augmente localement la célérité.

En dessous, on observe une thermocline principale, où la température chute rapidement, ce qui provoque une diminution marquée de la célérité malgré la pression croissante, cette zone conduit à la formation d'un chenal de propagation profond, où le son peut se propager sur de longues distances.

Enfin, dans la couche isotherme profonde, la température devient stable, et la célérité augmente de façon quasi linéaire sous l'effet de la pression. Dans les régions

polaires, où l'eau est froide de la surface jusqu'aux grandes profondeurs, la célérité décroît continuellement, comme indiqué par la courbe en pointillés sur la figure. Ce profil illustre ainsi la structure verticale typique de la propagation acoustique dans l'océan [Chauvaud et al., 2018].

2. **Canal à trajets multiples** : Le signal acoustique sous-marin est soumis à de multiples réflexions sur la surface de l'eau, le fond marin, ainsi que sur d'éventuelles couches de discontinuité dans la colonne d'eau. Ces phénomènes engendrent un canal à réponse impulsionnelle étendue, souvent caractérisé par plusieurs trajets (paths) distincts [Khan et al., 2014].

Mathématiquement, la réponse impulsionnelle du canal peut s'écrire comme la somme des contributions de chaque trajet i :

$$h(t) = \sum_{i=1}^N a_i \delta(t - \tau_i), \quad (1.2)$$

où :

- N est le nombre de trajets multiples,
- a_i représente l'amplitude complexe (intensité et phase) associée au i -ième trajet,
- τ_i est le délai de propagation relatif du i -ième trajet,
- $\delta(\cdot)$ est la fonction delta de Dirac.

Chaque trajet arrive à un instant différent, avec une amplitude et une phase variables, ce qui induit une dispersion temporelle (delay spread) importante.

En modélisant la statistique de ces amplitudes complexes, la distribution de la magnitude $|a_i|$ est souvent caractérisée par :

- **Distribution de Rayleigh** : Si aucun trajet direct dominant n'existe, ce qui correspond à un environnement fortement diffus.
- **Distribution de Rice** : Si un trajet direct ou quasi-direct est prédominant, avec un facteur de Rice K définissant le rapport de puissance entre le trajet direct et les trajets diffus.

La fonction d'autocorrélation temporelle du canal $R_h(\Delta t)$ peut alors être exprimée en fonction des retards et amplitudes, ce qui permet de caractériser la cohérence temporelle du canal.

Cette dispersion spatiale et temporelle des trajets impacte fortement la conception des systèmes de communication sous-marins, nécessitant des techniques adaptées telles que l'égalisation temporelle et la diversité spatiale.

3. **Pertes de propagation** : Les pertes de propagation subies par un signal se propageant dans un canal acoustique sous-marin dépendent à la fois de la fréquence f (en kHz) et de la distance de transmission d (en mètres). Ces pertes d'atténuation résultent de la diffusion géométrique et de l'absorption, et peuvent être exprimées comme suit [Lurton, 2002]-[Pan et al., 2012] :

$$L_{dB}(d, f) = k \cdot 10 \log_{10}(d) + 10^{-3} \cdot d \cdot \alpha_{dB}(f), \quad (1.3)$$

où le facteur de diffusion $k \in \{1, 2\}$ décrit la géométrie de propagation : $k = 1$ correspond à une diffusion cylindrique, typique des environnements peu profonds, tandis que $k = 2$ correspond à une diffusion sphérique, lorsque le signal se propage uniformément dans un milieu océanique ouvert. Le terme $\alpha_{dB}(f)$ représente le coefficient d'atténuation dépendant de la fréquence, exprimé en dB par kilomètre.

Plusieurs modèles existent pour ce coefficient, mais le plus couramment utilisé dans les communications acoustiques sous-marines est donné par la formule de Thorp [Lurton, 1998], qui suppose une propagation à trajet unique (sans effets de multi-trajets) et est généralement valide pour des fréquences supérieures à quelques centaines de hertz :

$$\alpha_{dB}(f) \approx \frac{0.11 \cdot f^2}{f^2 + 1} + \frac{44 \cdot f^2}{f^2 + 4100} + 2.7 \cdot 10^{-4} \cdot f^2 + 0.003, \quad (1.4)$$

où f désigne la fréquence (en kHz).

La perte d'atténuation est illustrée à la Figure 1.2, qui montre le comportement du canal en fonction de la fréquence, lorsque la fréquence varie de 0 à 100 kHz, la perte par atténuation augmente progressivement. Les basses fréquences présentent une atténuation minimale, tandis que les hautes fréquences subissent des pertes plus importantes dues à l'absorption et à la diffusion. Cette atténuation sélective en fréquence affecte fortement la propagation du signal et doit être prise en compte dans la modélisation du canal.

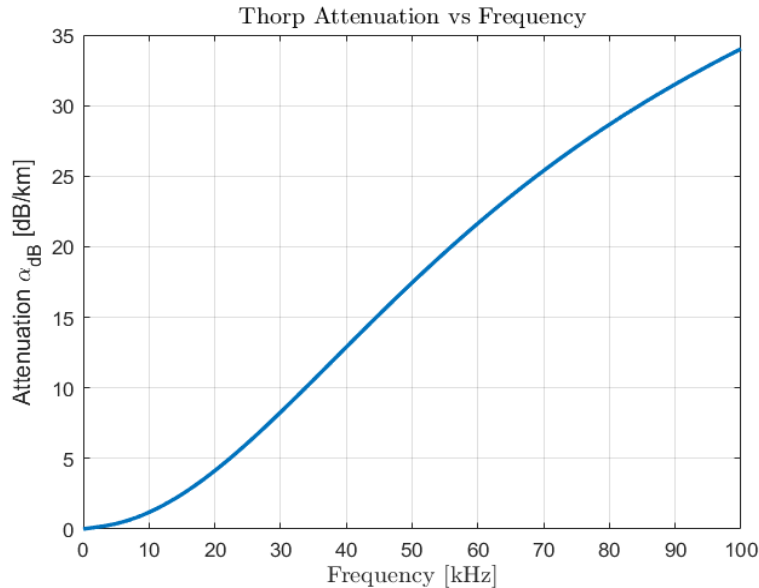


FIGURE 1.2 – Coefficient d'atténuation en fonction de la fréquence.

La réponse fréquentielle du canal acoustique peut être modélisée alors par la fonction de transfert suivante [Qarabaqi and Stojanovic, 2013] :

$$H(f, t) = \bar{H}_0(f) \sum_p \bar{h}_p \zeta_p(f, t) e^{-j2\pi f \tau_p(t)}, \quad (1.5)$$

où $\bar{H}_0(f)$ est la fonction de transfert du trajet principal, \bar{h}_p représente l'atténuation relative, $\zeta_p(f, t)$ le coefficient de diffusion, et $\tau_p(t)$ le retard temporel relatif associé au trajet p .

4. **Sources de bruit dans le canal acoustique sous-marin** : Le canal acoustique sous-marin est fortement influencé par différents types de bruits ambiants, qui dégradent la qualité de la transmission et limitent les performances des systèmes de communication. Parmi ces bruits, on distingue principalement : le bruit thermique, le bruit dû aux turbulences, le bruit généré par les vagues, et celui dû au trafic maritime [Pan et al., 2012].

Chaque type de bruit est dominant dans une certaine bande de fréquence et peut être modélisé à l'aide de densités spectrales de puissance (DSP) empiriques, exprimées en décibels par rapport à une pression de référence de $1 \mu Pa$ à un mètre de la source. Ces modèles sont issus de travaux classiques tels que ceux de [Knudsen et al., 1948]-[Coates and Coates, 1990], ce dernier étant particulièrement utilisé dans le domaine des communications acoustiques sous-marines.

Les expressions mathématiques des différentes contributions sont exprimées par le tableau suivant (fréquence f en kHz) :

Contribution	Expression mathématique
Turbulences	$N_T(f) = 17 - 30 \log_{10}(f)$
Trafic maritime	$N_S(f) = 40 + 20(S - 0.5) + 26 \log_{10}(f) - 60 \log_{10}(f + 0.03)$
Vagues	$N_W(f) = 50 + 7.5\sqrt{W} + 20 \log_{10}(f) - 40 \log_{10}(f + 0.4)$
Bruit thermique	$N_C(f) = -15 + 20 \log_{10}(f)$

TABLE 1.2 – Modèles empiriques des densités spectrales de bruit sous-marin.

où :

- f est la fréquence (en kHz),
- $S \in [0, 1]$ est un facteur décrivant l'intensité du trafic maritime,
- W est la vitesse du vent (en m/s).

Dans ce cadre, la densité spectrale totale de bruit peut être modélisée comme la somme des contributions précédentes :

$$N_{\text{total}}(f) = N_T(f) + N_S(f) + N_W(f) + N_C(f). \quad (1.6)$$

Ce bruit ambiant est donc dépendant de la fréquence et des conditions environnementales locales, et constitue un facteur limitant fondamental pour la conception des modems acoustiques et des stratégies de modulation/adaptation de bande passante dans les systèmes de communication acoustique sous-marin.

5. **Mobilité et Doppler** : La mobilité relative entre l'émetteur, le récepteur et le milieu (courants marins) introduit des effets Doppler significatifs. Cela entraîne un élargissement Doppler, voire un décalage fréquentiel global, qui affecte la synchronisation et la démodulation. Statistiquement, cela modifie la structure de cohérence du canal dans le temps et la fréquence [Zhang et al., 2013]-[Ahmad et al., 2018].

Dans le cas d'un canal bande étroite, où la fréquence porteuse f_c est très supérieure à la largeur de bande du signal B , l'effet Doppler peut être approximé par un simple décalage fréquentiel :

$$f_r = f_c \left(1 + \frac{v}{c}\right). \quad (1.7)$$

Ce qui n'affecte pas significativement le spectre du signal.

Cependant, dans un canal large bande, cette modélisation n'est plus valable. Le Doppler n'induit pas seulement un décalage en fréquence, mais une transformation plus complexe du signal, sous la forme d'une variation de l'échelle temporelle (compression ou étirement), affectant l'ensemble du spectre.

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

où $s(t)$ est le signal émis, $r(t)$ le signal reçu, et $\Delta = \frac{v}{c}$ est le facteur Doppler, défini comme le rapport entre la vitesse relative v entre source et récepteur et la célérité c du son dans l'eau.

L'effet de cette transformation dans le domaine fréquentiel est donné par :

$$\mathcal{F}\{r(t)\} = \frac{1}{|1 - \Delta|} S\left(\frac{f}{1 - \Delta}\right). \quad (1.9)$$

Ce qui implique à la fois un décalage fréquentiel, une compression ou expansion spectrale, ainsi qu'une modification d'amplitude spectrale.

Si le signal original $s(t)$ occupe une bande passante B autour d'une fréquence centrale f_c , alors après effet Doppler, le signal reçu occupera une bande élargie (ou réduite) donnée par :

$$B' = B(1 - \Delta), \quad (1.10)$$

où B' est centrée autour de la nouvelle fréquence $f'_c = (1 - \Delta)f_c$.

Cet effet est particulièrement problématique en environnement mobile ou fortement perturbé, car il affecte la synchronisation, la détection, et nécessite des méthodes de correction avancées (recalage temporel, estimation du facteur Doppler, synchronisation multi-échelle, etc.).

6. **Réponse impulsionnelle du canal ASM** : La réponse impulsionnelle d'un canal acoustique sous-marin notée $h(t, \tau)$ est une fonction fondamentale qui décrit comment un signal acoustique se propage entre un émetteur et un récepteur à travers le milieu sous-marin. Elle indique comment le canal modifie un signal envoyé à un instant donné, en tenant compte de l'écho, de la diffraction, des réflexions

multiples et elle est exprimée en fonction des retards τ et du temps t et définie par cette équation :

$$h(t, \tau) = \sum_{k=1}^K c_k(t) g_k(\tau - \tau_k(t)), \quad (1.11)$$

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.

Ces fluctuations temporelles du canal, représentées par la variation des coefficients $c_k(t)$ et des retards $\tau_k(t)$, traduisent l'évolution dynamique du milieu de propagation (mouvement des capteurs, turbulence, variations de température ou de salinité, etc.). Elles sont liées à la fonction de diffusion du canal présentée par l'équation 1.12, une représentation conjointe en retard τ et en fréquence Doppler, qui décrit la distribution de l'énergie du canal dans le domaine temps-fréquence. Un canal très variable dans le temps se manifeste par un étalement Doppler plus large, ce qui limite le temps de cohérence du canal défini dans la partie suivante, c'est-à-dire la durée pendant laquelle le canal peut être considéré comme invariant [Bello, 1963].

$$S_h(\tau, f) = \int R_h(\tau, u) e^{-2j\pi f u} du, \quad (1.12)$$

où

$$R_h(\tau, u) = \mathbb{E}[(h(\tau, t) - \mathbb{E}[h(\tau, t)])(h(\tau, t + u) - \mathbb{E}[h(\tau, t + u)])^*]. \quad (1.13)$$

Cette fonction de diffusion permet ainsi de caractériser à la fois l'étalement temporel et l'étalement Doppler, deux phénomènes responsables des Inter Symbol Interference (ISI) et Inter Carrier Interference (ICI) respectivement.

7. **Temps et bande de cohérence** : La qualité d'un lien acoustique sous-marin dépend fortement des caractéristiques dynamiques du canal, en particulier de sa dispersion temporelle et fréquentielle, qui se traduisent respectivement par la bande de cohérence et le temps de cohérence. La dispersion temporelle, liée aux multiples trajets empruntés par le signal (réflexions sur le fond, la surface, etc.), provoque des ISI. Celles-ci deviennent plus sévères lorsque la durée de l'étalement temporel dépasse celle d'un symbole, ce qui est souvent le cas dans les canaux sous-marins peu profonds où l'étalement peut atteindre plusieurs dizaines voire centaines de millisecondes.

Le canal peut alors être modélisé à partir de son profil de retard-puissance, qui permet de calculer la dispersion temporelle. La bande de cohérence, définie comme $B_C = \frac{1}{\tau_{rms}}$, donne une estimation de la largeur de bande au sein de laquelle le canal peut être considéré comme non sélectif en fréquence. Typiquement, cette bande est très étroite de l'ordre de quelques dizaines à la centaine de Hz ce qui pose un défi majeur pour la transmission large bande [van Walree, 2011].

Étant donné que l'étalement temporel du canal peut se définir comme suit :

$$\tau_{rms} = \sqrt{\frac{\int (\tau - \tau_g)^2 S_d(\tau) d\tau}{\int S_d(\tau) d\tau}}, \quad (1.14)$$

où $S_d(\tau)$ la dispersion temporelle du canal qui est défini par l'équation suivante :

$$S_d(\tau) = \int |S_h(\tau, f)|^2 df, \quad (1.15)$$

avec $S_h(\tau, f)$ est la fonction de diffusion présentée par l'équation 1.12.

La bande de cohérence, quant à elle, désigne l'intervalle fréquentiel sur lequel la réponse du canal peut être considérée comme plate, c'est-à-dire sans variation significative du gain ni de la phase. Si la largeur spectrale du signal dépasse cette bande, le canal devient sélectif en fréquence, ce qui engendre des **ISI**. Ces deux paramètres dépendent fortement des conditions physiques du milieu marin, telles que la vitesse relative entre l'émetteur et le récepteur, la profondeur, la température, ou encore la présence de turbulences et de réflexions multiples.

On peut définir l'étalement fréquentiel alors par :

$$D_{rms} = \sqrt{\frac{\int (f - f_g)^2 S_{\Delta}(f) df}{\int S_{\Delta}(f) df}}, \quad (1.16)$$

où $S_{\Delta}(f)$ caractérise la dispersion fréquentielle du canal

$$S_{\Delta}(f) = \int |S_h(\tau, f)|^2 d\tau. \quad (1.17)$$

On définit alors le temps de cohérence par :

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

Ce temps de cohérence présente la durée au-delà de laquelle les variations du canal deviennent significatives, rendant le canal non stationnaire dans le temps.

1.3.2 Contraintes pour les communications **ASM**

Dans le contexte des communications acoustiques sous-marines, les propriétés physiques du canal introduisent plusieurs contraintes critiques qui doivent être prises en compte lors de la conception des systèmes de communication, en particulier pour l'allocation efficace des ressources (fréquence, puissance, temps).

Ces contraintes incluent :

- **Bande passante fortement limitée** : En raison de l'atténuation croissante avec la fréquence, la bande utile disponible pour les communications acoustiques est extrêmement étroite, généralement limitée à quelques dizaines de kHz. Cette contrainte impose une planification spectrale rigoureuse et favorise l'utilisation de technologies comme l'**OFDM** pour exploiter au mieux le spectre disponible.

- **Sélectivité temporelle et fréquentielle** : La combinaison du multi-trajet et de la dispersion fréquentielle génère un canal fortement sélectif en fréquence et en temps. Cela rend la communication vulnérable à l'**ISI** et **ICI**, nécessitant des techniques robustes d'égalisation et de synchronisation.
- **Risque élevé d'interférences et de collisions** : La rareté des ressources en bande passante, combinée à l'absence de coordination centralisée dans de nombreuses configurations de réseaux sous-marins (réseaux ad hoc, capteurs autonomes), augmente le risque d'accès concurrentiel à la même sous-porteuse ou au même canal, ce qui engendre des collisions et de fortes interférences.
- **Variabilité rapide et imprévisible du canal** : Les mouvements de l'eau (vagues, courants) et les changements de température/salinité provoquent des fluctuations rapides. Cela rend toute stratégie d'allocation statique totalement inefficace, nécessitant des schémas dynamiques et adaptatifs.
- **Contraintes énergétiques strictes** : Les équipements sous-marins sont généralement alimentés par batterie, avec une capacité énergétique limitée difficilement rechargeable. Toute stratégie de gestion de ressources doit donc minimiser la consommation d'énergie tout en maintenant un niveau de performance acceptable.
- **Latence importante et synchronisation difficile** : La lenteur de propagation des ondes acoustiques (environ 1500 m/s) entraîne des délais de transmission importants, compliquant la synchronisation entre nœuds et limitant l'efficacité des protocoles d'accès classiques, tels que **TDMA**, **FDMA**, **CDMA**, **ALOHA**, **CSMA** ou **OFDMA**.

Dans ce contexte, l'accès au canal par plusieurs utilisateurs devient un enjeu fondamental, nécessitant des stratégies adaptées pour partager les ressources efficacement.

1.4 Partage des ressources : Problématique de l'accès multiple

1.4.1 État de l'art des protocoles d'accès multiple pour les réseaux acoustiques sous-marins

La gestion efficace du spectre dans un réseau acoustique sous-marin repose fondamentalement sur la manière dont les ressources (temps, fréquence, puissance) sont partagées entre plusieurs utilisateurs. Cette problématique est d'autant plus complexe que le canal acoustique est sévèrement contraint en termes de bande passante, de robustesse, et de disponibilité. L'accès multiple vise à permettre à plusieurs nœuds de communiquer sans provoquer d'interférences destructrices, tout en maximisant l'utilisation du spectre.

Les schémas d'accès multiple traditionnels, conçus pour les réseaux radio-fréquence, s'avèrent souvent inefficaces lorsqu'ils sont appliqués directement en **ASM**. Ils constituent cependant la base sur laquelle une littérature abondante a construit des protocoles adaptés. Nous classons ces approches en deux grandes familles : les protocoles basés sur la

contention (accès aléatoire) et ceux basés sur la réservation (sans contention), une brève définition de leurs principes est d’abord présentée, suivie d’un tableau comparatif mettant en évidence leurs avantages et leurs limites.

Protocoles basés sur la réservation (Sans contention) :

- **FDMA** : Ce schéma alloue une bande de fréquence statique à chaque utilisateur, son application en **ASM** est très limitée [Myung et al., 2007]. D’une part, la bande passante totale est déjà extrêmement faible. D’autre part, l’atténuation variant fortement avec la fréquence et l’effet Doppler sévère (imposant des bandes de garde larges) rendent cette division statique du spectre très inefficace.
- **TDMA** : Dans ce schéma, chaque utilisateur se voit attribuer un slot temporel fixe [Lam, 2003]. Si sa simplicité est un atout, le défi majeur en **ASM** est la synchronisation stricte requise. Les délais de propagation élevés, variables et inconnus (dans les réseaux ad hoc) nécessitent des temps de garde très importants entre les slots, gaspillant la ressource.

Pour répondre à ce défi fondamental, la littérature a exploré des allocations de slots qui contournent le besoin d’une horloge centralisée. L’exemple le plus notable est TDA-MAC (Transmit Delay Allocation MAC), développé à l’Université de York [Morozs et al., 2017]. Ce protocole parvient à organiser un accès temporel **TDMA** sans synchronisation d’horloge en allouant des délais de transmission relatifs. Les simulations montrent que cette approche rivalise avec les performances d’un **TDMA** synchronisé idéal et surpasse largement les protocoles à contention. Toutefois, bien que TDA-MAC résolve brillamment le problème de la synchronisation, il reste une approche basée sur la réservation. Il est très efficace pour un trafic constant (monitoring), mais il demeure peu flexible face à un trafic très sporadique ou des topologies ad hoc hautement dynamiques, où les slots peuvent rester inutilisés.

Protocoles basés sur la contention (Accès aléatoire) : Ces protocoles sont attractifs pour les réseaux ad hoc et les capteurs autonomes en raison de leur simplicité et de leur nature décentralisée.

- **ALOHA** : Protocole historique [Abramson, 2003] où les nœuds transmettent sans coordination. En **ASM**, les longs délais de propagation augmentent drastiquement la période de vulnérabilité du paquet, menant à des collisions fréquentes et une efficacité très faible. Des variantes comme Slotted-ALOHA (S-ALOHA) ont été étudiées, mais elles imposent une synchronisation temporelle globale difficile à maintenir.
- **CSMA** : Protocole, où un nœud écoute le canal avant d’émettre, réduit les collisions dans les réseaux Radio-Fréquence [Hwang et al., 2006]. Cependant, en **ASM**, le carrier sense est fondamentalement peu fiable : en raison des longs délais, l’information d’écoute du canal est déjà obsolète au moment où elle est reçue. Un nœud peut détecter un canal libre alors qu’un nœud distant a déjà commencé à émettre.

Pour contourner ce problème, une part importante de la littérature s'est concentrée sur des protocoles de réservation par poignée de main (handshake), souvent basés sur des échanges RTS/CTS (Request-to-Send/Clear-to-Send). Des protocoles comme FAMA (Floor Acquisition Multiple Access) [Molins and Stojanovic, 2006] ou MACA-U [Ng et al., 2008] tentent ainsi d'acquiescer le plancher (réserver le canal) avant de transmettre les données. Si elles gèrent mieux le problème de la station cachée, ces méthodes introduisent une surcharge (overhead) significative et une latence supplémentaire, ce qui est rédhibitoire pour des canaux à si faible débit.

Approches hybrides et avancées :

- **CDMA** : En attribuant un code pseudo-aléatoire unique à chaque utilisateur [Viterbi, 1995], le CDMA permet une transmission simultanée sur tout le spectre. Il est théoriquement robuste au bruit. Cependant, sa performance en ASM est sévèrement dégradée par deux facteurs : la perte d'orthogonalité des codes due aux multi-trajets et au Doppler, et une sensibilité extrême à l'effet "near-far" (un émetteur proche éclipsant un émetteur lointain), qui est particulièrement prononcé en ASM et nécessite une gestion de puissance complexe [Kang, 2021].
- **OFDMA** : Une extension multi-utilisateur de la modulation OFDM, dans laquelle la bande de fréquence est divisée en un ensemble de sous-porteuses orthogonales réparties dynamiquement entre les utilisateurs [Tu et al., 2011]. Contrairement au FDMA, les sous-porteuses peuvent être étroitement espacées sans interférer, ce qui améliore l'efficacité spectrale. De plus, l'attribution flexible des ressources en temps-fréquence permet d'adapter la transmission à la qualité du canal. Cependant, dans les environnements acoustiques sous-marins, la performance de l'OFDMA dépend fortement du maintien de l'orthogonalité entre sous-porteuses, ce qui nécessite des mécanismes de synchronisation et de compensation du décalage Doppler robustes.

Afin de mieux visualiser les principes décrits, les figures suivantes illustrent les différents schémas d'accès multiple (TDMA, FDMA, CDMA, ALOHA, CSMA) appliqués au canal acoustique sous-marin.

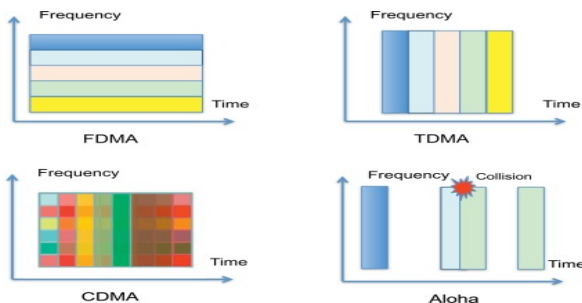


FIGURE 1.3 – Représentation schématique du TDMA, FDMA, CDMA et ALOHA.

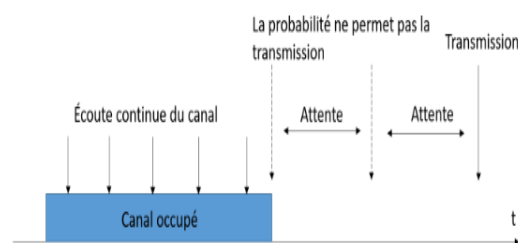


FIGURE 1.4 – Représentation schématique du protocole CSMA.

En complément, le tableau ci-après compare les différents schémas d'accès multiple en termes de leur principe, leurs avantages et leurs limitations en environnement sous-marin.

TABLE 1.3 – Comparaison des schémas classiques d'accès multiple dans le contexte ASM

Catégorie de Protocole	Exemples (Littérature ASM)	Principe d'adaptation	Avantages en ASM	Limites en ASM
Basé TDMA	TDA-MAC [Morozs et al., 2017]	Allocation dynamique des slots, synchronisation lâche	Évite les collisions (si synchronisé), bonne efficacité énergétique	Difficile à synchroniser, peu flexible, gaspillage de ressources si trafic faible.
OFDMA	[Tu et al., 2011]	Répartition dynamique de sous-porteuses orthogonales entre utilisateurs	Haute efficacité spectrale; allocation fine des ressources; flexibilité temporelle et fréquentielle	Sensibilité aux décalages Doppler; synchronisation stricte requise; complexité de gestion des ressources
Basé (CSMA /ALOHA)	FAMA, MACA-U [Ng et al., 2008]	Handshake (RTS/CTS) pour réserver le canal, backoff aléatoire	Simple, décentralisé, bon pour trafic sporadique (ad hoc)	Faible débit, overhead élevé (RTS/CTS), problème de la station cachée/exposée
Basé CDMA	U-CDMA [Kang, 2021]	Codes robustes, gestion de puissance (problème near-far)	Robuste au narrowband, LPI/LPD	Perte d'orthogonalité (Doppler), problème near-far sévère

Comme le montre le tableau 1.3, les schémas classiques d'accès multiple (TDMA, FDMA, CDMA, ALOHA, CSMA, OFDMA.) ont été conçus initialement pour des réseaux terrestres à large bande passante et faible latence, où les caractéristiques du canal sont relativement stables dans le temps. Leur transposition directe dans les environnements acoustiques sous-marins se heurte à de multiples limitations structurelles et physiques :

- **Temps de propagation élevé** : La vitesse du son dans l'eau (1500 m/s) est considérablement plus faible que celle des ondes radio (3×10^8 m/s), ce qui entraîne des délais importants. Les protocoles comme TDMA deviennent inefficaces, car le temps requis pour la coordination dépasse souvent la durée utile des créneaux de transmission.
- **Instabilité du canal** : Les propriétés acoustiques du milieu marin (température, salinité, profondeur) fluctuent en permanence, rendant le canal non stationnaire. Les schémas statiques d'allocation de ressources ne peuvent donc pas s'adapter aux variations rapides et imprévisibles des conditions de propagation.

- **Bande passante limitée** : Le canal acoustique sous-marin est fortement atténué aux hautes fréquences. L'utilisation de **FDMA** est donc contrainte par la bande passante exploitable (souvent de quelques kHz), avec une sélectivité fréquentielle marquée. Cela rend le partitionnement fréquentiel sous-optimal dans la plupart des cas.
- **Collisions et interférences** : Les protocoles à accès aléatoire (**ALOHA**, **CSMA**) souffrent d'un taux de collision élevé, aggravé par les délais de propagation. De plus, les transmissions simultanées de différents nœuds dans un canal partagé entraînent des interférences, nuisant à la fiabilité des communications.
- **Consommation énergétique** : Les nœuds sous-marins sont souvent alimentés par batterie, avec une capacité limitée de rechargement. Les protocoles inefficaces (par exemple, avec beaucoup de retransmissions ou d'écoute active inutile) entraînent une surconsommation d'énergie.
- **Difficulté de synchronisation** : En raison de la latence, de la variabilité temporelle des transmissions, ainsi que de l'absence d'accès au GPS pour les nœuds immergés, la synchronisation temporelle stricte, essentielle pour des schémas comme **TDMA**, est difficile à maintenir sans recourir à des mécanismes complexes et coûteux. Contrairement aux systèmes aériens où une horloge commune peut être facilement distribuée via le GPS, les nœuds sous-marins ne peuvent pas se synchroniser aussi précisément, ce qui complique la coordination temporelle des transmissions.

Ces limitations soulignent la nécessité d'adopter des stratégies plus souples, dynamiques et contextuelles pour la gestion des ressources dans les réseaux sous-marins. Cela implique de passer d'une logique d'allocation fixe ou centralisée à des mécanismes auto-organisés capables de :

- prendre en compte les dynamiques locales du canal et les états des nœuds ;
- coordonner efficacement l'accès au canal sans communication excessive ;
- gérer les conflits d'accès tout en minimisant les interférences ;
- optimiser l'utilisation de la bande passante disponible de manière adaptative.

C'est dans ce contexte que se pose la problématique d'allocation de ressources dans un réseau acoustique multi-utilisateurs, elle vise à répartir efficacement les ressources limitées (bande, puissance, temps) entre plusieurs nœuds actifs, tout en tenant compte de la nature fortement dynamique, distribuée et bruitée du canal. Le canal acoustique étant souvent un canal d'interférence, où les transmissions concurrentes dégradent mutuellement la qualité du signal, l'allocation de ressources devient un problème fondamental, à la croisée de l'optimisation, de la coordination décentralisée et de l'intelligence adaptative. Parmi les solutions techniques permettant de répondre à ces défis, la modulation **OFDM** qui se caractérise par sa capacité à diviser efficacement la bande passante disponible en sous-canaux orthogonaux, chacun étant moins sensible à l'évanouissement fréquentiel, elle constitue la base physique du schéma d'accès **OFDMA**, en permettant la transmission parallèle sur des sous-porteuses étroites et orthogonales.

La figure 1.5 montre une utilisation de la modulation **OFDM** dans un schéma d'accès par répartition dans le temps **TDMA**. Les ressources (la totalité des sous-porteuses) sont

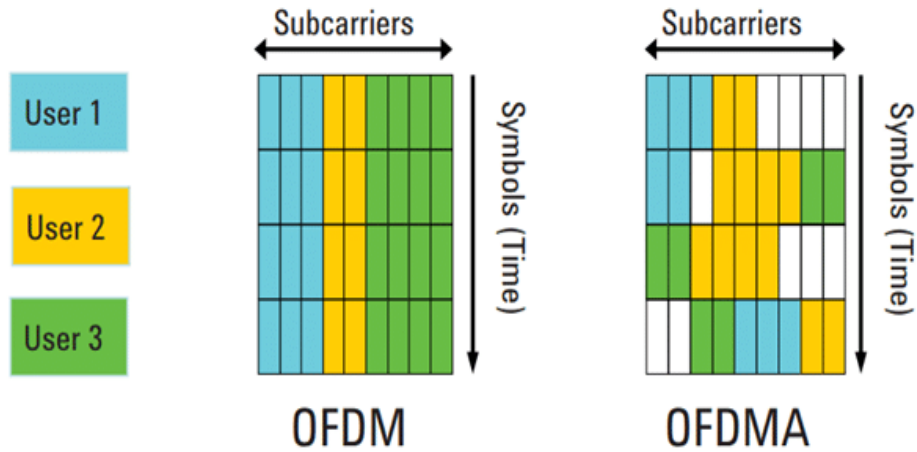


FIGURE 1.5 – Différence entre l’allocation des ressources en **OFDM** et en **OFDMA**

allouées à un seul utilisateur à la fois (d’abord l’utilisateur 1, puis le 2, puis le 3). À un instant t donné, le système est mono-utilisateur. Tandis que pour **OFDMA** c’est le véritable schéma d’accès multiple. Dans un même intervalle de temps (un même symbole), les sous-porteuses (la ressource fréquentielle) sont réparties entre plusieurs utilisateurs. L’utilisateur 1, l’utilisateur 2 et l’utilisateur 3 transmettent simultanément, chacun sur le sous-ensemble de sous-porteuses qui lui a été alloué.

Cette capacité de l’**OFDMA** à diviser finement la ressource temps-fréquence en blocs allouables est ce qui en fait un candidat idéal pour l’allocation de ressources multi-utilisateurs. L’**OFDM**, comme détaillé dans la section suivante, fournit la brique de modulation sous-jacente qui rend cette division possible et robuste. Sa flexibilité en fait une modulation particulièrement adaptée à l’environnement acoustique.

1.4.2 Technique de modulation **OFDM**

L’**OFDM** est une technique de modulation multi-porteuse particulièrement efficace pour la transmission de données à haut débit sur des canaux soumis à des évanouissements en fréquence. Il est basé sur la division du signal global en un grand nombre de signaux à bande étroite, appelés sous-porteuses, qui sont transmis en parallèle [Moose, 1994]. La particularité de l’**OFDM** réside dans l’orthogonalité entre ses sous-porteuses, d’où elles peuvent se chevaucher spectralement sans interférer entre elles, ce qui permet une utilisation optimale de la bande passante disponible, contrairement aux techniques classiques de multiplexage par répartition en fréquence, cette orthogonalité est assurée par une synchronisation stricte et un espacement des sous-porteuses.

Le débit binaire total R d’une transmission **OFDM** s’exprime en fonction du nombre de sous-porteuses N_{sc} et du débit par sous-porteuse R_{sc} selon :

$$R = N_{sc} \times R_{sc}. \quad (1.19)$$

La bande passante totale est approximativement donnée par

$$B = N_{sc} \times \Delta f, \quad (1.20)$$

où Δf est l'espacement entre deux sous-porteuses. Ainsi, le débit global dépend directement de la largeur de bande allouée.

Concernant la capacité théorique du canal, elle dépend du rapport signal sur bruit (SNR) sur chaque sous-porteuse. La capacité maximale (en bits par seconde) d'une transmission OFDM est donnée comme suit :

$$C = \sum_{n=1}^{N_{sc}} \Delta f \cdot \log_2(1 + \text{SNR}_n). \quad (1.21)$$

Si l'on suppose un SNR uniforme sur toutes les sous-porteuses, la capacité totale s'écrit alors

$$C = B \cdot \log_2(1 + \text{SNR}). \quad (1.22)$$

L'une des principales raisons qui justifient l'adoption de l'OFDM aussi pour les communications sous-marines est sa capacité à atténuer les effets du multi-trajet, phénomène particulièrement prononcé dans les environnements marins [Chithra et al., 2015]. Par ailleurs en insérant un intervalle de garde, à la fois cyclique et plus long que l'étalement temporel du canal, on évite d'une part que les copies retardées d'un symbole interfèrent avec le symbole suivant. Ce mécanisme, renforcé par un préfixe cyclique, permet non seulement de supprimer l'interférence inter-symbole mais aussi d'exploiter efficacement la bande passante limitée en autorisant un chevauchement contrôlé entre les sous-canaux.

Dans la transmission OFDM, la puissance d'émission également joue un rôle central puisqu'elle conditionne la qualité du lien et, en particulier, le SNR de chaque sous-porteuse. Le signal global peut être vu comme la superposition de plusieurs sous-porteuses orthogonales, chacune portant une fraction de la puissance totale. La puissance moyenne transmise s'exprime alors comme la somme des puissances allouées aux sous-porteuses disponibles :

$$P_{tot} = \sum_{n=1}^{N_{sc}} P_n, \quad (1.23)$$

où P_n représente la puissance attribuée à la n -ième sous-porteuse.

Dans le cas le plus simple d'une répartition uniforme, chaque sous-porteuse reçoit $n \in N_c$ la même part de puissance :

$$P_n = \frac{P_{tot}}{N_{sc}}. \quad (1.24)$$

Ainsi, plus le nombre de sous-porteuses utilisées N_{sc} est élevé, plus la puissance disponible par sous-porteuse diminue, ce qui impacte directement le SNR individuel et donc la capacité de chaque canal. Inversement, limiter le nombre de sous-porteuses disponibles permet de concentrer la puissance et d'améliorer la robustesse face au bruit, mais au prix d'une réduction de la bande passante exploitée.

La puissance d'émission et la répartition entre sous-porteuses jouent un rôle clé dans la performance individuelle de chaque lien OFDM, et dans un système multi-utilisateurs, ces décisions ne concernent plus un seul émetteur, mais plusieurs émetteurs ou paires émetteur-récepteur partageant simultanément les mêmes ressources spectrales. La nécessité de répartir de manière optimale la puissance et les sous-porteuses entre les utilisateurs

introduit alors un problème stratégique d'allocation des ressources, ce problème consiste à maximiser les performances globales du réseau (débit, fiabilité, efficacité spectrale) tout en tenant compte des interactions entre les utilisateurs, puisque l'action de chaque utilisateur influence les performances des autres.

1.4.3 Position du problème d'allocation de ressources en réseau multi-utilisateurs

Dans la continuité de l'adoption de la modulation **OFDM** comme solution adaptée aux contraintes du canal **ASM**, se pose naturellement la question de l'allocation des ressources dans un contexte multi-utilisateur. La gestion de ces ressources constitue un défi majeur, en raison des caractéristiques particulières du milieu sous-marin. Contrairement aux canaux terrestres, le canal acoustique sous-marin évolue dans un environnement non structuré, hétérogène et non réglementé, manquant d'infrastructure de réseau préexistante et de mécanismes de coordination standardisés. Cette absence de coordination entraîne une concurrence entre utilisateurs partageant des bandes de fréquences similaires, sans synchronisation, générant ainsi des interférences fréquentes et une utilisation inefficace du spectre disponible.

Dans ce contexte, où les nœuds du réseau (capteurs, véhicules autonomes, stations fixes ou mobiles) doivent partager un canal à bande passante limitée et fortement influencé par la latence et la variabilité des conditions de propagation, il devient essentiel de concevoir des mécanismes d'allocation capables d'optimiser l'utilisation des ressources de transmission (temps, fréquence, puissance, etc.). Ces mécanismes visent à maximiser les performances du réseau en termes de débit, de fiabilité et d'efficacité spectrale, tout en minimisant les interférences. La gestion des ressources dans un environnement acoustique sous-marin multi-utilisateurs requiert ainsi des stratégies adaptatives, décentralisées et robustes, capables de s'adapter à l'instabilité du canal et à l'absence de coordination centralisée.

Face à ces contraintes, le développement de solutions efficaces d'allocation des ressources représente un domaine de recherche actif, des approches variées, allant des méthodes conventionnelles aux techniques modernes basées sur l'apprentissage automatique et l'optimisation, ont été proposées pour relever les défis du canal acoustique sous-marin. La section suivante présente alors un état de l'art de ces solutions, en analysant leurs principes, avantages et limitations.

1.4.4 Solutions classiques et modernes : état de l'art

Les premières approches développées pour gérer l'accès au canal dans les réseaux acoustiques sous-marins s'inspirent largement des techniques éprouvées dans les réseaux terrestres, telles que **TDMA**, **FDMA** ou **CSMA**, bien que brièvement présentées en début de chapitre, il est important de souligner que ces méthodes, bien qu'intuitives et simples à mettre en œuvre, révèlent rapidement leurs limites dans un contexte sous-marin. Les contraintes spécifiques du canal acoustique comme la latence élevée, la bande passante réduite, et la variabilité du milieu limitent fortement leur efficacité, notamment en matière de synchronisation, d'évolutivité et de tolérance aux interférences, face à ces limitations, des approches plus récentes ont été explorées. Ces récentes solutions reposent sur

l'utilisation de modèles adaptatifs, capables d'apprendre, de s'adapter dynamiquement à l'environnement et de prendre des décisions autonomes, même en l'absence d'infrastructure centralisée. Dans ce contexte, les premières approches développées pour l'allocation de ressources dans les réseaux acoustiques sous-marins s'inspirent des techniques classiques utilisées dans les réseaux radio terrestres et ces méthodes peuvent être classées en deux grandes catégories : les méthodes classiques centralisées et les approches intelligentes décentralisées.

Approches classiques centralisées : Ces approches reposent sur une entité centrale (un nœud maître ou une station de surface) chargée de collecter les informations globales du réseau et d'optimiser de manière globale l'allocation des ressources. Ces méthodes, souvent basées sur des schémas d'accès sans contention, permettent une planification efficace et une coordination fine des transmissions pour éviter les collisions. Par exemple, plusieurs protocoles MAC pour les réseaux sous-marins s'appuient sur des adaptations de schémas TDMA, où la station centrale assigne des intervalles de temps (slots) à chaque nœud. D'autres protocoles utilisent une technique de Polling (interrogation), où la station de base interroge séquentiellement les nœuds pour leur donner le droit d'émettre, gérant ainsi explicitement l'accès au canal [Burrowes and Khan, 2009].

Cependant, plusieurs travaux ont montré que ces méthodes présentent également plusieurs limites dans les environnements sous-marins [Akyildiz et al., 2005]. En particulier, la latence de communication élevée due à la faible vitesse de propagation du son dans l'eau (environ 1500 m/s), la topologie dynamique des nœuds et l'imprécision des informations collectées souvent causée par des canaux acoustiques imprévisibles et changeants rendent difficile une coordination centralisée efficace.

Approches modernes décentralisées : Dans le contexte des réseaux sous-marins, où les contraintes de communication telles qu'une latence élevée, une bande passante limitée, une grande variabilité des canaux acoustiques et une topologie souvent dynamique rendent les approches centralisées peu efficaces, les méthodes décentralisées et heuristiques s'avèrent particulièrement adaptées pour l'allocation de ressources. Contrairement aux architectures centralisées qui dépendent d'un nœud maître pour agréger l'ensemble des informations et coordonner les décisions, les approches décentralisées répartissent l'intelligence du système entre les nœuds individuels. Chaque nœud prend des décisions de manière autonome en se basant uniquement sur des informations locales comme la qualité des canaux ou l'état des voisins tout en poursuivant un objectif global, tel que la maximisation du débit réseau, l'optimisation de la consommation énergétique, l'optimisation de ressources disponibles, ou la réduction de l'interférence inter-nœuds [Pompili and Akyildiz, 2009]-[Jornet et al., 2008]. Ces approches ont été largement explorées dans les réseaux sous-marins notamment en raison de leur résilience face à la mobilité des nœuds, de leur évolutivité et de leur capacité à fonctionner sans point de défaillance unique. Par exemple, [Pompili and Akyildiz, 2009] ont proposé des protocoles distribués de routage et de gestion d'énergie adaptés aux environnements 3D sous-marins, tandis que [Jornet et al., 2008] ont étudié des stratégies de coordination locale permettant une utilisation efficace de la bande passante acoustique dans des réseaux dynamiques. De plus, les approches bio-inspirées, comme les algorithmes de colonies de fourmis ou d'optimisation par essaim particulaire, en anglais Particle Swarm Optimization (PSO), ont également montré leur pertinence pour la prise de décision distribuée dans ces contextes complexes [Heidemann et al., 2012].

Deux grandes familles se distinguent particulièrement : les algorithmes inspirés par la nature, l'apprentissage par renforcement multi-agent, et les approches issues de la théorie des jeux :

1. **Algorithmes inspirés par la nature, en anglais Nature Inspired Algorithm (NIA)** : Les NIA tirent parti des mécanismes d'intelligence collective observés dans la nature, tels que la coopération d'insectes sociaux ou les dynamiques d'essaims, pour résoudre des problèmes complexes de manière distribuée [Soni et al., 2021]. Parmi ces algorithmes, l'optimisation par colonies de fourmis (Ant Colony Optimization — ACO) s'avère prometteuse pour l'allocation de ressources dans les réseaux acoustiques sous-marins. Dans ce cadre, les nœuds du réseau sont assimilés à des fourmis explorant diverses options d'allocation (par exemple, la sélection de canaux ou de niveaux de puissance). L'attractivité d'une option est représentée par une phéromone dont l'intensité est mise à jour localement en fonction de la qualité de la solution trouvée (ex. rapport signal/bruit, débit obtenu). Cette approche permet aux nœuds de converger vers des allocations efficaces sans coordination centrale, elle offre des avantages notables : auto-organisation, tolérance aux pannes et adaptabilité à la topologie du réseau [Kaveripakam and Chinthajjala, 2023, Chatterjee and Das, 2015].

De même, la **technique PSO** appliquée de manière distribuée permet à chaque nœud (particule) d'ajuster ses paramètres d'allocation (puissance, fréquence, etc.) en fonction de sa propre expérience et de celle de ses voisins. Cette forme d'apprentissage collaboratif favorise une convergence rapide vers des solutions satisfaisantes, avec une faible complexité computationnelle. PSO s'adapte particulièrement bien à des environnements dynamiques et incertains, tels que ceux rencontrés dans les réseaux sous-marins, notamment grâce à des approches comme celle proposée par [Krishnaswamy and Manvi, 2019], qui utilisent des versions discrètes de PSO pour le clustering dynamique et l'optimisation énergétique des réseaux acoustiques sous-marins.

2. **Apprentissage par renforcement multi-agents (MARL)** : Une approche décentralisée prometteuse peut améliorer la gestion des ressources dans les réseaux sans fil, notamment dans les environnements dynamiques et incertains. Dans cette approche, chaque nœud du réseau est modélisé comme un agent autonome capable d'apprendre par interaction avec son environnement, selon le paradigme d'essais-erreurs. L'objectif de chaque agent est de maximiser une récompense cumulative en prenant des décisions optimisées, telles que la sélection de canaux, le niveau de puissance, ou l'attribution de sous-porteuses à partir d'informations locales partielles, comme le niveau de bruit, la demande de trafic ou la présence de collisions [Hu et al., 2020]-[Rosenberger et al., 2021]-[Lee and Kim, 2023].

L'un des principaux avantages du **MARL** réside dans sa capacité à s'adapter en ligne, sans dépendre d'un modèle explicite du canal ou de l'environnement, ce qui le rend particulièrement pertinent dans des scénarios où la dynamique est difficile à modéliser a priori, comme les réseaux sous-marins, les véhicules autonomes, ou les drones communicants. Par ailleurs, contrairement aux approches centralisées, le **MARL** peut s'implémenter dans des systèmes entièrement distribués, réduisant

ainsi la surcharge de signalisation et les besoins en infrastructure.

Plusieurs études récentes illustrent l'efficacité du [MARL](#) dans des contextes concrets d'allocation de ressources. Dans le contexte des réseaux NR-V2X (New Radio Vehicle-to-Everything), les auteurs de [[Lee and Kim, 2023](#)] proposent une approche basée sur le [MARL](#) pour améliorer le taux de réception des paquets (Packet Reception Ratio - PRR) dans un environnement hétérogène où coexistent des paquets aux formats, priorités et objectifs variés. Grâce à des simulations, ils montrent que leur approche surpasse les méthodes d'allocation aléatoires et s'approche des performances d'un schéma optimal, tout en s'adaptant dynamiquement à la complexité du trafic véhiculaire.

Pour les systèmes [OFDM](#) multi-cellulaires, l'article [[Hu et al., 2020](#)] traite l'allocation conjointe des sous-porteuses et de la puissance comme un problème d'optimisation de l'efficacité spectrale. Après une première étape de décomposition en deux sous-problèmes, les auteurs introduisent une stratégie [MARL](#) permettant à chaque station de base (BS) d'adapter dynamiquement ses décisions en fonction de l'état du canal local. Leur approche permet d'obtenir jusqu'à 53,6% de gain d'efficacité spectrale par rapport aux méthodes classiques, avec une convergence plus rapide que les approches mono-agent.

Dans [[Gao et al., 2022](#)], les auteurs étudient un scénario de communication sous-marine impliquant plusieurs véhicules autonomes (autonomous underwater vehicle (AUVs)) pour l'acquisition collaborative de données. Chaque AUV prend ses décisions de manière autonome concernant l'affectation des tâches et le plan de trajectoire, sans coordination explicite avec les autres véhicules, afin d'optimiser l'utilisation du réseau acoustique sous-marin et réduire la consommation énergétique globale. Le problème est abordé par un algorithme intelligent basé sur l'optimisation de l'énergie, qui permet aux AUVs de sélectionner leurs actions de manière distribuée en fonction des conditions locales du réseau et de la mission. Les résultats de simulation montrent que cette approche permet de réduire efficacement la consommation énergétique du système multi-AUV et d'atteindre un bon compromis entre performance opérationnelle et coût énergétique.

3. **Approches basées sur la théorie des jeux (Game Theory) :** La théorie des jeux constitue un cadre mathématique puissant pour modéliser les interactions stratégiques entre les nœuds d'un réseau sous-marin lors de l'allocation de ressources critiques telles que la bande passante, la puissance de transmission ou les créneaux temporels [[Muhammed et al., 2018](#)]. Dans les approches non coopératives, chaque nœud cherche à maximiser sa propre utilité (par exemple, le débit ou l'efficacité énergétique) sans coordination avec les autres, ce qui conduit souvent à des situations d'équilibre de Nash où aucun nœud n'a intérêt à modifier unilatéralement sa stratégie [[Gharsalli et al., 2023a](#)]-[[Pottier et al., 2017](#)].

Ces jeux sont couramment utilisés pour l'allocation de puissance, où chaque nœud ajuste son niveau de transmission pour maximiser son Rapport signal sur interférence plus bruit (SINR) tout en tenant compte de l'impact sur ses voisins, ou pour l'accès au canal, où les nœuds doivent décider de leurs instants de

transmission pour minimiser les collisions. À l'inverse, dans les jeux coopératifs [Chee et al., 2006], les nœuds peuvent former des coalitions ou négocier afin de coordonner leurs décisions et partager équitablement les gains issus de la coopération, ce qui permet souvent d'améliorer l'équité, la robustesse et la performance globale du réseau. Des mécanismes tels que les jeux de formation de coalition, les jeux de négociation (Bargaining) [Gharsalli et al., 2024a], et les approches de routage coopératif incitatif ont ainsi été proposés pour les réseaux acoustiques sous-marins, en s'appuyant sur des critères de stabilité et d'équité [He et al., 2019]-[Bardhan and Ghose, 2013a].

Par ailleurs, les environnements marins étant dynamiques et incertains, les jeux dynamiques comme les jeux répétés, les jeux de Stackelberg (avec un leader et des followers), ou encore la théorie des jeux évolutionnaires, sont particulièrement adaptés pour modéliser l'évolution des comportements dans le temps [Tang et al., 2021]-[Rathi and Gupta, 2020]. De plus, les jeux stochastiques ou bayésiens permettent de gérer l'incertitude inhérente à l'environnement (par exemple, les variations du canal ou l'information incomplète entre nœuds) [Akkarajitsakul et al., 2011]-[Chang et al., 2016]. Enfin, dans des contextes où plusieurs objectifs doivent être optimisés simultanément (débit, énergie, délai), des modèles multi-objectifs issus de la théorie des jeux sont explorés, souvent en combinaison avec des approches d'MARL, permettant d'apprendre des stratégies d'allocation efficaces dans des environnements incertains et distribués.

L'approche de la théorie des jeux est particulièrement puissante car elle permet non seulement de modéliser les interactions stratégiques dans des environnements distribués et incertains, mais aussi de formaliser et d'analyser la notion même de compromis entre utilisateurs aux intérêts potentiellement conflictuels. En effet, dans un réseau acoustique sous-marin où les ressources sont limitées et partagées, les décisions de chaque utilisateur impactent les autres, créant des situations de concurrence, mais aussi des opportunités de coopération.

La théorie des jeux offre ainsi des outils permettant de trouver des équilibres stables, où aucun joueur ne peut améliorer sa situation sans détériorer celle d'un autre, garantissant une forme d'optimalité locale, mais aussi des mécanismes de négociation et de collaboration pour atteindre des solutions globalement plus efficaces et équitables. Ces compromis se traduisent par des stratégies qui maximisent non seulement les gains individuels, mais aussi le bien-être collectif du réseau, en équilibrant par exemple fonction d'utilité, l'optimisation de ressources disponible, et l'équité entre les utilisateurs.

De plus, la théorie des jeux permet d'aborder la complexité des environnements sous-marins à travers des modèles dynamiques et stochastiques, où les utilisateurs adaptent continuellement leurs stratégies en fonction des évolutions du canal et du comportement des autres, assurant ainsi robustesse et résilience face aux incertitudes. Cette capacité à intégrer à la fois compétition, coopération, et adaptation fait de la théorie des jeux un cadre essentiel pour concevoir des solutions d'allocation de ressources innovantes, performantes et justes.

Une discussion détaillée sur les types de jeux et surtout des jeux non coopératifs sera présentée au chapitre 2, suivie des jeux de négociation (Nash Bargaining) au chapitre 3,

puis des jeux de coalition au chapitre 4.

1.5 Conclusion

En résumé, ce chapitre a permis de poser les bases nécessaires à l'étude des communications acoustiques sous-marines en mettant en évidence leurs défis spécifiques, notamment la variabilité du canal, la bande passante limitée et l'absence d'infrastructures normalisées. Face à ces contraintes, l'allocation optimale des ressources apparaît comme un enjeu central pour garantir la fiabilité et l'efficacité des transmissions.

La modulation **OFDM** se révèle être une technique prometteuse grâce à sa flexibilité dans la gestion des sous-porteuses et de la puissance, tandis que la théorie des jeux offre un cadre adapté pour modéliser les interactions complexes entre utilisateurs et concevoir des mécanismes décentralisés d'allocation des ressources. Ces approches combinées ouvrent la voie à des solutions adaptatives et performantes, capables de s'adapter aux conditions dynamiques du canal acoustique sous-marin.

Le chapitre suivant approfondira les notions et les fondements théoriques de la théorie des jeux appliquée à notre contexte, en préparant le cadre nécessaire à la modélisation et à l'analyse des stratégies d'allocation.

Chapitre 2

Théorie des jeux dans les réseaux acoustiques sous marins

2.1 Théorie des jeux pour l'allocation de ressources

Dans de nombreux domaines : économie, politique, stratégie militaire, biologie ou encore intelligence artificielle, les individus prennent des décisions qui ne peuvent être comprises isolément : leurs choix dépendent de ce que feront les autres. C'est dans ce contexte que la théorie des jeux intervient.

La théorie des jeux est une branche des mathématiques appliquées et de l'économie qui étudie les interactions stratégiques entre des agents rationnels. Elle modélise des situations où le résultat pour chaque participant dépend non seulement de ses propres choix, mais aussi de ceux des autres.

Cette discipline propose un cadre rigoureux pour analyser les comportements en situation de conflit ou de coopération, en cherchant à identifier les équilibres ou les stratégies optimales pour chacun des acteurs.

La théorie des jeux trouve ses origines dans les réflexions économiques et philosophiques du XVIII^e siècle, mais sa formalisation mathématique moderne remonte au XX^e. Dès 1838, l'économiste français Antoine Augustin Cournot analyse le comportement stratégique de deux entreprises concurrentes dans son modèle de duopole, introduisant implicitement la notion d'équilibre stratégique [Cournot, 1863]. Toutefois, c'est en 1928 que John von Neumann publie un article fondateur dans lequel il démontre le théorème du minimax pour les jeux à deux joueurs à somme nulle, posant ainsi les bases mathématiques de la discipline. La véritable naissance de la théorie des jeux intervient en 1944 avec l'ouvrage majeur "Theory of Games and Economic Behavior", coécrit par von Neumann et l'économiste Oskar Morgenstern, où les auteurs développent un cadre général pour modéliser les interactions stratégiques entre agents rationnels [Von Neumann and Morgenstern, 1944]. Ce travail fondateur introduit notamment la théorie des jeux coopératifs et la formalisation des préférences individuelles à l'aide de fonctions d'utilité. Quelques années plus tard, John Nash révolutionne la discipline en introduisant en 1950 le concept d'équilibre de Nash, en anglais Nash Equilibrium NE, applicable à un ensemble plus large de jeux, y compris les jeux non coopératifs [Nash Jr, 1950]. Son approche ouvre la voie à une extension de la théorie à des situations avec information imparfaite ou dynamique.

Depuis les années 1970, la théorie des jeux s'est imposée comme un outil fondamen-

tal dans de nombreux domaines : économie, biologie évolutive, science politique, informatique, et plus récemment intelligence artificielle. Des ouvrages de référence tels que [Fudenberg and Tirole, 1991], ou d’Osborne et Rubinstein [Osborne and Rubinstein, 1994], ont contribué à structurer cette discipline, tandis que les travaux plus récents sur les systèmes multi-agents [Shoham and Leyton-Brown, 2008] permettent d’intégrer des notions comme l’apprentissage, l’incertitude ou encore la coopération dans des environnements complexes.

2.1.1 Définitions et types de jeux

Cette section a pour objectif d’introduire les concepts fondamentaux de la théorie des jeux, afin de fournir les bases nécessaires à la compréhension des chapitres suivants. L’accent est tout d’abord mis sur les jeux non coopératifs, en particulier ceux représentés sous forme normale, qui constituent le cadre d’analyse initial. Cette première partie permettra d’établir une compréhension rigoureuse des interactions stratégiques avant de se tourner, dans les chapitres ultérieurs, vers les jeux coopératifs, notamment les jeux de négociation et les jeux de coalition, qui seront au cœur de l’étude principale de notre thèse.

Jeux sous forme stratégique :

Définition : Un jeu stratégique (ou jeu en forme normale) est un modèle mathématique qui décrit une situation d’interaction entre plusieurs agents rationnels, appelés joueurs, où chaque joueur :

- choisit une stratégie de manière simultanée (ou sans connaître les choix des autres) ;
- cherche à maximiser son propre gain ou utilité ;
- prend ses décisions en tenant compte des possibles choix des autres joueurs.

Formellement, un jeu stratégique à K joueurs est défini par un triplet

$$\mathcal{G} = \{\mathcal{K}, \mathcal{S}, u\}, \quad (2.1)$$

où

- $\mathcal{K} = \{1, \dots, K\}$: est l’ensemble des acteurs (appelés joueurs ou bien utilisateur) qui interagissent dans la situation en considération,
- $\mathcal{S} = S_1 \times \dots \times S_K$: est l’ensemble des stratégies disponibles pour chaque joueur $i \in \mathcal{K}$,
- u : est la fonction d’utilité du joueur i , et on note $u_i(s)$ son degré de satisfaction vis-à-vis du profil d’action $s \in \mathcal{S}$.

Dans le contexte des environnements acoustiques sous-marins, les joueurs peuvent être modélisés comme un ensemble d’utilisateurs ou de liens émetteur-récepteur cherchant à optimiser l’utilisation des ressources disponibles. Dans cette thèse, l’attention est principalement portée sur deux dimensions stratégiques essentielles : l’allocation de puissance et le partage des sous-porteuses entre les utilisateurs. Les actions possibles des joueurs correspondent ainsi aux niveaux de puissance d’émission choisis et à la sélection des sous-porteuses à utiliser et à partager dans le spectre. Les fonctions d’utilité sont quant à elles définies en fonction de critères de performance pertinents pour les communications acoustiques sous-marines, tels que la capacité de transmission, le taux d’erreur binaire,

ou encore le **SNR**. Ce cadre permet de modéliser de manière rigoureuse les interactions concurrentes entre utilisateurs dans un canal acoustique complexe, caractérisé par des ressources spectrales limitées, une atténuation importante et une latence élevée.

Jeu fini : Un jeu est dit fini si :

- le nombre de joueurs est fini ;
- chaque joueur dispose d'un ensemble fini de stratégies.

Autrement dit, on a $|N| < \infty$ et $|S_i| < \infty$ pour tout $i \in K$. Dans un jeu fini, l'existence d'un **NE** (en stratégies mixtes) est toujours garantie, ce qui en fait une catégorie très étudiée en théorie des jeux.

Jeu à somme nulle : Un jeu à deux joueurs est dit à somme nulle lorsque le gain de l'un est exactement opposé à celui de l'autre, c'est-à-dire que :

$$\forall (s_1, s_2) \in S_1 \times S_2, \quad u_1(s_1, s_2) + u_2(s_1, s_2) = 0. \quad (2.2)$$

Ce type de jeu reflète des situations de pure compétition. Un joueur ne peut améliorer son gain qu'aux dépens de l'autre. Ces jeux admettent des propriétés spécifiques, comme l'existence d'un **NE** coïncidant avec l'équilibre de minimax, et peuvent être résolus par programmation linéaire.

Exemple : le dilemme du prisonnier :

Le **dilemme du prisonnier** est un jeu stratégique à deux joueurs, illustrant le conflit entre intérêt individuel et intérêt collectif. Chaque joueur a deux stratégies possibles : coopérer (**C**) ou trahir (**T**). Les gains associés sont donnés par la matrice suivante :

		Joueur 2	
		C	T
Joueur 1	C	(3, 3)	(0, 5)
	T	(5, 0)	(1, 1)

Dans cette configuration :

- Si les deux joueurs coopèrent, chacun obtient une récompense modérée (3, 3).
- Si un joueur trahit pendant que l'autre coopère, le traître obtient le gain maximal (5), et l'autre le gain minimal (0).
- Si les deux trahissent, chacun obtient un gain faible (1, 1).

Ce jeu possède un unique **NE** en stratégie pure : (**T**, **T**). Il montre que des décisions purement rationnelles peuvent conduire à un résultat sous-optimal pour l'ensemble des joueurs. Ce paradoxe est particulièrement pertinent dans des contextes où la coopération est souhaitable mais non garantie, comme dans la gestion des ressources partagées ou les communications multi-utilisateurs.

2.1.2 Équilibre de Nash (Nash Equilibrium)

NE est l'un des concepts fondamentaux de la théorie des jeux non coopératifs. Il permet de modéliser des situations où chaque joueur, conscient des stratégies des autres, choisit la sienne de manière à maximiser son propre intérêt, sans incitation à dévier unilatéralement. Le **NE** repose sur l'idée que chaque joueur adopte une meilleure réponse aux stratégies des autres.

Un profil de stratégies est un équilibre de Nash si chaque stratégie jouée est une meilleure réponse aux stratégies adverses.

Definition 1. *Meilleure réponse [en anglais : Best Response (**BR**)]* La meilleure réponse du joueur i face aux stratégies des autres joueurs $s_{-i} \in \prod_{j \neq i} S_j$ est l'ensemble des stratégies qui maximisent son utilité, soit :

$$BR_i(s_{-i}) = \arg \max_{s_i \in S_i} u_i(s_i, s_{-i}). \quad (2.3)$$

Autrement dit, $s_i^* \in BR_i(s_{-i})$ signifie que s_i^* est une stratégie optimale pour i étant donné que les autres joueurs jouent s_{-i} .

Definition 2 (NE). *Un profil de stratégies*

$$s^* = (s_1^*, s_2^*, \dots, s_K^*) \in \mathcal{S} \quad (2.4)$$

est un **NE** si, pour tout joueur $i \in \mathcal{K}$, aucun écart unilatéral ne permet d'améliorer son utilité. Autrement dit, pour tout $i \in \mathcal{K}$ et pour toute stratégie $s_i \in S_i$,

$$u_i(s_i^*, s_{-i}^*) \geq u_i(s_i, s_{-i}^*), \quad (2.5)$$

où $s_{-i}^* = (s_1^*, \dots, s_{i-1}^*, s_{i+1}^*, \dots, s_K^*)$ désigne le profil des stratégies des autres joueurs.

Application : Trouver l'Équilibre de Nash par Meilleures Réponses

Appliquons les définitions ci-dessus à l'exemple du dilemme du prisonnier pour trouver formellement l'équilibre.

1. Meilleures réponses (BR) du Joueur 1 : Nous analysons les choix du Joueur 1 (gains à gauche) en fonction des actions du Joueur 2.

- **Si le Joueur 2 joue C :** Le Joueur 1 compare son gain s'il joue C (obtient 3) à son gain s'il joue T (obtient 5). Puisque $5 > 3$, $BR_1(C) = \{T\}$.
- **Si le Joueur 2 joue T :** Le Joueur 1 compare son gain s'il joue C (obtient 0) à son gain s'il joue T (obtient 1). Puisque $1 > 0$, $BR_1(T) = \{T\}$.

2. Meilleures réponses (BR) du Joueur 2 : Nous analysons les choix du Joueur 2 (gains à droite) en fonction des actions du Joueur 1.

- **Si le Joueur 1 joue C :** Le Joueur 2 compare son gain s'il joue C (obtient 3) à son gain s'il joue T (obtient 5). Puisque $5 > 3$, $BR_2(C) = \{T\}$.
- **Si le Joueur 1 joue T :** Le Joueur 2 compare son gain s'il joue C (obtient 0) à son gain s'il joue T (obtient 1). Puisque $1 > 0$, $BR_2(T) = \{T\}$.

3. Identification de l'Équilibre de Nash : Un NE est un profil de stratégies (s_1^*, s_2^*) où $s_1^* \in BR_1(s_2^*)$ et $s_2^* \in BR_2(s_1^*)$.

- Profil (C, C) : $C \notin BR_1(C)$. Pas un NE.
- Profil (C, T) : $C \notin BR_1(T)$. Pas un NE.
- Profil (T, C) : $C \notin BR_2(T)$. Pas un NE.
- Profil (T, T) :
 - $T \in BR_1(T)$ (vérifié).
 - $T \in BR_2(T)$ (vérifié).

Les deux stratégies sont des meilleures réponses mutuelles. C'est l'unique NE.

Dans un réseau de communication acoustique sous-marin, un NE désigne une situation d'équilibre où chaque utilisateur adopte une stratégie d'allocation (comme la distribution de sa puissance ou de sa bande passante) telle qu'aucune modification individuelle ne lui procurerait d'amélioration, tant que les autres maintiennent leurs choix.

Remarque : Un profil de stratégies $s^* = (s_1^*, \dots, s_K^*)$ est un NE si et seulement si chaque stratégie s_i^* est une meilleure réponse aux stratégies des autres joueurs s_{-i}^* , c'est-à-dire :

$$s_i^* \in BR_i(s_{-i}^*), \quad \forall i \in \mathcal{K}.$$

Les concepts fondamentaux de la théorie des jeux non coopératifs, et en particulier la notion d'Équilibre de Nash, offrent désormais un cadre pertinent pour étudier leur application à la communication ASM. En effet, la problématique de gestion des ressources (temps, bande passante, puissance) dans ces réseaux se prête particulièrement bien à une modélisation sous forme de jeu. Chaque joueur, cherchant à maximiser son propre débit et à optimiser l'allocation de ressources disponibles, doit tenir compte des décisions des autres joueurs partageant le même canal.

Dans la section suivante, nous illustrons comment ces concepts théoriques peuvent être exploités pour concevoir des mécanismes d'accès multiple décentralisés et adaptatifs, reposant sur l'Équilibre de Nash comme solution stable pour l'allocation des ressources.

2.2 Application à la gestion des ressources en ASM

Dans les réseaux ASM, la gestion efficace des ressources, telles que la puissance d'émission et les sous-porteuses disponibles, est un enjeu crucial. La théorie des jeux fournit un cadre pertinent pour modéliser les interactions entre utilisateurs autonomes, souvent concurrents, dans un environnement contraint et bruyant. Dans ce contexte, les joueurs sont généralement les émetteurs (ou les paires émetteur-récepteur), chacun ayant pour objectif de maximiser sa propre performance (débit, fiabilité) tout en minimisant les interférences causées aux autres.

Les actions possibles sont :

- l'allocation de puissance d'émission ;

— le partage des bandes de fréquences (ou sous-porteuses) disponible.

Dans la partie suivante, nous commencerons par une étude des jeux non coopératifs appliqués à l'allocation des ressources pour un système [ASM](#), avant de nous orienter vers les jeux coopératifs, notamment les jeux de négociation et de coalition associés au Q-learning, qui seront traités respectivement dans les chapitres 3 et 4.

2.2.1 Jeux non coopératifs pour l'allocation des ressources

La gestion efficace des ressources dans les réseaux [ASM](#) constitue un enjeu majeur, en raison des contraintes physiques et environnementales propres à ce type de communication, parmi ces contraintes figurent la forte atténuation des signaux acoustiques, la propagation lente des ondes sonores dans l'eau (environ 1500 m/s), les effets de multi-trajets, ainsi que la présence d'interférences causées par le partage du canal entre plusieurs utilisateurs. Dans ce contexte, optimiser l'allocation des ressources, notamment la puissance d'émission et la gestion du spectre, est indispensable pour assurer la fiabilité et l'efficacité des communications et face à l'absence d'une coordination centralisée dans la plupart des réseaux [ASM](#), la théorie des jeux non coopératifs s'impose comme un cadre théorique pertinent. Elle permet de modéliser les interactions stratégiques entre des agents autonomes, typiquement des émetteurs ou des paires émetteur-récepteur, chacun cherchant à maximiser sa propre utilité (débit, fiabilité) tout en tenant compte des actions des autres utilisateurs, des précisions supplémentaires sont disponibles dans [[Pottier, 2018](#)].

Les jeux non coopératifs sont particulièrement adaptés pour modéliser les situations où les utilisateurs doivent partager des ressources limitées tout en étant potentiellement en compétition. Divers types de jeux peuvent être considérés, comme les jeux de partage de ressources, les jeux à somme non nulle, ou encore les jeux de puissance, ces modèles permettent d'identifier des équilibres stables, en particulier les équilibres de Nash où aucun joueur n'a intérêt à modifier unilatéralement sa stratégie, cependant, un équilibre de Nash ne garantit pas toujours une distribution équitable des ressources.

Par ailleurs, nous intégrons la modulation [OFDM](#) comme solution technique complémentaire dans ce contexte, grâce à sa capacité à diviser la bande passante en multiples sous-porteuses orthogonales, l'[OFDM](#) offre une meilleure gestion du spectre et atténue les effets liés au retard de propagation et à la dispersion du canal. Cela en fait une technologie particulièrement adaptée à l'environnement acoustique sous-marin, comme l'ont montré les travaux de [[Zhou and Wang, 2014](#)].

Afin de formaliser rigoureusement le problème, la section suivante introduit la modélisation mathématique du jeu non coopératif appliqué à l'allocation de ressources. Nous y définissons précisément les variables du système, les objectifs des agents, ainsi que les contraintes techniques à respecter.

2.2.2 Modélisation du problème par un jeu non coopératif

Avant de formaliser mathématiquement le problème d'allocation de ressources dans un environnement [ASM](#), il est essentiel de définir le cadre général du jeu non coopératif dans lequel les différents joueurs (nœuds du réseau) évoluent. Dans ce type de jeu, chaque joueur agit de manière autonome, sans coordination explicite avec les autres, et cherche

où, pour chaque émetteur i sur la $n^{\text{ème}}$ sous-porteuse, les définitions et notations suivantes sont adoptées :

- $x_i[n, m]$ représente le symbole de modulation ;
- $h_{ii}[n, m]$ correspond aux coefficients de réponse en fréquence du canal direct ;
- $h_{ji}[n, m]$ représente les coefficients du canal d'interférence ;
- $z_i[n, m]$ est un bruit gaussien complexe supposé à moyenne nulle et variance $\sigma_{i,n}^2$.

Le canal **ASM** est caractérisé par une sélectivité fréquentielle et une forte variabilité temporelle. Pour y remédier, nous supposons que le préfixe cyclique T_g est supérieur à la dispersion temporelle du canal et que l'espacement entre les sous-porteuses Δf est plus grand que la plus large dispersion Doppler du canal, on considère également que les délais de propagation sont supérieurs au temps de cohérence des évanouissements rapides du canal et que les émetteurs sont synchronisés.

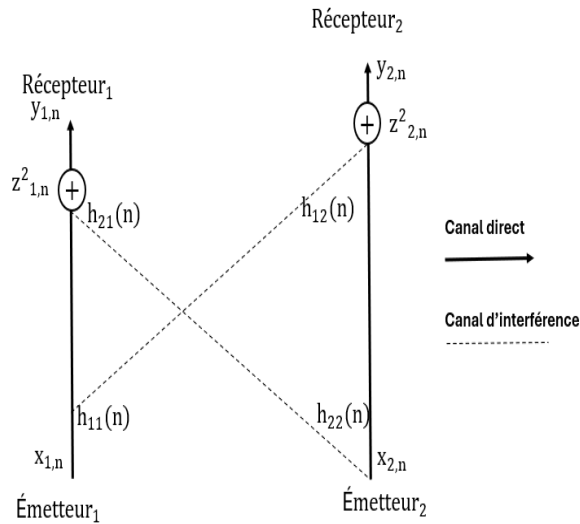


FIGURE 2.2 – Canal acoustique à interférences avec 2 paires émetteur-récepteur

Considérant ces hypothèses, le système étudié est formulé comme non coopératif dans notre analyse, en raison de l'incapacité des utilisateurs à communiquer et à parvenir à un accord sur le partage des ressources. Ainsi, chaque utilisateur transmet son signal de manière indépendante sur les N_{sc} sous-porteuses et alloue sa puissance uniquement en fonction des statistiques du canal, du bruit et des interférences des autres utilisateurs, sous la contrainte de puissance suivante, pour chaque utilisateur i :

$$\sum_{n=1}^{N_{sc}} p_{i,n} \leq P_i^{max}, \forall i \in \{1, \dots, K\}, \quad (2.7)$$

où $p_{i,n}$ est la puissance allouée par l'utilisateur i sur la $n^{\text{ème}}$ sous-porteuse et P_i^{max} est la puissance maximale disponible pour chaque utilisateur i .

Nous formulons ensuite le problème mathématique proposé afin d'allouer efficacement la puissance $p_{i,n}$. Nous définissons le vecteur de puissance \mathbf{P} alloué à tous les utilisateurs sur l'ensemble des sous-porteuses comme suit :

$$\mathbf{P} = (p_{i,j})_{i \in \{1, \dots, K\}, j \in \{1, \dots, N_{sc}\}} = (\mathbf{p}_i, \mathbf{P}_{-i}), \quad (2.8)$$

où $\mathbf{P}_{-i} = (\mathbf{p}_1, \dots, \mathbf{p}_{i-1}, \mathbf{p}_{i+1}, \dots, \mathbf{p}_{N_{sc}})$, $\forall i \in \{1, \dots, K\}$.

2.2.4 Choix de la fonction d'utilité

Le problème de partage des ressources dans un environnement [ASM](#) avec plusieurs utilisateurs peut être modélisé comme un jeu non coopératif. Dans ce cadre, chaque utilisateur cherche à maximiser son propre taux de transmission en allouant de manière stratégique sa puissance de transmission sur les sous-porteuses d'un système [OFDM](#), tout en tenant compte des interférences générées par les autres utilisateurs. La résolution de ce jeu devient particulièrement complexe dans un environnement partiellement observable, comme c'est souvent le cas dans les réseaux sous-marins, où chaque utilisateur n'a qu'une connaissance limitée des canaux des autres.

Dans un cadre idéal où les utilisateurs auraient une connaissance parfaite des canaux et des actions des autres, la formulation classique d'un jeu non coopératif consisterait à maximiser le taux de transmission de chaque utilisateur et optimiser la distribution de puissance. Pour chaque utilisateur i , ce taux peut être exprimé comme suit :

$$u_i(\mathbf{P}) = \frac{1}{BT} \sum_{n=1}^{N_{sc}} \log_2(1 + \text{SINR}_{i,n}(\mathbf{P})), \quad (2.9)$$

où $\text{SINR}_{i,n}(\mathbf{P})$ représente le [SINR](#) pour l'utilisateur i sur la $n^{\text{ème}}$ sous-porteuse qui s'exprime comme suit :

$$\text{SINR}_{i,n}(P) = \frac{\mathbb{E}|h_{ii}(n)|^2 p_{i,n}}{\sum_{j \neq i} \mathbb{E}|h_{ji}(n)|^2 p_{j,n} + \sigma_{i,n}^2}. \quad (2.10)$$

Cependant, cette approche idéale est peu applicable dans un environnement acoustique sous-marin, où les canaux sont souvent inconnus ou partiellement observables. Dans ce contexte, la connaissance complète des autres utilisateurs est rarement disponible, ce qui contraint les utilisateurs à prendre des décisions basées sur des informations locales et des estimations probabilistes des canaux. Pour traiter cette incertitude, une solution courante consiste à reformuler le problème sous la forme d'un **jeu bayésien**.

Jeu bayésien : gestion de l'incertitude Dans ce cadre, chaque joueur forme des croyances probabilistes sur les canaux et les actions des autres, ce qui permet d'intégrer l'incertitude dans les stratégies. Cependant, les jeux bayésiens souffrent de plusieurs limitations :

- **Simplification du modèle de canal** : Les canaux sont souvent modélisés de manière simplifiée (par exemple, en utilisant des modèles de Bernoulli).
- **Croyances partagées** : Les jeux bayésiens supposent que tous les joueurs partagent les mêmes distributions a priori, ce qui est rarement réaliste dans un environnement dynamique comme celui des réseaux sous-marins.
- **Limitation du nombre de joueurs/actions** : Cette approche est souvent limitée à un petit nombre de joueurs et d'actions, ce qui peut restreindre son application à des systèmes plus complexes.

Ainsi, bien que les jeux bayésiens soient une approche utile dans des contextes de faible incertitude, leur application dans les réseaux acoustiques sous-marins est limitée par la variabilité et l'incertitude des canaux.

Approche robuste (une solution plus réaliste) :

Cette méthode a été initialement proposée par [Pottier, 2018] qui vise à surmonter les limitations des jeux bayésiens classiques et proposer une approche robuste qui est plus adaptée. Contrairement aux jeux bayésiens, cette approche ne repose pas sur des croyances probabilistes partagées, mais sur l'optimisation de la stratégie de puissance en anticipant le pire scénario d'interférence. Concrètement, chaque utilisateur prend en compte l'incertitude liée aux canaux des autres utilisateurs et ajuste sa stratégie en conséquence, sans avoir besoin d'une connaissance directe des actions des autres. Ce modèle est plus réaliste dans les réseaux sous-marins où les utilisateurs ne peuvent pas observer directement les stratégies des autres, mais peuvent estimer les interférences sur la base de statistiques locales.

Cette approche robuste trouve sa place dans le cadre des jeux non-coopératifs, en effet, chaque utilisateur optimise sa stratégie sans coopération explicite, cherchant à maximiser son propre gain tout en prenant en compte l'impact de l'incertitude sur les interférences. Dans ce contexte, les utilisateurs choisissent leurs allocations de puissance de manière indépendante, ce qui est au cœur du modèle du jeu non-coopératif.

La fonction d'utilité robuste peut alors être exprimée comme suit :

$$u_i(\mathbf{p}_i, \mathbf{p}_{-i}) = \min_{f_{H_{ji}}} \alpha \sum_{n=1}^{N_{sc}} \mathbb{E} \left[\log \left(1 + \frac{|H_{ii,n}|^2 p_{i,n}}{\sigma_{i,n}^2 + \sum_{j \neq i} |H_{ji,n}|^2 p_{j,n}} \right) \right], \quad (2.11)$$

où $f_{H_{ji}}$ représente la distribution des canaux interférents H_{ji} et $\alpha = (N_{sc} T \Delta f)^{-1}$, avec T désigne le temps symbole.

En assumant des interférences gaussiennes et en appliquant l'inégalité de Jensen [Pottier, 2018], cette fonction peut être réécrite comme suit :

$$u_i(\mathbf{p}_i, \mathbf{p}_{-i}) \approx \alpha \sum_{n=1}^{N_{sc}} \log (1 + \gamma_{i,n}(\mathbf{p}_{-i}) \cdot p_{i,n}), \quad (2.12)$$

où $\gamma_{i,n}(\mathbf{p}_{-i})$ est un paramètre dépendant des estimations locales du canal et des interférences des autres utilisateurs et exprimé comme suit :

$$\gamma_{i,n}(\mathbf{p}_{-i}) = \frac{e^{\mathbb{E}[\log(|H_{ii,n}|^2)]}}{\sigma_{i,n}^2 + \sum_{j \neq i} \mathbb{E}[|H_{ji,n}|^2] p_{j,n}}. \quad (2.13)$$

D'après l'équation 2.12, la résolution non coopérative du jeu ne nécessite pour chaque joueur que des informations locales : les statistiques de son canal direct H_{ii} et la densité spectrale de puissance du bruit total (incluant les interférences). Cette formulation repose sur une estimation locale des canaux et des interférences, ce qui permet aux utilisateurs d'adopter des stratégies de puissance optimales sans avoir besoin de connaître les stratégies exactes des autres.

Cette approche est ainsi particulièrement pratique et robuste face à l'incertitude inhérente aux canaux sous-marins. Son intérêt principal réside dans le fait que les données d'entrée utilisées pour la décision stratégique sont des statistiques du canal, et non des valeurs instantanées d'estimation du canal. En pratique, obtenir ces valeurs instantanées est difficile dans un système **ASM**, car cela nécessiterait que le récepteur renvoie à l'émetteur l'estimation du canal calculée à la réception. Or, en raison de la variabilité temporelle et spatiale du canal acoustique sous-marin, ces estimations sont souvent obsolètes ou peu fiables. L'utilisation de statistiques permet donc aux utilisateurs d'un système **ASM** de partager le spectre de manière décentralisée avec une connaissance minimale de leur environnement, tout en maintenant une allocation de ressources efficace et robuste.

Comparaison entre les approches bayésienne et robuste : Une comparaison des deux approches est présentée par le tableau 2.1. Dans un jeu bayésien, les canaux sont simplifiés et les croyances des utilisateurs sont partagées, ce qui peut entraîner des inefficacités dans un environnement dynamique. En revanche, l'approche robuste offre une plus grande flexibilité et une meilleure adaptation aux incertitudes des canaux en prenant en compte les interférences de manière locale et distribuée.

TABLE 2.1 – Comparaison des jeux bayésien et robuste

Critère	Jeu Bayésien	Approche Robuste
Modélisation du canal	Simplifiée (Bernoulli)	Statistiques locales
Connaissance requise	Distributions a priori partagées	Moyennes locales + bruit
Observabilité des autres joueurs	Requise	Non requise
Complexité analytique	Élevée	Modérée
Applicabilité en ASM	Limitée	Réaliste
Robustesse aux incertitudes	Faible	Élevée

Ainsi, l'approche robuste est plus adaptée aux réseaux sous-marins, car elle ne nécessite pas une connaissance précise des canaux des autres utilisateurs et offre une meilleure résilience face aux incertitudes.

Optimisation de l'allocation de puissance pour chaque utilisateur

Chaque utilisateur i doit déterminer une allocation de puissance optimale \mathbf{p}_i pour maximiser son propre taux de transmission, tout en respectant la contrainte de puissance maximale P_i^{max} . Ce problème d'optimisation peut être formulé comme suit :

$$(Pb_1) \begin{cases} \max_{\mathbf{p}_i} u_i(\mathbf{P}), & \forall i \in \{1, \dots, K\} \\ \text{(s.t)} : & \sum_{n=1}^{N_{sc}} p_{i,n} \leq P_i^{max}, \quad \forall i \in \{1, \dots, K\} \end{cases} \quad (2.14)$$

où K est le nombre total d'utilisateurs, et P_i^{max} est la contrainte de puissance maximale pour l'utilisateur i . L'objectif est de maximiser l'utilité de chaque utilisateur, en tenant compte de l'impact des autres utilisateurs, ce qui rend le problème difficile à résoudre de

manière centralisée.

Le problème de partage des ressources décrit ci-dessus peut être vu comme un jeu non-coopératif \mathcal{G} entre les utilisateurs, qui est défini comme suit :

$$\mathcal{G} = \{\mathcal{N}, \mathcal{P}, u\}, \quad (2.15)$$

où :

- $\mathcal{K} = \{1, \dots, K\}$: ensemble des joueurs (utilisateurs),
- \mathcal{P} : ensemble des stratégies des joueurs, représentant les allocations de puissance possibles,
- $u_i(\mathbf{p}_i, \mathbf{P}_{-i})$: fonction d'utilité de l'utilisateur i , qui dépend de sa propre stratégie d'allocation de puissance \mathbf{p}_i et des stratégies des autres utilisateurs \mathbf{P}_{-i} .

Dans ce jeu non-coopératif, chaque utilisateur choisit une stratégie d'allocation de puissance qui maximise son utilité individuelle, sans coopération avec les autres utilisateurs. Le **NE** de ce jeu est le profil de puissances où aucun utilisateur n'a intérêt à modifier sa stratégie de manière unilatérale.

2.2.5 Condition de **NE**

NE est défini comme suit : un profil de puissances \mathbf{p}_i^* est un **NE** si, pour chaque utilisateur i , sa fonction d'utilité ne peut pas être améliorée en changeant uniquement sa propre stratégie, tout en gardant les stratégies des autres utilisateurs constantes. Mathématiquement, cela peut être formulé par la condition suivante :

$$u_i(\mathbf{p}_i^*, \mathbf{p}_{-i}^*) \geq u_i(\mathbf{p}_i, \mathbf{p}_{-i}^*), \forall i \in \{1, \dots, K\}. \quad (2.16)$$

Cela signifie que l'utilisateur i ne peut pas augmenter son utilité en changeant sa stratégie de manière unilatérale, ce qui définit un équilibre dans lequel chaque utilisateur maximise son utilité tout en tenant compte des décisions des autres.

2.2.6 Meilleures Réponses, Résolution du Jeu et Algorithme d'Allocation de Puissance

Dans un jeu non-coopératif, pour trouver **NE**, il est suffisant de rechercher le profil des meilleures réponses BR_i pour chaque joueur i . Une meilleure réponse est la stratégie qui maximise l'utilité d'un joueur donné, étant donné les stratégies des autres joueurs. En se basant sur la définition de **BR** donnée par l'équation 2.3, on définit la meilleure réponse relative à \mathcal{G} pour l'utilisateur i comme suit :

$$BR_i(\mathbf{P}) = \{\mathbf{p}_i^*, u_i(\mathbf{p}_i^*, \mathbf{P}_{-i}) \geq u_i(\mathbf{p}_i, \mathbf{P}_{-i}), \forall \mathbf{p}_i\}. \quad (2.17)$$

Le profil de **NE** correspond au point où toutes les meilleures réponses se croisent, c'est-à-dire où chaque utilisateur choisit sa meilleure réponse en fonction des autres. Pour résoudre ce jeu non-coopératif, un algorithme d'allocation de puissance basé sur la méthode de water-filling [Cover, 1999] peut être utilisé ce qui permet de déterminer l'allocation de puissance optimale pour chaque utilisateur en maximisant son taux de transmission, tout en tenant compte de l'interférence générée par les autres utilisateurs. Ce mécanisme est

bien adapté pour les problèmes de partage des ressources dans des systèmes de communication multi-utilisateurs.

La prochaine partie proposera une solution pratique pour implémenter cet algorithme et résoudre le problème d'allocation de puissance en utilisant le jeu non-coopératif. Tout d'abord, chaque utilisateur i doit résoudre le problème d'optimisation (Pb_1) afin de maximiser sa fonction d'utilité et trouver la meilleure réponse BR_i . En effet, l'ensemble des BR_i est dérivé à partir de la technique Iterative water filling (*IWF*) (en utilisant les conditions de Karush-Kuhn-Tucker (KKT)), comme cela a déjà été montré dans [Scutari et al., 2008].

Proposition 1. *La solution optimale $p_{i,n}^*$ pour l'utilisateur i sur la $n^{\text{ème}}$ sous-porteuse par rapport au jeu \mathcal{G} et au problème d'optimisation (Pb_1) (voir 2.14) peut être dérivée comme suit :*

$$p_{i,n}^* = \left[\frac{1}{\lambda_i} - \frac{1}{\text{SINR}_{i,n}(\mathbf{P})} \right]^+, \quad (2.18)$$

où

$$\text{SINR}_{i,n}(\mathbf{P}) = \frac{\mathbb{E}[|h_{ii}(n)|^2] p_{i,n}}{\sigma_{i,n}^2 + \sum_{j \neq i} \mathbb{E}[|h_{ji}(n)|^2] p_{j,n}}, \quad (2.19)$$

avec λ_i est le multiplicateur de Lagrange obtenu en satisfaisant la contrainte de puissance (6), et $[a]^+ = \max(a, 0)$.

Preuve :

La solution *IWF* est obtenue en utilisant la méthode du multiplicateur de Lagrange. Pour commencer, nous définissons la fonction lagrangienne $L(p_{i,n}, \lambda_i)$ dérivée du problème d'optimisation comme suit :

$$L(p_{i,n}, \lambda_i) = \sum_{n=1}^{N_{sc}} \log_2(1 + \text{SINR}_{i,n}(\mathbf{P})) - \lambda_i \left(\sum_{n=1}^{N_{sc}} p_{i,n} - P_i^{\max} \right). \quad (2.20)$$

L'allocation de puissance optimale $p_{i,n}$ est dérivée en imposant que la dérivée partielle de la fonction lagrangienne par rapport à $p_{i,n}$ soit nulle :

$$\frac{\partial L(p_{i,n}, \lambda_i)}{\partial p_{i,n}} = 0. \quad (2.21)$$

En prenant la dérivée de la fonction lagrangienne et en résolvant pour $p_{i,n}$, nous obtenons la condition suivante :

$$\frac{\mathbb{E}|h_{ii}(n)|^2}{\sigma_{i,n}^2 + \sum_{j \neq i} \mathbb{E}|h_{ji}(n)|^2 p_{j,n} + \mathbb{E}|h_{ii}(n)|^2 p_{i,n}} - \lambda_i = 0. \quad (2.22)$$

De cette manière, nous pouvons facilement dériver l'expression suivante pour l'allocation de puissance optimale :

$$p_{i,n}^* = \left[\frac{1}{\lambda_i} - \frac{\sigma_{i,n}^2 + \sum_{j \neq i} \mathbb{E}|h_{ji}(n)|^2 p_{j,n}}{\mathbb{E}|h_{ii}(n)|^2} \right]^+. \quad (2.23)$$

Selon l'algorithme *IWF*, au début, tous les utilisateurs répartissent uniformément leur puissance de transmission sur les N_{sc} sous-porteuses. Ensuite, à chaque itération, chaque utilisateur met à jour son profil d'allocation de puissance en fonction des stratégies des

autres utilisateurs, jusqu'à ce que l'algorithme converge vers un **NE** qui doit satisfaire deux conditions essentielles : l'unicité et la convergence de l'algorithme. La démonstration suivante illustre ces deux conditions.

2.2.7 Unicité et convergence vers **NE**

Dans le contexte de l'algorithme **IWF** appliqué à un jeu non coopératif \mathcal{G} de partage de spectre, il est essentiel d'établir des conditions suffisantes assurant l'unicité de **NE** ainsi que la convergence de l'algorithme vers cet équilibre. Ces résultats garantissent la stabilité et la prévisibilité du système décentralisé.

Unicité de **NE** : On définit la matrice $Q_n \in \mathbb{R}^{I \times I}$ associée à la sous-porteuse n comme suit :

$$[Q_n]_{ij} = \begin{cases} \frac{\mathbb{E}[|h_{ji}(n)|^2]}{\exp(\mathbb{E}[\log(|h_{ii}(n)|^2)])} & \text{si } i \neq j \\ 0 & \text{sinon} \end{cases} \quad (2.24)$$

NE du jeu \mathcal{G} est unique si, pour toute sous-porteuse n , le rayon spectral $\rho(Q_n)$ vérifie :

$$\rho(Q_n) < 1 \quad \forall n \in \{1, \dots, N_{sc}\}. \quad (2.25)$$

Interprétation : Cette condition traduit le fait que le niveau d'interférence mutuelle entre les utilisateurs reste suffisamment faible pour empêcher l'apparition de multiples équilibres. Si les liens de communication sont fortement couplés (i.e., $\rho(Q_n) \geq 1$), plusieurs équilibres peuvent exister, dépendant de l'état initial de l'algorithme.

Convergence de l'algorithme : De même, une condition suffisante garantissant la convergence géométrique de l'algorithme **IWF** vers l'unique **NE** est donnée à partir de la matrice suivante :

$$[Q_{\max}]_{ij} = \begin{cases} \max_n \frac{\mathbb{E}[|h_{ji}(n)|^2]}{\exp(\mathbb{E}[\log(|h_{ii}(n)|^2)])} & \text{si } i \neq j \\ 0 & \text{sinon} \end{cases} \quad (2.26)$$

Si $\rho(Q_{\max}) < 1$, alors l'algorithme **IWF** converge géométriquement vers l'unique **NE** du jeu \mathcal{G} .

Conséquence importante : Cette condition de convergence implique également l'unicité de l'équilibre (car $\rho(Q_n) \leq \rho(Q_{\max})$), ce qui renforce sa portée. La preuve de ces deux conditions est illustrée dans cet article [Scutari et al., 2008].

Remarque : Lorsque l'interférence inter-utilisateur augmente (par exemple en raison de la proximité géographique des nœuds ou de canaux fortement corrélés), les conditions ci-dessus peuvent ne plus être satisfaites. Dans ce cas, l'algorithme peut encore converger, mais vers un équilibre potentiellement non unique et l'allocation de puissance converge alors vers une configuration **FDMA** où chaque utilisateur évite les sous-porteuses déjà utilisées par les autres. Cette tendance d'exclusion mutuelle est naturellement induite par la structure du jeu, qui pénalise les configurations fortement conflictuelles.

Dans la section suivante, l'allocation de puissance proposée par rapport au jeu non coopératif est simulée à travers des simulations de Monte-Carlo, en termes d'allocation de puissance distribuée et de région de capacité tout en tenant compte de l'impact du niveau d'interférence.

2.3 Résultats de simulation et analyse

2.3.1 Paramètres et hypothèses de simulation

Dans cette section, on a concentré sur un système composé de deux utilisateurs partageant la même bande passante B . Nous utiliserons MATLAB afin de tester les performances de l'algorithme proposé ainsi que la solution de NE. Le modèle de canal acoustique sous-marin utilisé est basé sur celui présenté dans [Qarabaqi and Stojanovic, 2013]. L'ensemble des paramètres utilisés est présenté dans le tableau suivant.

TABLE 2.2 – Paramètres

Paramètre	Valeur
Profondeur des émetteurs (TX)	65 m
Profondeur des récepteurs (RX)	2 m
Profondeur de l'eau	70 m
Distance $TX1 \rightarrow RX1$	1000 m
Distance $TX2 \rightarrow RX2$	1200 m
Distance entre $TX1$ et $TX2$	150 m
Distance entre $RX1$ et $RX2$	150 m
Nombre d'utilisateurs (K)	2
Bande passante totale	14 kHz
Nombre de sous-porteuses (N_{sc})	800
Espacement entre sous-porteuses (Δf)	17.5 Hz
Puissance maximale par utilisateur	1 W
SNR	20 dB
Nombre d'itérations (L)	15

2.3.2 Jeu à deux joueurs : impact de l'interférence

Dans les communications acoustiques sous-marines, la réponse fréquentielle du canal joue un rôle fondamental dans la conception et l'optimisation des systèmes de transmission. Contrairement aux canaux radio classiques, les canaux acoustiques présentent une sélectivité en fréquence très marquée en raison de la variabilité de la profondeur, de la salinité, de la température et des phénomènes de multi-trajets. Par conséquent, le gain du canal et le niveau d'interférence varient significativement d'une sous-porteuse à une autre.

Dans ce contexte, analyser la réponse en fréquence permet ainsi de mieux comprendre la structure du canal et d'identifier les bandes de fréquences favorables ou défavorables à la transmission pour chaque utilisateur. Cela s'avère particulièrement crucial dans un contexte multi-utilisateur où les ressources doivent être partagées intelligemment pour maximiser le rendement global du système tout en minimisant les interférences. C'est

dans ce cadre que l’observation fine de la réponse fréquentielle guide les décisions d’allocation de puissance dans notre algorithme basé sur NE.

Les canaux directs et d’interférence, notés h_{ii} et h_{ji} avec $i, j \in \{1, 2\}$, sont générés conformément au modèle de canal sous-marin proposé dans [Qarabaqi and Stojanovic, 2013] et la réponse fréquentielle de ces canaux est représentée sur la Fig. 2.3. La réponse en fréquence du canal (Fig. 2.3) permet d’extraire des informations précieuses concernant l’allocation optimale des ressources (sous-porteuses). On observe clairement que les gains des utilisateurs ainsi que le niveau d’interférence varient significativement d’une sous-porteuse à une autre parmi les $N_{sc} = 800$ disponibles. Cette variabilité justifie l’approche de type jeu non coopératif où chaque joueur adapte son comportement.

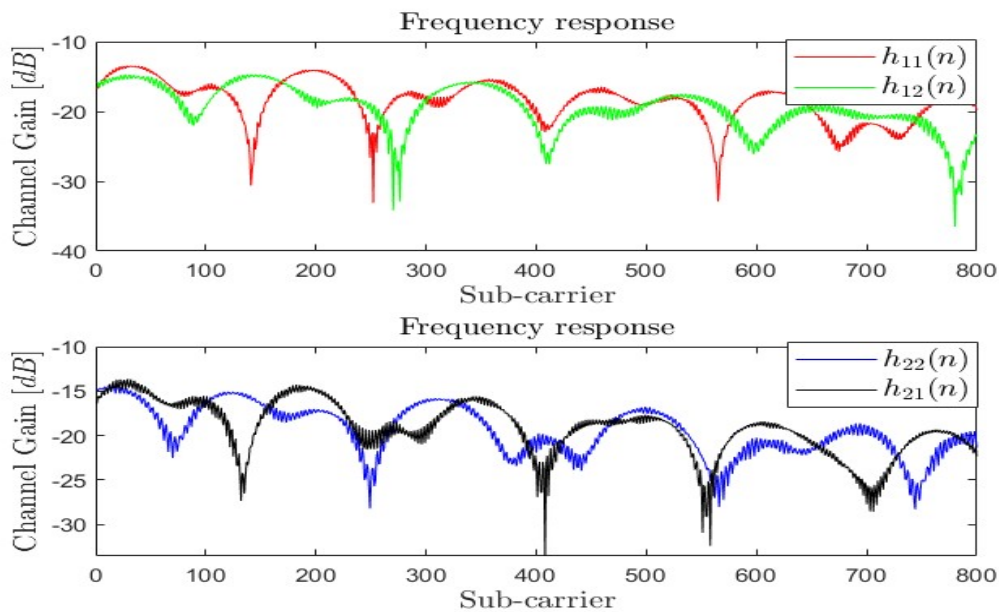


FIGURE 2.3 – Réponse fréquentielle des canaux directs et d’interférences simulés selon le modèle de [Qarabaqi and Stojanovic, 2013].

La Figure 2.4 illustre la distribution de puissance résultante, qui est la stratégie d’Équilibre de Nash atteinte par l’algorithme IWF (Eq. 2.18). Cette allocation est une réponse directe aux conditions de canal de la Figure 2.3 : un joueur n’alloue de la puissance que sur les sous-porteuses où son gain de canal direct est suffisamment élevé par rapport à l’interférence qu’il reçoit.

On peut identifier trois scénarios stratégiques clairs en comparant les deux figures :

- **Cas 1 : Allocation unilatérale (Joueur 1)** : Sur les sous-porteuses $n \approx 230$ à $n = 260$, la Fig. 2.4 montre que le Joueur 1 ($p_{1,n}$) alloue sa puissance maximale, tandis que le Joueur 2 ($p_{2,n}$) n’alloue rien. La Fig. 2.3 l’explique :
 - **Joueur 1 (transmet)** : Il bénéficie d’un bon canal direct ($h_{11} \approx -13$ dB, rouge) et d’une interférence reçue très faible ($h_{21} \approx -30$ dB, noire).
 - **Joueur 2 (se tait)** : Il subit un mauvais canal direct ($h_{22} \approx -20$ dB, bleu) et une interférence potentielle très élevée ($h_{12} \approx -15$ dB, verte). Transmettre serait inefficace.

- **Cas 2 : Allocation unilatérale (Joueur 2) :** La situation inverse s’observe sur les sous-porteuses $n \approx 600$ à $n = 630$. Ici, $p_{2,n}$ est maximale et $p_{1,n}$ est nulle.
 - **Joueur 2 (transmet) :** Il possède un bon canal direct ($h_{22} \approx -15$ dB, bleu) et une faible interférence reçue ($h_{12} \approx -25$ dB, verte).
 - **Joueur 1 (se tait) :** Il subit un mauvais canal direct ($h_{11} \approx -20$ dB, rouge) et une forte interférence ($h_{21} \approx -15$ dB, noire).
- **Cas 3 : Compétition (Allocation des deux joueurs) :** Enfin, sur des plages comme $n \approx 28$ à $n = 62$, les deux joueurs allouent une puissance élevée. Cela se produit car les deux joueurs ont des conditions favorables : leurs gains directs (h_{11} et h_{22}) sont tous deux élevés (proches de -12 dB) et les interférences mutuelles (h_{12} et h_{21}) sont gérables (autour de -15 dB). Les deux joueurs luttent donc pour ces sous-porteuses.

Ainsi, la distribution de puissance non uniforme de la Fig. 2.4 n’est pas aléatoire, mais une illustration directe de l’équilibre de jeu non coopératif où chaque joueur exploite les sous-porteuses qui maximisent sa propre utilité en fonction des conditions de canal et des actions de son adversaire.

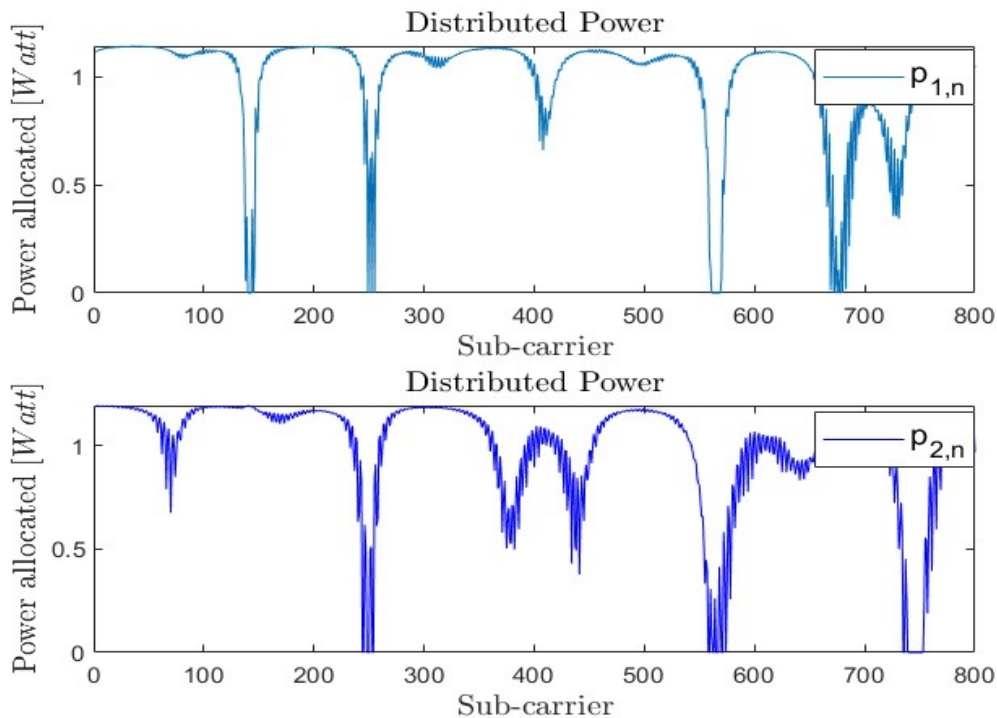


FIGURE 2.4 – Répartition de puissance sur les sous-porteuses pour 2 utilisateurs obtenue par l’algorithme itératif water-filling dans un jeu non coopératif.

Dans les réseaux sous-marins multi-utilisateurs, la distance entre les paires émetteur-récepteur joue un rôle fondamental sur le niveau d’interférence mutuelle. Cette interférence conditionne à la fois la performance du système (à travers la fonction d’utilité) et la convergence des algorithmes d’allocation de ressources ce qui est illustré dans la partie suivante.

2.3.3 Impact de la distance sur la fonction d'utilité et la région de capacité

D'après les résultats théoriques présentés dans [Pottier et al., 2016], lorsque les joueurs sont suffisamment éloignés et que l'interférence est minimisée, la solution de l'algorithme converge vers un NE unique. Afin de valider expérimentalement cette propriété, nous étudions dans cette section l'évolution de la fonction d'utilité pour différents niveaux d'interférence, liés à la distance entre les paires émetteur-récepteur.

Trois cas sont considérés et illustrés dans la Fig. 2.5, où chaque scénario correspond à une configuration particulière de distance inter-paires :

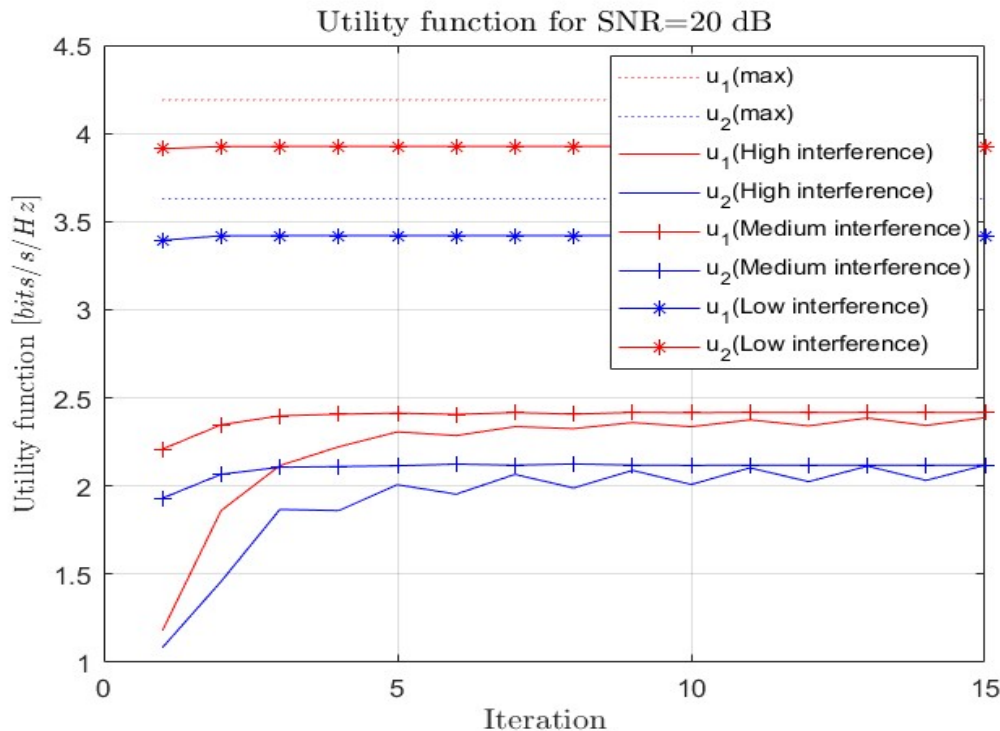


FIGURE 2.5 – Évolution de la fonction d'utilité de chaque joueur au cours des 30 itérations, selon la distance entre les paires.

- **Cas d'interférence élevée** : Ce scénario correspond à la configuration par défaut décrite précédemment, où la distance entre les paires émetteur-récepteur est fixée à 150 m. Dans ce cas, les puissances moyennes des canaux directs et des canaux d'interférence sont du même ordre de grandeur. Une telle proximité physique génère un fort couplage entre les deux joueurs, ce qui se traduit par une forte interférence mutuelle sur l'ensemble des sous-porteuses. En conséquence, la fonction d'utilité évolue lentement et reste relativement faible, même après plusieurs itérations. On observe une saturation précoce de la performance, indiquant que les joueurs atteignent un NE possible, mais non optimal. Ce type d'équilibre peut être multiple, car les stratégies ne sont pas suffisamment disjointes pour éliminer les effets d'interférence. On note également que l'allocation initiale uniforme (à la première itération) conduit à une performance très médiocre, démontrant la nécessité

d'une allocation adaptative.

- **Cas d'interférence modérée** : Ici, la distance est augmentée à 1500 m, ce qui permet une réduction significative de l'interférence, bien que cette dernière soit encore présente. La fonction d'utilité de chaque joueur augmente de manière plus marquée par rapport au premier cas. Cependant, comme le niveau d'interaction entre les stratégies demeure non négligeable, la convergence vers un **NE** unique n'est pas assurée. On observe des variations dans la dynamique d'apprentissage, révélant un comportement instable dans certaines zones du spectre. Néanmoins, ce scénario montre une amélioration notable des performances globales.
- **Cas d'interférence faible** : Dans ce dernier cas, la distance entre les paires est portée à 4500 m. La grande séparation spatiale permet aux joueurs d'évoluer quasiment de manière indépendante. L'interférence mutuelle devient quasi nulle, ce qui favorise une convergence rapide et stable vers un **NE** unique. Les courbes d'utilité montrent une augmentation abrupte dès les premières itérations, avec une saturation à des niveaux très élevés, reflétant une exploitation optimale du spectre. Ce cas représente une situation idéale où la compétition entre les utilisateurs est faible, et où l'allocation spectrale s'effectue sans conflit.

Pour compléter cette analyse, nous représentons la **région de capacité** (Capacity Region (CR)) associée à ces trois scénarios dans la Fig. 2.6. La région de capacité décrit l'ensemble des taux de transmission r_1 , r_2 que les deux joueurs peuvent atteindre simultanément sous contrainte d'interférence :

$$r_1 \leq \log_2(1 + \text{SNR}_1) = u_1 \quad (2.27)$$

$$r_2 \leq \log_2(1 + \text{SNR}_2) = u_2 \quad (2.28)$$

où u_1 et u_2 représentent les capacités maximales atteignables respectivement par le joueur 1 et le joueur 2, en l'absence d'interférence. Ces taux dépendent directement du niveau d'interférence subi.

Dans la Fig. 2.6, chaque courbe délimite la région de capacité correspondant à un certain niveau d'interférence :

- le point (**x**) indique le NE obtenu dans le cas d'interférence élevée ;
- le point (+) correspond à un NE sous interférence modérée ;
- le point (**o**) montre le NE atteint en cas d'interférence faible.

On y ajoute également une courbe de référence représentant une stratégie **TDMA**, dans laquelle chaque utilisateur occupe l'ensemble du spectre pendant une fraction de temps exclusive. Cette stratégie élimine l'interférence, mais au prix d'une occupation partielle du temps.

Tous les **NE** présentés ici sont obtenus par la stratégie d'allocation d'eau dite **Water-Filling**, représentée par l'équation (2.18). On constate que : lorsque l'interférence est élevée, la région de capacité est **fortement réduite** : les joueurs sont contraints par la compétition, à l'inverse, lorsque l'interférence est faible, la région de capacité s'agrandit et tend vers son enveloppe convexe, permettant d'atteindre des débits proches de l'optimum.

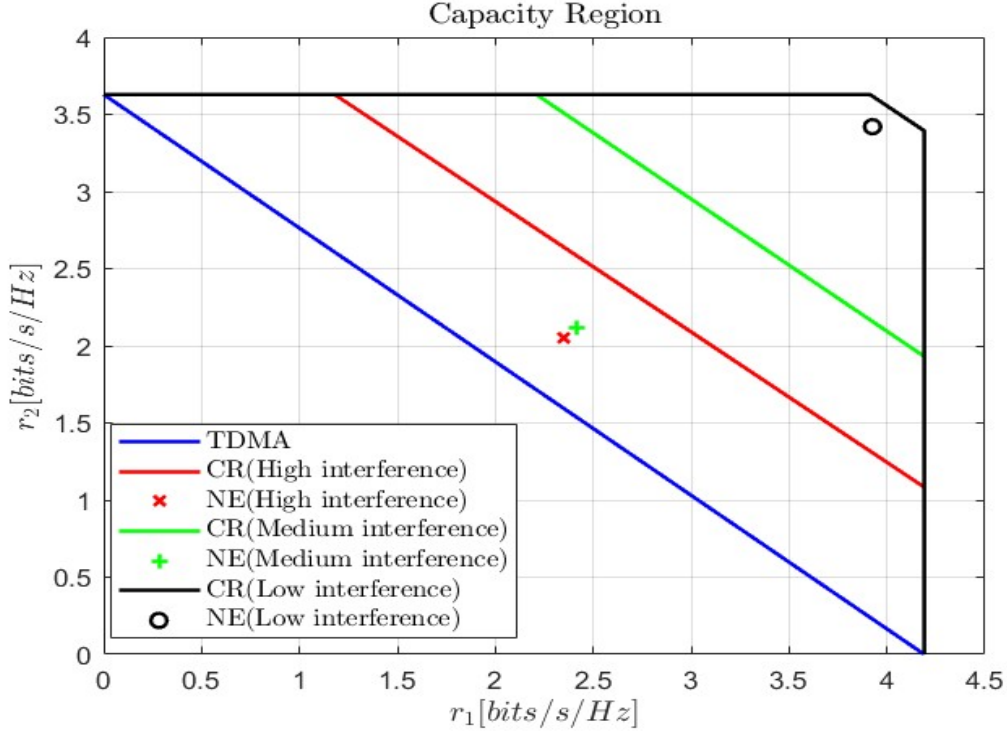


FIGURE 2.6 – Régions de capacité et points de NE en fonction du niveau d’interférence.

Ces résultats démontrent l’importance d’une gestion spatiale des transmissions dans les réseaux sous-marins, et motivent l’intégration de techniques adaptatives capables d’exploiter les propriétés physiques du canal, en fonction des positions relatives des nœuds. Tous les résultats présentés ici sont détaillés et expliqués dans notre article [Gharsalli et al., 2023b].

2.4 Conclusion

En conclusion, ce chapitre a montré que l’allocation décentralisée de puissance, fondée sur la théorie des jeux non-coopératifs et la modulation OFDM, constitue une approche efficace pour optimiser la performance individuelle des utilisateurs dans un canal acoustique sous-marin complexe et variable. Notre travail s’est appuyé sur les algorithmes d’allocation proposés dans [Pottier, 2018], utilisés comme base méthodologique. La principale contribution a consisté à faire varier le niveau d’interférence entre utilisateurs et à calculer la région de capacité du système, ce qui a permis d’évaluer précisément l’impact des interférences mutuelles sur les performances globales.

Les résultats ont mis en évidence que, malgré l’optimisation individuelle obtenue grâce à la stratégie de type Water-Filling, l’absence de coordination et de partage d’informations entre les joueurs limite la performance globale et ne garantit pas une allocation équitable des ressources. Ces constats soulignent l’importance de considérer des approches coopératives ou hybrides, permettant une meilleure coordination entre utilisateurs. En intégrant le partage d’informations sur l’état du canal et les stratégies d’allocation, de tels modèles coopératifs ouvrent la voie à des communications sous-marines plus fiables, robustes et équitables, tout en préservant les avantages de la modulation OFDM et de l’adaptation

des ressources.

Chapitre 3

Stratégies d'allocation coopérative basées sur la négociation dans les CASM

3.1 Introduction

La théorie des jeux coopératifs étudie les situations dans lesquelles plusieurs acteurs, appelés joueurs ou agents, choisissent de collaborer afin d'atteindre des objectifs communs ou de maximiser leurs bénéfices mutuels. Contrairement aux approches non coopératives où chaque joueur agit de manière autonome et stratégique pour optimiser son propre intérêt, l'approche coopérative s'intéresse aux scénarios où la coordination et l'échange des informations entre les parties sont possibles et nécessaires.

Dans de nombreux environnements, qu'il s'agisse de réseaux de communication, de systèmes économiques, de négociations politiques ou de plateformes collaboratives, la coopération permet souvent d'obtenir de meilleurs résultats que l'action individuelle [Chatain, 2016]. Dans un environnement acoustique, cette coopération est particulièrement pertinente car les ressources disponibles telles que la bande passante, la puissance d'émission, ou les canaux de communication sont limitées et sujettes à des interférences. En coordonnant les décisions entre les utilisateurs (par exemple, capteurs ou nœuds de réseau), il est possible d'optimiser l'utilisation du spectre, de réduire les interférences, et d'améliorer la qualité de la transmission. Ainsi, la collaboration permet d'exploiter au mieux les ressources partagées dans un milieu acoustique souvent contraint, cependant, cette coopération soulève une question essentielle : comment répartir équitablement les bénéfices obtenus grâce à la collaboration ?

Dans le contexte spécifique des communications *ASM*, l'application de la théorie des jeux coopératifs reste encore peu explorée. La majorité des travaux s'appuie principalement sur des approches non coopératives ou hiérarchiques, telles que les jeux de Stackelberg pour le contrôle de puissance et l'optimisation énergétique [Luo and Wang, 2025]. Toutefois, quelques contributions récentes introduisent explicitement des mécanismes coopératifs. Par exemple, [Muhammed et al., 2018] propose une revue des techniques issues de la théorie des jeux appliquées aux réseaux de capteurs acoustiques sous-marins, incluant des modèles coopératifs et de négociation. De manière plus appliquée, [Ye et al., 2024] exploitent la coopération pour la sélection de relais afin d'améliorer la qualité de service

dans les réseaux [ASM](#), tandis que [[Pan et al., 2024](#)] s'intéressent à la coopération pour le brouillage spatio-temporel sécurisé à travers une sélection conjointe des nœuds et une allocation optimale de puissance.

Bien qu'il existe quelques travaux consacrés aux jeux coopératifs en [ASM](#), la littérature reste limitée et fragmentée. Cela met en évidence l'intérêt de poursuivre l'étude de mécanismes coopératifs et équitables pour la gestion des ressources dans ces environnements contraints, afin de modéliser les interactions entre les utilisateurs, d'identifier les conditions de coopération stable et de proposer des solutions efficaces pour le partage des ressources, permettant ainsi d'obtenir des gains significatifs. La théorie des jeux coopératifs offre également un cadre rigoureux pour analyser les relations de coopération entre agents rationnels et évaluer les mécanismes d'entente, d'échange ou de négociation susceptibles d'émerger. Il existe par ailleurs différentes formes de jeux coopératifs, selon la nature de l'interaction entre les joueurs, le type d'information disponible ou encore les modalités de coopération.

On peut distinguer plusieurs grandes catégories :

- **Jeux avec transfert d'utilité (TU)** : Sont des jeux coopératifs dans lesquels les bénéfices issus de la coopération peuvent être librement transférés et redistribués entre les joueurs, souvent sous forme monétaire ou de ressources quantifiables. Cette flexibilité facilite la formation d'accords et l'analyse du partage des gains, comme détaillé par ce travail de [[Kranich et al., 2005](#)] sur les concepts centraux des jeux dynamiques TU.
- **Jeux sans transfert d'utilité (NTU)** : Concernent des situations où les gains ne sont pas directement échangeables, soit en raison de préférences non monétaires, soit à cause de contraintes contextuelles qui limitent la transférabilité. Ces jeux nécessitent des méthodes d'analyse spécifiques qui tiennent compte des préférences multidimensionnelles des joueurs, comme l'explique cette étude approfondie des jeux NTU [[Peters, 2009](#)].
- **Jeux à information complète** : Sont caractérisés par le fait que tous les joueurs connaissent parfaitement les paramètres du jeu, notamment les stratégies possibles, les gains associés et les préférences des autres joueurs. Cette connaissance totale permet d'analyser précisément les interactions stratégiques et de déterminer des équilibres stables, comme l'a exposé Gibbons dans son introduction aux fondements de la théorie des jeux [[Gibbons, 1997](#)].
- **Jeux à information incomplète** : Certains éléments essentiels du jeu restent inconnus, tels que les préférences ou les types des adversaires. Cette incertitude oblige les joueurs à former des croyances probabilistes, souvent modélisées par des approches bayésiennes, et conduit à des solutions telles que l'équilibre bayésien. Cet article [[Ausubel et al., 2002](#)] illustre bien comment ces aspects influencent les stratégies de négociation dans des contextes d'information imparfaite.
- **Jeux statiques** : Ce sont des jeux dans lesquels tous les joueurs prennent leurs décisions simultanément ou, du moins, sans connaître les choix des autres. La co-

opération ou la compétition se déroule en une seule étape, sans évolution dans le temps ni possibilité de modifier les stratégies après observation des actions adverses [Benchekroun and Van Long, 2011].

- **Jeux dynamiques** : Au contraire, ces jeux modélisent des situations où les décisions des joueurs s'étendent dans le temps. Les joueurs peuvent prendre leurs décisions de manière séquentielle, ajuster leurs stratégies en fonction des actions passées des autres, ou en fonction de l'évolution d'un état dynamique [Avni et al., 2016]-[Benchekroun and Van Long, 2011].
- **Jeux centralisés** : La coordination des décisions ainsi que la répartition des ressources sont assurées par un mécanisme ou une entité centrale qui supervise et optimise la coopération entre les agents. Ce cadre facilite souvent l'allocation optimale des ressources, comme illustré par Fang, Li dans son étude sur l'allocation centralisée basée sur une approche de négociation [Fang et al., 2021].
- **Jeux décentralisé** : Reposent sur une prise de décision autonome des agents, qui coopèrent localement sans intervention d'une autorité centrale. Ces agents négocient ou s'auto-organisent pour parvenir à des accords, ce qui rend ces systèmes plus flexibles et adaptés à des environnements distribués, comme démontré par [Sharma and Teneketzis, 2011] dans le contexte de l'allocation optimale de puissance dans les réseaux cellulaires.
- **Jeux de négociation** : Ces jeux s'intéressent aux mécanismes par lesquels les agents parviennent à un accord sur la manière de coopérer et de répartir les bénéfices issus de leur collaboration. Contrairement aux modèles où la coopération est imposée ou déterminée à l'avance, les jeux de négociation modélisent un processus interactif, dans lequel les parties échangent, argumentent et ajustent leurs positions pour parvenir à un compromis mutuellement acceptable [Zhang et al., 2015]-[Madani et al., 2022]. Ils reposent sur des notions fondamentales telles que l'équité, la rationalité des agents et l'efficacité des accords, des concepts qui seront détaillés et approfondis dans les sections suivantes.

3.2 Modélisation coopérative basée sur la négociation de Nash

3.2.1 Définition, hypothèses et contraintes

Définition : Dans les systèmes de communication multi-utilisateurs, notamment dans des environnements contraints tels que les réseaux acoustiques sous-marins ou les réseaux sans fil partagés, les utilisateurs doivent se partager un ensemble limité de ressources (bande passante, puissance, sous-porteuses, etc.). L'approche coopérative s'impose comme une solution pertinente lorsque les utilisateurs peuvent communiquer, échanger des informations, et coordonner leurs actions afin d'optimiser collectivement leur performance. Dans ce contexte, la théorie des jeux coopératifs fournit un cadre rigoureux pour modéliser ces interactions et proposer des solutions équitables et efficaces.

Parmi les solutions classiques de la théorie des jeux coopératifs, le modèle de négociation de Nash ou Nash Bargaining (NB) en anglais occupe une place importante, ce type est proposé par John Nash en 1950 [Nash et al., 1950], fournit un cadre mathématique pour résoudre les conflits d'intérêts entre plusieurs parties souhaitant parvenir à un accord mutuellement bénéfique. La solution de Nash repose sur un certain nombre d'axiomes (efficacité, symétrie, invariance par transformation affine, indépendance des alternatives non pertinentes) qui permettent de garantir une solution unique et équitable [Vegas Segura, 2020].

L'idée centrale repose sur un compromis entre équité et efficacité : la solution de négociation doit profiter à tous les participants, sans favoriser excessivement un joueur au détriment des autres.

Dans le contexte des systèmes multi-agents ou des réseaux de communication, la solution de NB peut être utilisée pour résoudre des problèmes de partage de ressources, comme l'allocation de puissance, de bande passante ou de temps d'accès, de manière collaborative. L'idée clé est de maximiser le produit des utilités supplémentaires que chaque agent retire de la coopération, relativement à un point de désaccord (ou d'isolement), appelé disagreement point. Ce point représente la performance minimale que chaque joueur peut garantir sans coopération.

Mathématiquement, la solution de négociation de Nash est définie comme suit :

$$\text{NBS} = \arg \max_{u \in U} \prod_{i=1}^N (u_i - d_i). \quad (3.1)$$

où U est l'ensemble des utilités accessibles par coopération, u_i est l'utilité obtenue par le joueur i , et d_i représente son utilité en cas d'absence de coopération.

Cette approche est particulièrement pertinente dans les environnements où les utilisateurs sont capables de communiquer ou de coordonner leurs actions, mais restent autonomes et rationnels. Le produit présent dans la formule de Nash Bargaining assure une répartition équitable des gains, car il cherche à maximiser simultanément les bénéfices de chaque partie. C'est cette propriété qui garantit l'équité du résultat final tout en tenant compte les incitations individuelles.

Hypothèses et contraintes appliquées : Pour notre cas, on considère le même modèle proposé dans le chapitre précédent, mais ici le problème d'allocation de puissance est définie en tant qu'un jeu de négociation coopératif afin d'étudier l'existence de la solution de Nash Bargaining par rapport à ce jeu.

On suppose dans ce cas que :

- Le préfixe cyclique T_g est supérieur à l'étalement de retard du canal : cela garantit que les symboles OFDM ne subissent pas d'ISI, ce qui permet une réception correcte même en présence de multi-trajets.
- L'espacement entre les porteuses Δf est supérieur à la largeur de bande Doppler maximale du canal : cette condition assure l'orthogonalité entre les sous-porteuses, même en présence de variations de canal dues à la mobilité ou aux fluctuations environnementales (effet Doppler).

- Les utilisateurs (joueurs) peuvent coopérer et coordonner leurs actions ainsi que leurs allocations de ressources : cette hypothèse est essentielle dans le cadre d’une approche coopérative telle que la négociation de Nash, où les décisions sont prises collectivement pour maximiser les gains de chacun.
- Les utilisateurs connaissent seulement les statistiques des canaux [ASM](#) qu’elles varient beaucoup plus lentement.
- Il existe des contraintes sur la puissance de transmission et un débit minimal à respecter : chaque utilisateur doit satisfaire un seuil minimal de qualité de service (QoS), ce qui impose des contraintes sur les ressources qu’il peut allouer.
- Une sous-porteuse $n \in N_{sc}$ peut être partagée entre plusieurs utilisateurs : contrairement aux schémas d’accès exclusifs, cette hypothèse permet une superposition des transmissions sur une même sous-porteuse, ce qui reflète des scénarios de partage spectral souple.

3.2.2 Modèle mathématique et axiomes de Nash

Formellement, un **jeu de négociation à K joueurs** est défini par une paire $(\mathbf{R}, \mathbf{R}^{\min})$, où :

- $\mathbf{R} = (R_1, \dots, R_K)$ représente l’ensemble faisable : l’ensemble des utilités que les joueurs peuvent atteindre via la coopération, compte tenu des ressources disponibles.
- $\mathbf{R}^{\min} = (R_1^{\min}, \dots, R_K^{\min})$ est le point de désaccord : il représente le gain minimum garanti à chaque utilisateur en cas d’absence de coordination (c’est-à-dire si aucune coopération n’est trouvée).

La solution de [NB](#) est alors définie comme l’unique point $\mathbf{R}^* = (R_1^*, \dots, R_K^*)$ qui maximise le produit des gains relatifs, mesurés par rapport au point de désaccord :

$$\mathbf{R}^* = \arg \max_{\mathbf{R} \in \mathcal{R}} \prod_{i=1}^K (R_i - R_i^{\min}). \quad (3.2)$$

La solution de Nash repose sur quatre axiomes fondamentaux qui permettent de garantir son unicité, sa pertinence et son équité. Ces propriétés sont essentielles dans le contexte des jeux coopératifs, où les utilisateurs cherchent une solution bénéfique et équitable à la fois.

- **(A1) Invariance par transformation affine** : La solution de négociation est invariante vis-à-vis de toute transformation affine (changement d’échelle ou de décalage) des fonctions d’utilité :

$$\text{Si } \hat{R}_i = a_i R_i + b_i, \quad \text{avec } a_i > 0, b_i \in \mathbb{R}, \quad \forall i. \quad (3.3)$$

La solution de négociation pour les utilités transformées $\hat{R}_i = \{\hat{R}_1, \dots, \hat{R}_K\}$ correspond à l’image affine de la solution originale. Cela signifie que la solution ne dépend pas de l’unité de mesure ou de l’échelle choisie pour exprimer les gains des utilisateurs.

- **(A2) Symétrie** : Si les joueurs sont interchangeable, c'est-à-dire qu'ils ont des ensembles de stratégies identiques et des fonctions d'utilité symétriques, alors la solution doit leur attribuer des gains égaux :

$$\text{Si } \mathbf{R} \text{ et } \mathbf{R}^{\min} \text{ sont symétriques, alors } R_i^* = R_j^* \quad \forall i, j. \quad (3.4)$$

Cela garantit l'équité dans le traitement des utilisateurs lorsque leurs conditions sont similaires.

- **(A3) Indépendance des alternatives non pertinentes (Independence of Irrelevant Alternatives)** : Si une solution est obtenue à partir d'un ensemble faisable \mathcal{R} , alors restreindre l'ensemble faisable à une partie $\mathcal{R}' \subseteq \mathcal{R}$ qui contient cette solution ne doit pas modifier la solution :

$$\text{Si } \mathbf{R}^* \in \mathcal{R}' \subseteq \mathcal{R}, \text{ alors } \mathbf{R}^* \text{ reste la solution optimale sur } \mathcal{R}'. \quad (3.5)$$

Ce principe évite que des options inutiles (non optimales) puissent affecter la décision finale.

- **(A4) Efficacité de Pareto** : La solution de Nash doit se trouver sur la frontière de Pareto de l'ensemble faisable \mathcal{R} : aucun utilisateur ne peut améliorer son gain sans détériorer celui d'un autre. Mathématiquement :

$$\nexists \mathbf{R}' \in \mathcal{R} \text{ tel que } R'_i \geq R_i^* \quad \forall i \text{ et } \exists j : R'_j > R_j^*. \quad (3.6)$$

Cela garantit l'efficacité globale du système : aucune ressource n'est gaspillée, et tous les utilisateurs atteignent un compromis optimal.

3.2.3 Application à l'allocation de ressources par un jeu coopératif

Dans le cadre de notre système à K utilisateurs, chaque utilisateur i cherche à maximiser un taux de transmission $R_i(\mathbf{P})$ qui dépend de sa propre allocation de puissance sur les N_{sc} sous-porteuses, représentée par le vecteur $\mathbf{P} = \{p_{i,n}\}$.

L'objectif est alors de résoudre le problème suivant, où l'on cherche la meilleure répartition de la puissance entre les utilisateurs, selon la solution de négociation de Nash :

$$(Pb_1) : \quad \max_{\mathbf{P}} U(\mathbf{P}) = \max_{\mathbf{P}} \left(\prod_{i=1}^K [R_i(\mathbf{P}) - R_i^{\min}(\mathbf{P})]^{1/K} \right). \quad (3.7)$$

Sous les contraintes suivantes :

$$\begin{cases} \sum_{n=1}^{N_{sc}} p_{i,n} \leq P^{\max}, & \forall i \in \{1, \dots, K\} \quad (\mathbf{C}_1) \quad (5.a) \\ R_i(\mathbf{P}) \geq R_i^{\min}(\mathbf{P}), & \forall i \in \{1, \dots, K\} \quad (\mathbf{C}_2) \quad (5.b) \end{cases} \quad (3.8)$$

- La contrainte (\mathbf{C}_1) impose une limite sur la puissance totale allouée par chaque utilisateur à l'ensemble des sous-porteuses.
- La contrainte (\mathbf{C}_2) garantit que chaque utilisateur obtienne un taux au moins égal à son minimum garanti.

3.2.4 Hypothèse simplificatrice

Dans notre approche, nous considérons une hypothèse simplificatrice : chaque utilisateur a un point de désaccord nul, i.e. $R_i^{\min} = 0$. Cela signifie qu'en l'absence de coopération, l'utilisateur ne bénéficie d'aucune ressource. Ce choix permet de :

- réduire la complexité de modélisation ;
- favoriser une solution pleinement coopérative où les utilisateurs cherchent à maximiser uniquement l'utilité globale ;
- refléter des cas concrets où l'accès aux ressources est impossible ou inefficace sans coordination (ex : environnement hostile, interférences élevées).

Ainsi, le problème se réduit à maximiser la fonction :

$$U(\mathbf{P}) = \prod_{i=1}^K R_i(\mathbf{P})^{1/K}. \quad (3.9)$$

Ce qui correspond à une maximisation du produit géométrique des taux, critère bien connu pour refléter une allocation équitable et efficiente entre les utilisateurs.

3.2.5 Théorie de la négociation de Nash : cas de deux utilisateurs

La théorie de la négociation de Nash constitue une approche pertinente pour parvenir à une allocation équitable et efficace des ressources limitées, telles que les sous-porteuses et la puissance de transmission. Le **protocole de négociation proposé** repose sur une stratégie itérative dans laquelle chaque utilisateur ajuste progressivement sa propre stratégie (allocation de puissance et de sous-porteuses), dans le but de maximiser son utilité tout en tenant compte de l'impact de ses décisions sur l'autre utilisateur.

Principe général de la négociation : Chaque utilisateur dispose d'une utilité associée à la qualité de sa transmission, qui dépend de son taux de transmission effectif, des conditions de canal, et du niveau d'interférence reçu. L'objectif est de trouver un compromis (ou point de négociation) qui maximise l'écart entre le débit R_i de chaque utilisateur i et son débit minimal R_i^{\min} , requis pour achever une certaine qualité de service.

On définit la fonction d'utilité de Nash suivante, pour un système à deux utilisateurs et une stratégie de puissance \mathbf{P} :

$$U(\mathbf{P}) = \prod_{i=1}^2 (R_i - R_i^{\min}), \quad (3.10)$$

avec

$$R_i(\mathbf{P}) = \frac{1}{N_{sc}} \sum_{n \in \mathcal{N}_i} \log_2 (1 + \text{SINR}_{i,n}(\mathbf{P})), \quad (3.11)$$

où, \mathcal{N}_i représente les sous-porteuses attribuées à l'utilisateur i après coopération, $\mathbf{P} = (P_{ij})_{\substack{1 \leq i \leq K \\ 1 \leq j \leq K}}$ et SINR est défini par l'équation suivante :

$$\text{SINR}_{i,n}(\mathbf{P}) = \frac{\mathbb{E}[|h_{ii}(n)|^2] p_{i,n}}{\sum_{j \neq i} \mathbb{E}[|h_{ji}(n)|^2] p_{j,n} + \sigma_{i,n}^2}. \quad (3.12)$$

Étapes d’algorithme de négociation pour deux utilisateurs : En se basant sur la fonction d’utilité précédente 3.10, l’algorithme de négociation de Nash proposé est présenté en détail en annexe (voir Algorithme 1) et repose sur un processus itératif de négociation et d’échange, dans lequel les utilisateurs proposent et ajustent dynamiquement leurs allocations de puissance sur les sous-porteuses disponibles. Ce processus itératif vise à améliorer la fonction d’utilité de chaque utilisateur ainsi que l’utilité globale du système avec une meilleure exploitation des ressources disponibles.

L’algorithme commence par une phase d’initialisation qui établit les paramètres nécessaires à la simulation, avant de mettre en œuvre une technique adaptée d’allocation des sous-porteuses. En fonction du nombre d’offres autorisées N_{offer} , des propositions d’allocation sont générées entre les utilisateurs ; chaque utilisateur peut accepter ou refuser les propositions reçues.

Pour les sous-porteuses communes proposées N_{Csc} , jugées également importantes (c’est-à-dire ayant un SNR élevé) pour tous les utilisateurs, celles-ci peuvent être partagées entre eux selon des poids, comme indiqué dans l’équation ??, afin de déterminer la proportion de sous-porteuses allouées à chaque utilisateur lors de la distribution des ressources. Ce poids α_i dépend du gain de canal de l’utilisateur associé à chaque sous-porteuse proposée. Quant aux sous-porteuses restantes, jugées un peu moins avantageuses, elles sont distribuées une par une entre les utilisateurs, en commençant par celui qui a initié la proposition. Les dernières sont réparties de manière équitable entre tous les utilisateurs, en commençant par celui ayant le gain de canal le plus faible.

Un aspect crucial de l’algorithme est l’ajustement dynamique des allocations de puissance via la fonction d’actualisation de la puissance, qui adapte les puissances selon les propositions finales de chaque utilisateur, tout en tenant compte des niveaux d’interférence. Enfin, les débits de chaque utilisateur ainsi que la fonction d’utilité globale sont calculés.

Pour illustrer le déroulement de l’algorithme décrit dans l’annexe (Algorithme 1) et en faciliter la compréhension, nous proposons un exemple pratique impliquant deux utilisateurs et un nombre restreint de sous-porteuses. Cette configuration permet d’observer les étapes clés de la négociation et la logique de répartition des ressources.

Exemple illustratif de négociation : deux utilisateurs, 4 sous-porteuses

Considérons un scénario simplifié avec deux utilisateurs U_1 et U_2 et $N_{\text{sc}} = 4$ sous-porteuses. Les gains de canal directs moyens pour chaque utilisateur sont donnés dans le tableau suivant :

TABLE 3.1 – Gains de canal directs moyens pour chaque sous-porteuse

Sous-porteuse	$\mathbb{E}[h_{11,n} ^2]$ (Utilisateur 1)	$\mathbb{E}[h_{22,n} ^2]$ (Utilisateur 2)
1	0.9	0.4
2	0.7	0.6
3	0.5	0.8
4	0.3	0.9

Étape 1 : Détection des sous-porteuses communes importantes On considère qu’une sous-porteuse est « importante » pour un utilisateur si son gain moyen est supé-

rieur à 0.6. Ainsi :

- U_1 considère importantes : sous-porteuses 1 et 2.
- U_2 considère importantes : sous-porteuses 2, 3 et 4.
- Sous-porteuse commune importante : **sous-porteuse 2**.

Étape 2 : Partage pondéré de la sous-porteuse commune (sous-porteuse 2) :

Pour la sous-porteuse 2, le partage entre les utilisateurs est calculé sur la base des moyennes statistiques des gains de canal. On définit :

$$\alpha_1 = \frac{\mathbb{E}[|h_{11,2}|^2]}{\mathbb{E}[|h_{11,2}|^2] + \mathbb{E}[|h_{22,2}|^2]} = \frac{0.7}{0.7 + 0.6} \approx 0.538, \quad (3.13)$$

d'où

$$\alpha_2 = 1 - \alpha_1 \approx 0.462. \quad (3.14)$$

Ainsi, la sous-porteuse 2 est partagée entre U_1 et U_2 selon ces poids.

Étape 3 : Attribution des sous-porteuses restantes :

Les sous-porteuses restantes (1, 3, 4) sont attribuées comme suit :

- la sous-porteuse 1 est fortement favorable à U_1 → allouée à U_1 ;
- la sous-porteuse 4 est très favorable à U_2 → allouée à U_2 ;
- la sous-porteuse 3 : partagée ou allouée à U_2 (gain plus élevé).

Étape 4 : Résumé de l'allocation finale :

Après la phase de négociation et d'évaluation des gains de canal, les sous-porteuses ont été attribuées de la manière suivante :

TABLE 3.2 – Allocation finale des sous-porteuses

Sous-porteuse	Utilisateur
1	U_1
2	Partagée ($\alpha_1 = 0.538, \alpha_2 = 0.462$)
3	U_2
4	U_2

Cette allocation reflète un compromis basé sur la qualité des canaux pour chaque utilisateur. La sous-porteuse 2, considérée importante pour les deux utilisateurs, est partagée proportionnellement à leurs gains respectifs, tandis que les autres sont attribuées au bénéficiaire le plus pertinent. Ainsi, U_2 reçoit plus de ressources car il dispose de gains plus élevés sur un plus grand nombre de sous-porteuses.

Étape 5 : Mise à jour de la puissance :

Une fois l'allocation des sous-porteuses déterminée, chaque utilisateur répartit sa puissance maximale P^{\max} sur les sous-porteuses qui lui ont été attribuées, en tenant compte

du partage éventuel.

- **Utilisateur** U_1 dispose des sous-porteuses 1 (exclusive) et 2 (partagée à 53.8%). Il répartit donc sa puissance P^{\max} sur ces deux sous-porteuses comme suit :
 - $P_{1,1} = \frac{1}{2}P^{\max}$ (sous-porteuse 1 exclusive)
 - $P_{1,2} = \frac{1}{2} \cdot 0.538 \cdot P^{\max}$
- **Utilisateur** U_2 dispose des sous-porteuses 2 (partagée à 46.2%), 3 et 4. Il répartit P^{\max} sur trois sous-porteuses :
 - $P_{2,2} = \frac{1}{3} \cdot 0.462 \cdot P^{\max}$
 - $P_{2,3} = \frac{1}{3}P^{\max}$
 - $P_{2,4} = \frac{1}{3}P^{\max}$

Cette stratégie garantit que la puissance totale allouée par chaque utilisateur respecte la contrainte de puissance maximale, tout en adaptant la répartition selon la négociation et la qualité des canaux.

Étape 6 : Calcul des taux et de l'utilité de Nash :

En supposant un modèle de type Shannon :

$$R_i = \sum_{n \in \mathcal{N}_i} \log_2(1 + \text{SNR}_{i,n}) \quad (3.15)$$

d'où

$$U(\mathbf{P}) = [(R_1 - R_1^{\min})(R_2 - R_2^{\min})]^{1/2} \quad (3.16)$$

Cela permet d'estimer l'efficacité de la solution par rapport au point de désaccord.

Illustration de l'équité dans l'allocation finale : Contrairement à une approche non coopérative où chaque utilisateur chercherait à maximiser son propre débit au détriment de l'autre, l'algorithme de négociation de Nash aboutit à une allocation plus équitable.

Chaque utilisateur reçoit :

- un ensemble de sous-porteuses où il a un avantage (meilleur canal) ;
- une part équitable des sous-porteuses communes, proportionnelle à la qualité de son canal.

Par exemple, bien que U_2 ait de meilleurs gains globaux sur plusieurs sous-porteuses, U_1 a pu conserver la sous-porteuse 1 (la plus favorable pour lui) et obtenir une part significative de la sous-porteuse 2, au lieu d'en être totalement exclu.

Cette solution satisfait les deux utilisateurs :

- **aucun utilisateur n'est lésé** : chacun dépasse son débit minimal R_i^{\min} ;
- **la solution est Pareto-optimale** : aucune amélioration possible pour un utilisateur sans pénaliser l'autre ;

- **accord équilibré** : le produit $(R_1 - R_1^{\min})(R_2 - R_2^{\min})$ est maximisé.

3.2.6 Conclusion

La première partie de notre étude s’est concentrée sur les jeux de négociation appliqués à la gestion des ressources dans les réseaux acoustiques sous-marins pour le cas de deux utilisateurs. Ces modèles offrent un cadre coopératif dans lequel les joueurs cherchent à parvenir à des accords mutuellement bénéfiques, en tenant compte de leurs intérêts respectifs. Bien que cette approche permette d’améliorer l’équité et la coordination par rapport aux jeux non coopératifs, elle repose souvent sur des hypothèses fortes concernant la connaissance des préférences et la rationalité des joueurs, ainsi que sur des capacités de communication relativement stables entre les nœuds.

Dans un environnement aussi dynamique, et complexe, il devient essentiel d’explorer des méthodes plus flexibles et adaptatives, capables d’apprendre directement des interactions avec l’environnement. C’est dans ce contexte que l’intégration de l’apprentissage par renforcement, et plus spécifiquement **QL**, constitue une évolution naturelle. Avant d’aborder en détail les scénarios multi-utilisateurs intégrant cette approche dans le cadre coopératif, nous introduisons dans la section suivante les concepts fondamentaux de l’apprentissage par renforcement, ainsi que leurs spécificités d’application dans les réseaux **ASM**.

3.3 Apprentissage par renforcement dans les réseaux ASM

L’apprentissage par renforcement est une branche de l’intelligence artificielle inspirée de la psychologie comportementale, qui s’intéresse à la manière dont un joueur autonome peut apprendre à interagir avec un environnement inconnu afin de maximiser une récompense cumulative. Contrairement à l’apprentissage supervisé, qui repose sur des données étiquetées, ou à l’apprentissage non supervisé, qui détecte des structures latentes dans les données, l’apprentissage par renforcement ou Reinforcement Learning (**RL**) en anglais repose sur un processus d’essais-erreurs. L’utilisateur apprend par expérience, en observant les conséquences de ses actions, et affine progressivement sa stratégie pour atteindre un comportement optimal [Vengerov, 2007]-[Staddon, 2020].

3.3.1 Introduction à l’apprentissage par renforcement

Le **RL** est généralement modélisé par un Processus de Décision Markovien (Markov Decision Process (**MDP**)), défini par un quintuplet $(\mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{R}, \gamma)$:

- \mathcal{S} : ensemble des états décrivant les configurations du réseau (ex. : qualité de canal, occupation des sous-porteuses) ;
- \mathcal{A} : ensemble des actions possibles (ex. : puissance transmise, choix de sous-porteuses) ;
- $\mathcal{T}(s, a, s')$: fonction de transition donnant la probabilité de passer de l’état s à s' après action a ;

- $\mathcal{R}(s, a, s')$: fonction de récompense reflétant un objectif local ou global (débit, équité) ;
- γ : facteur d'actualisation ($0 < \gamma \leq 1$) pondérant les récompenses futures.

L'objectif de l'utilisateur est de découvrir une politique optimale $\pi^*(s)$ qui maximise l'espérance de la récompense cumulée :

$$\pi^* = \arg \max_{\pi} \mathbb{E} \left[\sum_{t=0}^{\infty} \gamma^t R(s_t, a_t, s_{t+1}) \right].$$

Dans le contexte des réseaux **ASM** utilisant la technologie **OFDM**, chaque nœud (joueur) peut apprendre à allouer dynamiquement sa puissance de transmission et partager ses sous-porteuses en fonction des conditions du canal acoustique, et de l'état du réseau environnant.

L'environnement étant partiellement observable, fortement dynamique, et sujet à des interférences, les approches de **RL** se montrent particulièrement bien adaptées.

Méthodes courantes adaptées aux réseaux **ASM** :

- **QL** : Une méthode d'apprentissage par renforcement sans modèle (model-free) et hors-politique (off-policy). Elle repose sur la mise à jour d'une table de valeurs $Q(s,a)$, où chaque paire état-action est associée à un score représentant la récompense cumulative attendue. Le Q-learning sélectionne l'action qui maximise cette fonction, tout en explorant occasionnellement d'autres options via une politique ϵ -greedy. Cette méthode est particulièrement bien adaptée aux environnements discrets, où le nombre d'états et d'actions est limité. Dans les réseaux **ASM**, le **QL** est efficace lorsque l'environnement (nombre de canaux, puissances disponibles) peut être discrétisé de manière raisonnable. Il offre une bonne stabilité et une convergence théorique, par exemple [Kim, 2018] a proposé un protocole de routage distribué fondé sur le Q-learning dans le cadre des réseaux sous-marins dynamiques, montrant une double amélioration sensible en efficacité énergétique et en débit.
- **State-Action-Reward-State-Action (SARSA)** : Une variante de Q-learning dite en politique (on-policy), car elle met à jour la valeur de $Q(s,a)$ en utilisant l'action réellement choisie par la politique de l'agent, et non pas nécessairement la meilleure action possible, cela rend **SARSA** plus conservateur dans l'apprentissage, ce qui est souvent souhaitable dans des environnements dynamiques ou incertains. Dans le contexte des réseaux **ASM**, **SARSA** permet d'apprendre en suivant une politique d'exploration douce, en intégrant directement les choix de l'agent au cours de l'apprentissage, cela le rend utile lorsque le système évolue fréquemment (ex. : variations rapides du spectre, utilisateurs mobiles) et que la stabilité est plus critique que la performance maximale. Par ailleurs, cet article [Martin et al., 2018] a proposé SPGP-SARSA, une extension basée sur les processus gaussiens, et montré à travers des essais réels sur un robot sous-marin que cette approche améliore à la fois la précision des prédictions et la robustesse dans des environnements maritimes complexes.
- **Deep Q-Network (DQN)** : Une extension du **QL** qui utilise un réseau de neurones profond pour approximer la fonction $Q(s,a)$. Cela permet de traiter des

espaces d'états de grande dimension ou continus, ce qui est typiquement le cas dans les systèmes [OFDM](#), où les états peuvent inclure des informations complexes comme le [SNR](#), les coefficients de fading, ou l'occupation spectrale. Le [DQN](#) introduit plusieurs techniques pour stabiliser l'apprentissage, comme le replay buffer (mémoire d'expériences) et les cibles figées. Il est particulièrement adapté aux réseaux [ASM](#) de nouvelle génération, où la complexité du canal nécessite des représentations non linéaires. Néanmoins, son apprentissage est sensible aux hyperparamètres et nécessite un grand volume d'interactions, ce qui peut être un frein dans des systèmes à ressources limitées. Par exemple, un schéma de modulation adaptative basé sur [DQN](#) a été proposé dans [[Cui et al., 2023](#)] pour les communications [OFDM](#) sous-marines. Il permet de sélectionner dynamiquement le mode de modulation en fonction de l'état du canal, démontrant une réduction du taux d'erreur binaire et une amélioration du débit du système.

- **Actor-Critic** : Le modèle Actor-Critic combine les avantages des approches basées sur la fonction de valeur et celles basées sur les politiques. Il se compose de deux entités : l'acteur, qui sélectionne les actions selon une politique paramétrée, et le critique, qui estime la valeur des états ou des paires état-action pour guider l'apprentissage de l'acteur. Cette méthode est particulièrement utile dans les environnements à actions continues, comme la sélection dynamique du niveau de puissance ou de la bande passante dans un réseau [ASM](#). Elle permet une optimisation plus fine que les méthodes basées sur une discrétisation. Cependant, les modèles Actor-Critic peuvent souffrir d'instabilités et de variance élevée, notamment dans des environnements non stationnaires ou fortement bruités, ce qui nécessite souvent un ajustement fin des paramètres et une surveillance attentive du processus d'apprentissage. Des travaux récents ont démontré l'efficacité des architectures Actor-Critic dans les réseaux sous-marins spécifiquement [[Rahman et al., 2025](#)] ont proposé un protocole MAC économe en énergie pour les réseaux de capteurs sous-marins, combinant CSMA/CA, TDMA et Actor-Critic, améliorant significativement l'efficacité énergétique et la fiabilité dans des environnements acoustiques complexes.

Le tableau [3.3](#) compare les approches d'apprentissage par renforcement pertinentes pour les réseaux [ASM](#). Notre choix s'est porté sur le [QL](#) après une analyse par élimination basée sur la nature de notre problème. Premièrement, notre problème d'allocation de ressources [OFDM](#) (sous-porteuses, niveaux de puissance) se définit par un espace d'actions discret. Cette caractéristique fondamentale rend les méthodes conçues pour les actions continues, telles que l'Actor-Critic, inutilement complexes et mal adaptées. Deuxièmement, en comparant le [QL](#) aux autres méthodes pour actions discrètes comme le [DQN](#), un arbitrage crucial apparaît entre complexité et performance. Bien que le [DQN](#) puisse gérer des espaces d'états vastes grâce aux réseaux de neurones, son coût computationnel élevé et son besoin en données sont des contraintes majeures pour des nœuds [ASM](#) aux ressources limitées. Le [QL](#) (similaire à [SARSA](#) sur ce point) offre un compromis bien supérieur pour notre contexte : il est sans modèle (model-free), simple à implémenter, et robuste, tout en étant parfaitement optimisé pour les actions discrètes. Nous reconnaissons la limite du [QL](#) classique en environnement multi-agent, c'est précisément pour surmonter cet obstacle que nous l'avons intégré dans le cadre de jeu coopératif de négociation multi-utilisateur détaillé à la section suivante.

TABLE 3.3 – Comparaison des méthodes d’apprentissage par renforcement pour les réseaux [ASM](#)

Méthode	Type de politique	Type d’actions	Avantages	Limites / Contraintes
QL [Kim, 2018]	Off-policy	Discrètes	Simple à implémenter, convergence garantie en environnement discret stationnaire	Inefficace dans les environnements à grand espace d’états/actions
SARSA [Martin et al., 2018]	On-policy	Discrètes	Meilleure stabilité en environnement stochastique, apprend en suivant la politique réelle	Moins performant que QL si la politique est sous-optimale
DQN [Cui et al., 2023]	Off-policy	Discrètes ou continues (états)	Gère des états complexes, adapté aux environnements OFDM , capable de généralisation	Besoin élevé en données et en ressources de calcul, apprentissage instable sans précautions
Actor-Critic [Rahman et al., 2025]	On-policy ou off-policy	Continues	Supporte directement les actions continues, meilleure adaptation dans les réseaux dynamiques	Instabilité potentielle, sensibilité à la variance

Cas d’usage pour un environnement sous-marin : Dans le contexte spécifique des réseaux [ASM](#) exploitant la technologie [OFDM](#), le [RL](#) offre plusieurs avantages déterminants pour l’allocation conjointe de la puissance et des sous-porteuses. Premièrement, il permet une adaptabilité aux conditions extrêmes du canal acoustique, caractérisé par une forte variabilité, une latence élevée, et la présence de bruit impulsif. Deuxièmement, la nature décentralisée du [RL](#) rend possible une prise de décision autonome par chaque nœud, réduisant ainsi la nécessité d’une coordination centralisée coûteuse en bande passante. Troisièmement, le [RL](#) favorise un apprentissage continu, permettant aux politiques de s’ajuster dynamiquement aux changements de topologie ou à la variation de charge du réseau. Enfin, il est capable de gérer des objectifs multiples de manière naturelle, en intégrant simultanément des critères tels que la maximisation de débit, la fiabilité et l’équité entre les nœuds.

Face à la complexité inhérente des environnements sous-marins et aux défis posés par la gestion fine des ressources dans des réseaux OFDM, l'apprentissage par renforcement donc se révèle être une approche particulièrement prometteuse, complémentaire voire supérieure aux méthodes classiques d'optimisation. L'intégration de la théorie des jeux et de l'apprentissage par renforcement, en particulier du QL, permet de concevoir des stratégies d'allocation de ressources décentralisées plus robustes et intelligentes dans les réseaux ASM. Dans un environnement multi-agent, chaque nœud peut être vu comme un joueur rationnel (au sens de la théorie des jeux), autonome, interagissant avec d'autres, les agents n'ont pas une connaissance a priori des fonctions d'utilité des autres ni de la structure exacte du jeu. Le Q-learning multi-agent (ou MARL) permet alors à chaque agent d'apprendre, par interactions répétées avec l'environnement et les autres agents, une stratégie optimale qui converge vers un équilibre.

Ceci présente plusieurs avantages, d'une part, il hérite de la capacité d'adaptation dynamique du RL dans un environnement partiellement observable et bruité, typique des réseaux acoustiques, d'autre part, la structure de jeu fournit un cadre théorique pour modéliser les interactions stratégiques entre agents, notamment dans des scénarios de compétition pour des ressources rares (puissance, sous-porteuses). Cette approche conjointe permet une allocation conjointe puissance-sous-porteuses efficace et distribuée, en tenant compte à la fois de la dynamique du canal acoustique et des comportements adaptatifs des autres joueurs. L'objectif est non seulement d'atteindre un équilibre stable, mais aussi de maximiser des métriques globales comme l'efficacité, le débit total du réseau et l'équité entre tous les utilisateurs.

3.3.2 Théorie de la négociation de Nash avec Q-Learning pour un scénario multi-utilisateurs

Après avoir étudié en détail le scénario simplifié à deux utilisateurs, qui permettait d'illustrer les principes fondamentaux de la négociation coopérative via la solution de Nash, nous étendons désormais notre analyse à un environnement plus complexe et réaliste, impliquant plusieurs utilisateurs partageant les mêmes ressources spectrales dans un canal acoustique sous-marin. Dans un contexte multi-utilisateurs, les défis liés à l'allocation efficace des ressources sont renforcés. Les interférences entre utilisateurs deviennent plus significatives, rendant la coopération essentielle pour garantir une utilisation optimale du spectre et une amélioration des performances globales du système.

Le recours à un jeu coopératif de négociation s'avère particulièrement pertinent pour gérer ces interactions concurrentes tout en tenant compte des contraintes propres au canal acoustique (longs délais de propagation, faible bande passante, etc.), d'où les contributions principales de cette section peuvent être résumées comme suit :

- Nous développons une stratégie itérative d'allocation coopérative des ressources, basée sur la solution de NB combinée à RL, afin d'optimiser conjointement l'allocation de puissance et le partage des sous-porteuses.
- Nous analysons l'évolution de la fonction d'utilité de chaque utilisateur dans un cadre coopératif, en la comparant à une approche non coopérative, dans le but de mettre en évidence les gains apportés par la coopération, notamment dans des environnements fortement sujets aux interférences.

- Nous visons à minimiser les interférences générées par les utilisateurs partageant les mêmes sous-porteuses, soulignant ainsi l'importance de mécanismes d'allocation robustes pour améliorer l'efficacité du système.

Cette partie présente donc la généralisation de notre modèle aux réseaux multi-utilisateurs, en intégrant des outils d'optimisation coopérative et d'apprentissage distribué, adaptés aux contraintes spécifiques des réseaux acoustiques sous-marins. Dans ce scénario, l'objectif est de formuler le problème d'allocation de ressources pour plusieurs utilisateurs (agents) en optimisant conjointement l'allocation de puissance et le partage des sous-porteuses. Nous introduisons également une contrainte supplémentaire, à savoir un taux minimum que chaque utilisateur doit atteindre.

Pour un système à K utilisateurs, le problème est formulé comme suit, visant à maximiser l'utilité globale du système :

$$\max_{\mathbf{P}} U(\mathbf{P}) = \max_{\mathbf{P}} \prod_{i=1}^K [R_i(\mathbf{P}) - R_i^{\min}(\mathbf{P})]^{1/K}. \quad (3.17)$$

sous la contrainte :

$$\begin{cases} \sum_{n=1}^{N_{sc}} p_{i,n} \leq P^{\max}, & \forall i \in \{1, \dots, K\} & (3.17.a). \\ R_i(\mathbf{P}) \geq R_i^{\min}(\mathbf{P}), & \forall i \in \{1, \dots, K\} & (3.17.b). \end{cases}$$

La contrainte (3.17.a) représente la contrainte de puissance et (3.17.b) présente la contrainte du taux minimum défini par (3.18). En définissant le taux minimum comme étant le taux non coopératif, les utilisateurs peuvent évaluer les avantages de la coopération. Il est utilisé comme référence pour évaluer les bénéfices et compromis possibles entre les techniques coopératives et non coopératives.

$$R_i^{\min}(\mathbf{P}) = R_i^{NC}(\mathbf{P}). \quad (3.18)$$

Dans la section suivante, nous décrivons la solution proposée ainsi que l'algorithme permettant d'optimiser la distribution des ressources entre les utilisateurs.

Modélisation de l'allocation des ressources par Q-learning : Le Q-learning est un algorithme de RL permettant d'apprendre une politique optimale de sélection d'actions dans un processus de décision de Markov fini. Un processus de décision de Markov (ou MDP) est une extension du processus de Markov classique, qui modélise la prise de décision séquentielle. Il est défini par un ensemble d'états, un ensemble d'actions, une fonction de transition, et une fonction de récompense, il suppose que la probabilité d'atteindre un état futur dépend uniquement de l'état actuel et de l'action choisie, une propriété appelée propriété de Markov.

Dans notre contexte, cette méthode se révèle prometteuse pour résoudre le problème d'allocation optimale conjointe de la puissance et des sous-porteuses dans les systèmes acoustiques sous-marins. Le problème alors est formulé comme un processus d'apprentissage où un agent interagit avec l'environnement, sélectionne des actions, et met à jour une table de valeurs $Q(s, a)$ en fonction des récompenses immédiates et des gains futurs estimés. Le processus se déroule de manière itérative afin de converger vers une politique

d'allocation optimale des ressources.

Pour équilibrer l'exploration et l'exploitation des actions, nous adoptons la stratégie *epsilon-greedy*, où l'agent choisit aléatoirement une action avec une probabilité ε (exploration), ou suit l'action la plus prometteuse selon la fonction Q (exploitation). Cela permet d'explorer de nouvelles allocations tout en affinant progressivement les choix les plus efficaces et la mise à jour de la valeur Q est représentée par la formule suivante :

$$Q(s^t, a^t) \leftarrow Q(s^t, a^t) + \alpha \left(re + \gamma \max_{a^{t+1}} Q(s^{t+1}, a^{t+1}) - Q(s^t, a^t) \right), \quad (3.19)$$

avec :

- Agent : un nœud sous-marin intelligent (capteur, AUV).
- $Q(s^t, a^t)$: valeur estimée pour l'état s^t et l'action a^t à l'instant t , représentant ici une configuration d'allocation de puissance et de sous-porteuses.
- État (s) : peut inclure le gain de canal, l'indice de sous-porteuse disponible.
- Action (a) : ajustement de la puissance transmise et sélection d'une ou plusieurs sous-porteuses [OFDM](#).
- α : taux d'apprentissage, contrôlant la vitesse de mise à jour.
- re : récompense immédiate reçue après l'action a^t dans l'état s^t .
- $\gamma \in [0, 1]$: facteur d'actualisation, pondérant l'importance des récompenses futures par rapport aux récompenses immédiates.
- $\max_{a^{t+1}} Q(s^{t+1}, a^{t+1})$: estimation maximale de la valeur Q dans l'état suivant s^{t+1} , indiquant la meilleure récompense attendue.

3.3.3 Algorithme proposé de Nash Bargaining Multi-utilisateurs avec Q-Learning

L'algorithme proposé qui traite le problème d'allocation de puissance dans un réseaux acoustiques sous-marins multi-utilisateurs, où plusieurs nœuds transmetteurs (agents) se disputent des ressources spectrales limitées est donné en annexe (Algorithme 2). Cette compétition est formulée sous forme de processus décisionnel multi-agents de Markov (Multi-Agent Markov Decision Process ([MAMDP](#))), où chaque agent optimise indépendamment son allocation de puissance tout en tenant compte de l'interférence générée par les autres agents. L'espace d'états est défini par l'information sur le [CSI](#) de chaque agent et une estimation de l'interférence causée par les autres agents. L'espace d'actions consiste en des niveaux discrets d'allocation de puissance sur les sous-porteuses, représentant les différentes stratégies de puissance de transmission disponibles pour chaque agent. La structure des récompenses est conçue pour garantir que chaque agent maximise son utilité tout en maintenant l'équité dans le système. Chaque agent cherche à sélectionner des valeurs d'allocation de puissance qui augmentent son utilité tout en tenant compte de l'interférence causée par les agents voisins. Contrairement aux approches centralisées qui supposent une connaissance globale du système, notre approche distribuée réduit le besoin de telles informations globales. Chaque agent doit connaître son propre [CSI](#) ainsi qu'une estimation de l'interférence causée par les agents voisins, ce qui réduit considérablement la charge de communication et de traitement. Cette connaissance locale rend l'algorithme plus réalisable pour les déploiements réels, où les informations globales peuvent ne pas toujours être disponibles en raison de contraintes environnementales ou techniques.

Pour commencer le processus de négociation, chaque agent propose des sous-porteuses en fonction du taux réalisable pour chaque sous-porteuse, en veillant à ce que les taux proposés dépassent un point de désaccord estimé dynamiquement, ce point de désaccord est dérivé du cadre théorique du jeu non coopératif, qui présente une référence d'utilité que les joueurs doivent dépasser pour entrer dans une négociation équitable. Cela garantit que chaque solution proposée par un joueur est compétitive et empêche les résultats d'utilité faible. Lorsque plusieurs utilisateurs (joueurs) sélectionnent la même sous-porteuse, des conflits apparaissent. Pour résoudre ces conflits, l'algorithme utilise le Nash Q-learning, une méthode adaptative permettant aux joueurs d'apprendre les comportements optimaux au fil du temps. Plus précisément, l'algorithme adopte une stratégie de sélection d'actions dite **Softmax**, qui équilibre exploration et exploitation. Contrairement à une politique purement gourmande qui choisit systématiquement l'action ayant la plus haute valeur Q, la stratégie Softmax attribue à chaque action une probabilité de sélection proportionnelle à une fonction exponentielle de sa valeur Q. Formellement, la probabilité de choisir une action a dans l'état s est donnée par :

$$P(a | s) = \frac{e^{Q(s,a)/\tau}}{\sum_b e^{Q(s,b)/\tau}}, \quad (3.20)$$

où τ est un paramètre de température qui contrôle le degré d'exploration : des valeurs élevées de τ favorisent une sélection plus uniforme des actions (exploration), tandis que des valeurs faibles conduisent à privilégier les actions avec les valeurs Q les plus élevées (exploitation). Cette approche probabiliste permet aux agents d'explorer une gamme variée d'allocations de puissance, évitant ainsi d'avoir des solutions sous-optimales.

Dans la section suivante, nous présentons et analysons les résultats de simulation issus de deux scénarios distincts : l'algorithme de **NB** appliqué à un cas à deux utilisateurs sans **RL**, et d'autre part, une extension multi-utilisateurs intégrant le **QL** pour une allocation dynamique des ressources.

3.4 Résultats de simulation et analyse

3.4.1 Cas à 2 utilisateurs

Les résultats de la simulation sont fournis pour étudier l'évolution de la fonction d'utilité des utilisateurs en fonction des différents niveaux d'interférence, des valeurs distinctes des sous-porteuses, et en prenant le cas où R_i^{\min} est fixé à zéro. Le logiciel MATLAB est utilisé pour simuler l'algorithme proposé et un modèle de canal **ASM** est généré après 20 itérations, ce qui correspond à deux paires de transmetteurs et récepteurs extraites aléatoirement dans un rayon de 2000 m, avec une profondeur d'eau de 100 m, où la même bande passante B est partagée entre les deux.

Les paramètres du système sont représentés par le tableau suivant :

Étude de la fonction d'utilité individuelle :

Dans notre étude, nous avons abordé un défi crucial dans les réseaux **ASM** : l'optimisation de l'allocation des ressources basée sur la théorie des jeux de négociation coopératifs. Pour ce faire, nous avons comparé l'efficacité de notre algorithme proposé avec

TABLE 3.4 – Paramètres.

Désignation	Paramètres	Valeurs
Nombre d'utilisateurs	K	2
Nombre de sous-porteuses	N_{sc}	[8, 16, 32, 50, 64]
Bande passante	B	14 kHz
Offres autorisées	N_{offer}	3
Puissance maximale	P^{\max}	1 Watt
Rapport signal sur bruit	SNR	15 dB
Itérations	L	20
Profondeur d'eau	WD	100 m
Profil de puissance du canal direct (exemple)	$\sum_{n=1}^{N_{sc}} \mathbb{E} h_{ii}(n, t) ^2$	[0.1343, 0.7807]
Profil de puissance du canal d'interférence (exemple)	$\sum_{n=1}^{N_{sc}} \mathbb{E} h_{ji}(n, t) ^2$	[0.0689, 0.2067]
Interférence forte	$X_h h_{ji}(n, t) ^2$	$X_h = 0.5$
Interférence faible	$X_l h_{ji}(n, t) ^2$	$X_l = 0.01$

deux références majeures : d'abord, une approche non coopérative étudiée précédemment, et ensuite, un algorithme optimal. Cette comparaison nous a permis de démontrer les avantages potentiels de notre approche proposée en termes d'efficacité de l'allocation des ressources.

La Fig. 3.1 démontre que l'approche coopérative surpasse largement l'approche non-coopérative en termes de fonction d'utilité individuelle, confirmant les bénéfices d'une coordination. Pour valider l'efficacité de notre algorithme de négociation, nous l'avons comparé à la solution optimale théorique (points noir et vert), qui n'a pu être calculée qu'à $N_{sc} = 8$ sous-porteuses en raison de la complexité de calcul exponentielle de l'optimum. La forte proximité entre les résultats de notre algorithme coopératif et ces points optimaux pour $N_{sc} = 8$ confirme qu'il est un heuristique performant capable d'approcher l'optimum. Par ailleurs, le croisement des courbes d'utilité entre les Utilisateurs 1 et 2, visible après $N_{sc} = 15$, est la manifestation directe du mécanisme d'équité intégré. Ce mécanisme garantit que l'utilisateur initialement désavantagé par les conditions de canal reçoit une allocation préférentielle sur les nouvelles sous-porteuses, entraînant un rééquilibrage dynamique des utilités à mesure que les ressources augmentent. L'algorithme atteint ainsi un double objectif : maximiser l'utilité globale tout en assurant une distribution des ressources juste et efficace.

Dans une stratégie non coopérative, chaque utilisateur est motivé par une action égoïste, cherchant à maximiser sa propre utilité indépendamment des intérêts collectifs. Cela peut entraîner une forte concurrence entre les utilisateurs pour l'accès aux ressources, souvent au détriment de la performance globale du système. En revanche, la coopération permet aux utilisateurs d'échanger des informations pertinentes, telles que leurs besoins, préférences et contraintes environnementales. Cette collaboration favorise une prise de décision plus éclairée, où les propositions d'allocation des ressources tiennent compte des besoins et demandes de chaque utilisateur. Dans notre scénario, cette collaboration abou-

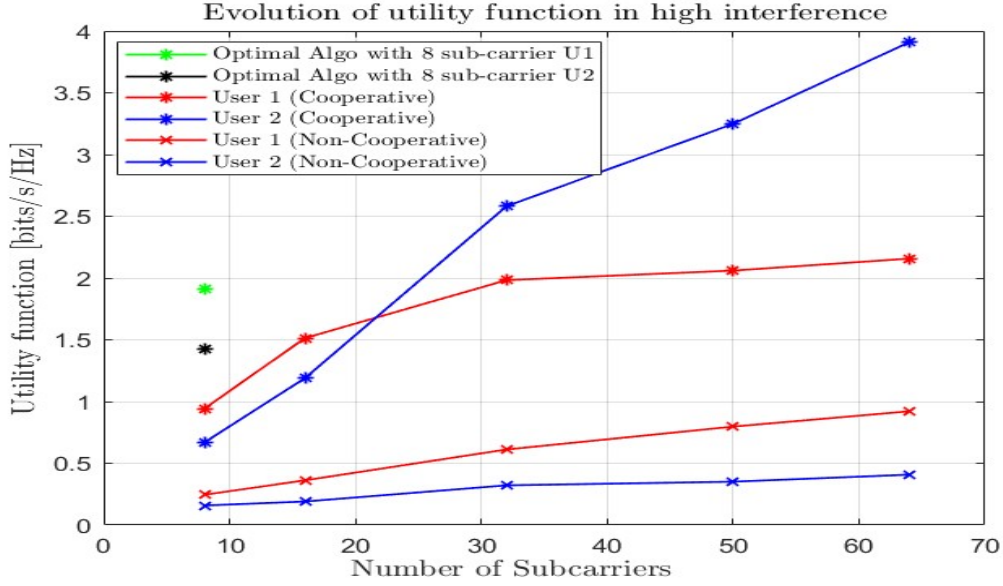


FIGURE 3.1 – Évolution de la fonction d'utilité des utilisateurs (Coopératif vs Non coopératif).

tit à une allocation des ressources plus efficace et équitable, facilitée par l'échange continu de propositions entre les utilisateurs. Ainsi, plutôt que de se concentrer uniquement sur la maximisation de leur propre utilité, les utilisateurs visent à optimiser la fonction d'utilité globale du système, ce qui peut mener à des opportunités plus efficaces pour explorer différentes allocations.

Étude d'impact des interférences :

Cependant, il est important de noter que les niveaux d'interférence peuvent toujours influencer l'évolution de la fonction d'utilité en raison notamment de la distance variable entre les émetteurs et les récepteurs, comme l'illustre la Fig. 3.2.

La performance individuelle peut varier en fonction du niveau d'interférence et du degré de coopération entre les utilisateurs, mais elle reste toujours meilleure que dans le jeu non coopératif. Ainsi, dans notre étude, nous trouvons que la fonction d'utilité de l'utilisateur 1 évolue en fonction du niveau d'interférence, montrant de meilleures performances en cas de faible interférence et des performances moins optimales en cas de forte interférence. Étant donné que l'utilisateur 2 bénéficie d'un gain plus important dans ce scénario, il se voit attribuer un poids plus élevé sur certaines sous-porteuses, représenté par α_i , par rapport à l'utilisateur 1.

Cette allocation de poids permet à l'utilisateur 2 d'allouer les ressources disponibles de manière plus efficace, lui permettant de maintenir des performances stables même en présence d'une interférence significative. On peut conclure alors que cette allocation des ressources garantit que chaque utilisateur peut maintenir un niveau de performance satisfaisant en fonction de ses besoins, contribuant ainsi à la stabilité et à la fiabilité globales du système. En optimisant l'allocation des ressources en fonction des conditions changeantes des canaux et des besoins individuels des utilisateurs, notre approche favorise

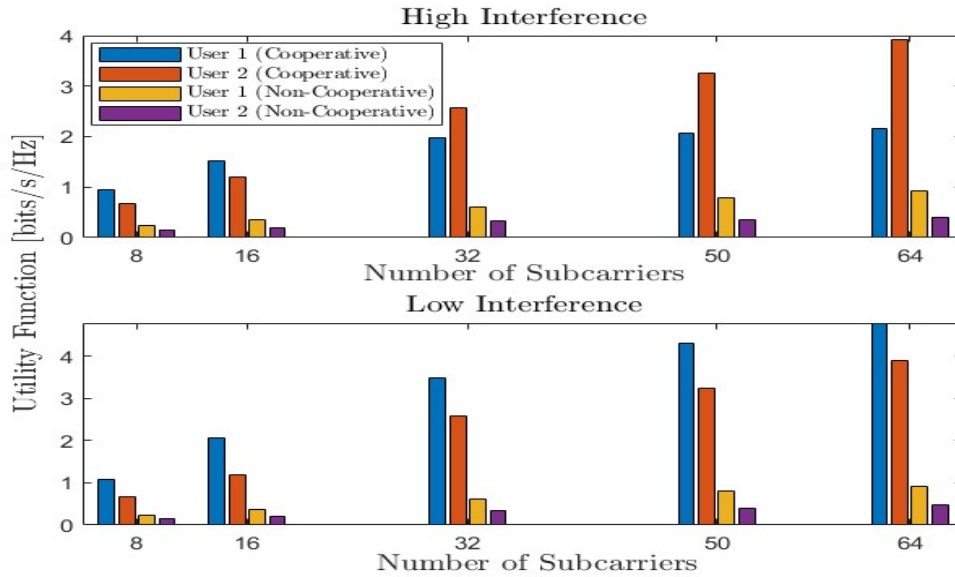


FIGURE 3.2 – Évolution de l'utilité en fonction du niveau d'interférence (Coopératif vs Non coopératif).

une utilisation efficace du spectre et une meilleure répartition de la puissance pour les deux utilisateurs par rapport aux jeux non-coopératifs.

Étude de la fonction d'utilité globale :

L'efficacité de la communication dépend fortement de la disponibilité et de la gestion des ressources, notamment le nombre de sous-porteuses allouées aux utilisateurs. En augmentant ce nombre, le système peut offrir une plus grande flexibilité pour optimiser l'allocation des ressources et ainsi améliorer ses performances. Cette tendance est clairement illustrée par la Fig. 3.3, qui montre les gains observés lorsque le nombre de sous-porteuses augmente.

En pratique, cette idée peut orienter la conception des protocoles de communication [ASM](#) en mettant l'accent sur l'importance d'allouer un nombre adéquat de sous-porteuses pour répondre aux exigences de communication variables et aux conditions environnementales. En garantissant un nombre suffisant de sous-porteuses, les systèmes de communication peuvent mieux s'adapter aux conditions changeantes des canaux et aux demandes des utilisateurs, améliorant ainsi la fiabilité de la communication dans des contextes réels. De plus, la convergence de la fonction d'utilité globale avec un nombre croissant de sous-porteuses suggère qu'il existe un équilibre optimal dans l'allocation de ressources, au-delà duquel des ressources supplémentaires peuvent ne pas apporter d'améliorations significatives des performances. Cette constatation souligne l'importance de l'optimisation des ressources et met en évidence la nécessité d'algorithmes efficaces capables d'ajuster dynamiquement l'allocation de ressources en fonction des exigences changeantes du système.

Dans des contextes pratiques, cette idée peut guider le développement d'algorithmes adaptatifs qui ajustent dynamiquement l'allocation de ressources en fonction des conditions réelles du système. En surveillant en continu les performances du système et en ajustant les stratégies d'allocation de ressources en conséquence, ces algorithmes peuvent

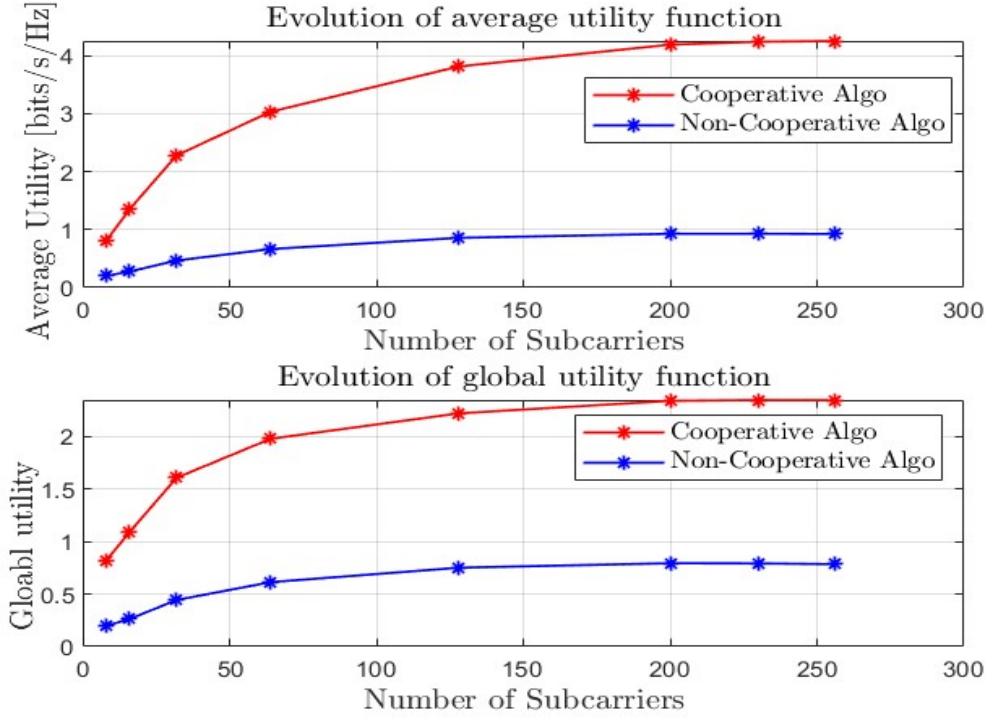


FIGURE 3.3 – Évolution de la fonction d'utilité moyenne et globale (Coopératif vs Non coopératif).

optimiser l'utilisation des ressources et améliorer l'efficacité globale du système. En outre, il est toujours observé que la comparaison entre les stratégies d'allocation des ressources coopérative et non coopérative met en évidence les avantages de la coopération des utilisateurs pour améliorer les performances du système et garantir l'équité dans l'allocation des ressources. Pratiquement, cette idée peut orienter la conception de protocoles de communication coopératifs qui encouragent le partage d'informations et la prise de décision collaborative parmi les utilisateurs afin d'améliorer les performances du système. En favorisant la coopération des utilisateurs, ces protocoles peuvent améliorer la fiabilité de la communication, améliorer les performances du système et garantir une allocation équitable des ressources dans les systèmes [ASM](#).

L'équité entre les utilisateurs est évaluée à l'aide de l'indice proposé par Jain *et al.* [[Jain et al., 1984](#)], défini par :

$$J(x_1, x_2, \dots, x_n) = \frac{(\sum_{i=1}^n x_i)^2}{n \sum_{i=1}^n x_i^2} \quad (3.21)$$

où x_i représente la ressource attribuée à l'utilisateur i . La valeur de J varie entre $1/n$ (équité minimale) et 1 (équité parfaite).

D'autre part, lorsque nous avons calculé cet indice d'équité, comme spécifié dans [[Jain et al., 1984](#)], avec un nombre plus élevé de sous-porteuses, par exemple 256, nous avons constaté que l'approche coopérative proposée aboutissait à un indice d'équité de 0.989, tandis que le jeu non coopératif donnait un indice de 0.855. Cette comparaison des valeurs des indices d'équité souligne l'efficacité accrue de notre algorithme coopératif

pour garantir l'équité entre les utilisateurs par rapport à la méthode non coopérative. Cette différence significative démontre que notre approche coopérative est plus efficace pour garantir une distribution équitable des ressources entre les utilisateurs. Ces résultats sont la contribution de l'article [Gharsalli et al., 2024b].

Discussion :

Cette étude examine l'existence d'une solution de NB pour un jeu de coopération concernant l'allocation de puissance dans les réseaux de communication acoustique sous-marine. Un algorithme de NB a été proposé afin d'allouer de manière optimale la puissance entre les utilisateurs et de partager efficacement les sous-porteuses disponibles, basé sur un schéma de négociation. Différents niveaux d'interférence ont été adoptés lors de la simulation afin de montrer le comportement de l'approche proposée.

Les résultats de la simulation révèlent des améliorations significatives dans la fonction d'utilité de chaque utilisateur grâce à la coopération et à la mise en œuvre de la solution de NB, ce qui conduit à une amélioration globale des performances du système. En effet, la coopération permet une allocation plus équilibrée et efficace des ressources, et les utilisateurs, en travaillant ensemble, parviennent à optimiser l'utilisation du spectre et de la puissance disponible. Ces résultats soulignent le rôle crucial de la coopération entre utilisateurs pour l'optimisation de l'efficacité des réseaux de communication sans fil acoustique, en particulier lorsqu'on les compare à des scénarios non coopératifs où chaque utilisateur agit de manière égoïste.

Cependant, il est nécessaire de faire l'extension de l'approche au cas multi-utilisateurs qui représente une étude intéressante, car elle permettrait d'explorer les dynamiques de coopération à plus grande échelle, avec un nombre d'utilisateurs plus important et des interactions plus complexes, ce qui pourrait enrichir davantage l'efficacité du système global et ce qui est va illustrer dans la partie suivante.

3.4.2 Cas multi-utilisateurs

Cette section présente la simulation d'un système multi-utilisateurs, composé spécifiquement de 8 puis 12 utilisateurs, partageant la même bande passante $B = 14$ kHz, avec une puissance maximale $P^{\max} = 2$ W et une puissance minimale $P^{\min} = 0.2$ W. MATLAB est utilisée pour simuler l'algorithme proposé ainsi que la solution de négociation de Nash, la simulation s'appuie sur le modèle de canal acoustique sous-marin spécifique décrit dans [Qarabaqi and Stojanovic, 2013], qui considère plusieurs paires émetteur-récepteur réparties aléatoirement dans un rayon de 2000 m et une profondeur d'eau de 80 m.

Nous effectuons des simulations sous des conditions de forte interférence afin d'évaluer l'efficacité de l'algorithme et nous évaluons l'utilité individuelle de chaque utilisateur à la convergence, dans le cadre de jeux coopératifs et non coopératifs basés sur le Q-learning, comme illustré dans la Fig. 3.4. La performance de l'utilité globale est représentée dans la Fig. 3.5, tandis que la Fig. 3.6 présente une matrice illustrant l'allocation des sous-porteuses disponibles. Il est important de noter ici que l'algorithme suppose que les utilisateurs peuvent échanger les informations nécessaires à l'apprentissage de manière fiable. Bien que ces échanges puissent, en pratique, introduire un surcoût temporel ou

communicationnel, ils ne sont pas modélisés dans notre simulation et les résultats présentés ici sont issus d'une simulation idéalement coordonnée, ce qui permet de se concentrer sur la dynamique d'apprentissage.

Comparaison de l'utilité individuelle dans le processus de négociation

L'utilité individuelle de chaque utilisateur reflète son intention dans le processus de négociation. L'objectif est de maximiser sa propre performance tout en tenant compte des préférences et de l'utilité des autres joueurs. Dans un jeu non coopératif, chaque utilisateur optimise indépendamment son utilité, ce qui conduit à des disparités importantes dans les niveaux de satisfaction individuels, cela se produit parce que les utilisateurs se concentrent uniquement sur la maximisation de leurs propres bénéfices sans tenir compte de l'impact de leurs décisions sur les autres. Un tel comportement entraîne souvent une allocation des ressources déséquilibrée, où certains utilisateurs reçoivent des allocations plus favorables que d'autres.

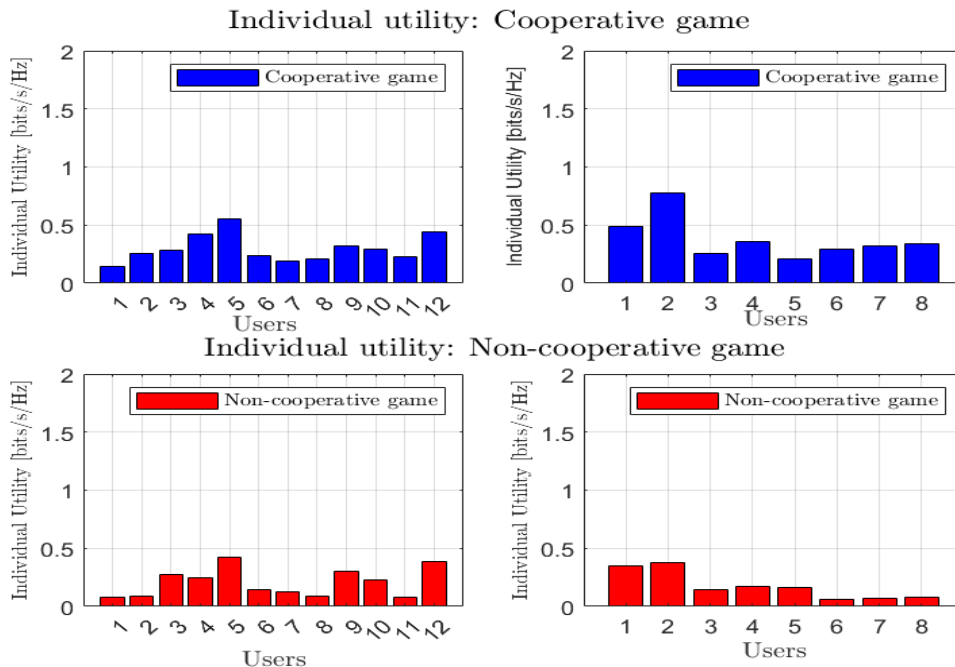


FIGURE 3.4 – Comparaison de la fonction d'utilité individuelle (Jeu coopératif vs Jeu non-coopératif).

En revanche, l'objectif du jeu de négociation dans notre contexte est de trouver une solution qui maximise le produit des utilités obtenues par chaque joueur, cette solution cherche à équilibrer la distribution de la puissance entre les utilisateurs et à allouer les sous-porteuses de manière à bénéficier à tous les joueurs. Ainsi, l'algorithme proposé facilite ce processus coopératif grâce au Q-learning, qui permet une redistribution dynamique et efficace de la puissance en ajustant les allocations aux utilisateurs là où elles sont les plus nécessaires. Cette approche améliore l'efficacité globale du système et garantit une distribution plus équilibrée des ressources par rapport au jeu non coopératif, conduisant à une utilité individuelle plus équitable, comme le montre la Fig. 3.4. En cherchant à optimiser leur propre utilité tout en prenant en compte celle des autres, les utilisateurs contribuent à réduire les disparités dans les niveaux de satisfaction. Cette approche coopérative améliore

l'efficacité globale du système et mène à une distribution des ressources plus équitable, comme l'indiquent des indices d'équité tels que l'indice d'équité de Jain [Jain et al., 1984].

Par exemple, dans le jeu coopératif avec 8 utilisateurs, l'indice d'équité est de 0,91, ce qui indique une distribution relativement équitable des ressources, cependant, pour 12 utilisateurs, l'indice d'équité chute à 0,78, suggérant une distribution moins équitable en raison de la concurrence accrue pour les ressources. De plus, l'amélioration moyenne de l'utilité à travers les utilisateurs dans le cadre coopératif met en évidence les avantages de la coopération. Pour le cas de 8 utilisateurs, l'amélioration moyenne par utilisateur est d'environ 187,6%, tandis que pour 12 utilisateurs, l'amélioration moyenne est d'environ 80,29%. Ces résultats démontrent l'efficacité de la stratégie coopérative pour améliorer la distribution de l'utilité, car le jeu coopératif conduit non seulement à une allocation plus équilibrée des ressources, mais améliore également de manière significative l'utilité individuelle par rapport au scénario non coopératif.

Cependant, à mesure que le nombre d'utilisateurs augmente, un déclin de l'utilité individuelle devient inévitable, cela est dû à la concurrence accrue pour des ressources limitées, à l'augmentation des interférences et aux contraintes de coopération.

Évolution de l'utilité globale

La fonction d'utilité globale est une mesure de la performance de notre système de communication, obtenue par itérations de l'algorithme coopératif présentée par la figure 3.5.

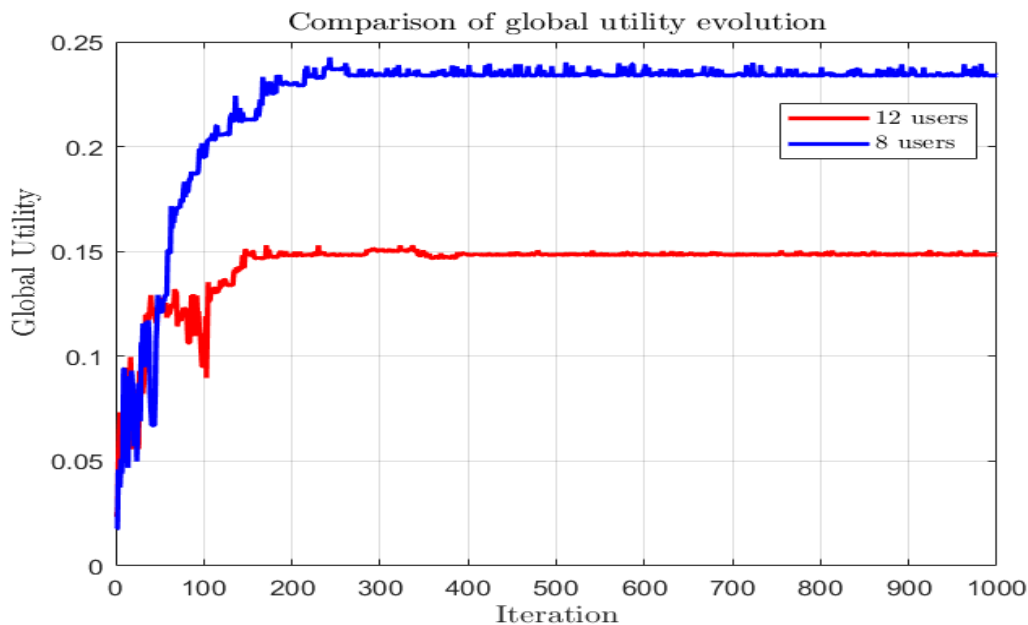


FIGURE 3.5 – Convergence de la fonction d'utilité globale.

L'amélioration des performances observée est la moyenne des nombreuses dynamiques environnementales et caractéristiques aléatoires vécues lors du processus d'apprentissage. À mesure que les itérations progressent, l'algorithme proposé améliore progressivement l'efficacité du système en apprenant et en optimisant l'allocation des ressources. Initialement, la fonction d'utilité globale commence à une valeur faible, reflétant une distribution

sous-optimale des ressources. Au fil du temps, elle augmente progressivement, indiquant le processus d'apprentissage du système pour affiner sa stratégie d'allocation, tant dans le scénario à 8 utilisateurs que celui à 12 utilisateurs. Cette croissance de l'utilité globale est principalement déterminée par la phase d'exploration de l'algorithme Q-learning, pendant laquelle diverses possibilités d'allocation des ressources sont évaluées afin d'identifier celles qui maximisent l'utilité. Les ajustements itératifs et les étapes d'apprentissage garantissent que l'algorithme proposé converge vers des solutions plus efficaces.

Les fluctuations observées de l'utilité globale entre les itérations 100 et 400 mettent en évidence les dynamiques de l'algorithme pendant la phase d'exploration. Ces variations reflètent les ajustements effectués par les agents sur leur stratégie de puissance et la résolution des conflits. En tenant compte du temps, il est constaté que le processus de convergence est plus rapide pour 8 utilisateurs par rapport à 12 utilisateurs, ce qui pourrait être lié à la complexité accrue dans la gestion d'un plus grand nombre d'utilisateurs et de sous-porteuses, ce qui pourrait nécessiter plus d'itérations pour que le système atteigne l'équilibre. De plus, la convergence de l'algorithme peut prendre du temps en raison de la nature inhérente du processus Q-learning, qui comporte à la fois des phases d'exploration et d'exploitation. Pendant l'exploration, les agents testent diverses stratégies de puissance, ce qui peut prolonger le temps de convergence au-delà du temps de cohérence du canal. Le temps nécessaire pour achever un processus de négociation, bien qu'affecté par les délais de propagation et les conditions dynamiques du canal, résulte de l'adaptation continue du système pour trouver une solution optimale.

Matrice d'allocation

Les deux matrices suivantes montrent comment l'algorithme proposé attribue les sous-porteuses aux joueurs. Les valeurs binaires (0 et 1) dans la matrice indiquent si une sous-porteuse est allouée (1) ou non allouée (0).

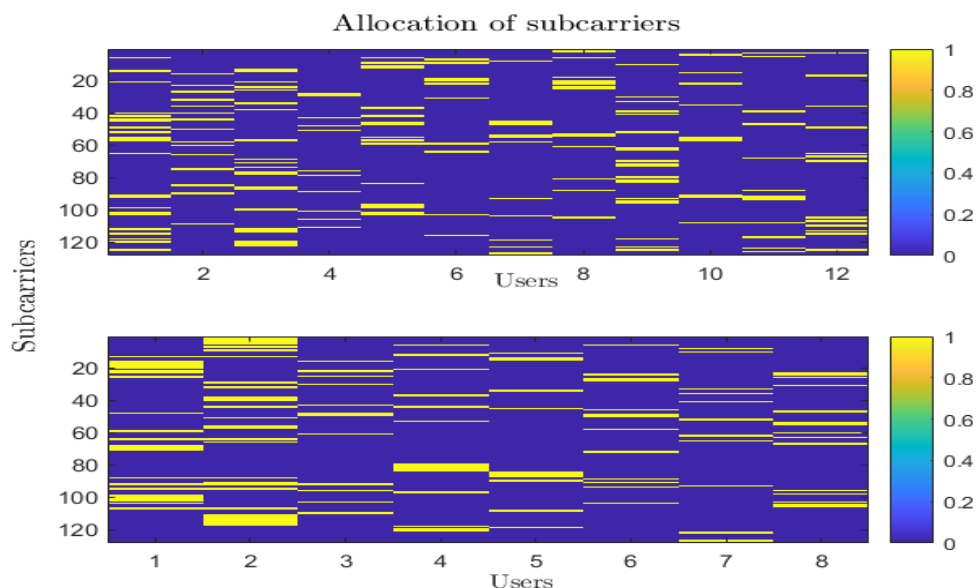


FIGURE 3.6 – Matrice d'allocation des sous-porteuses pour 12 utilisateurs (en haut) et 8 utilisateurs (en bas).

Comme montré dans cette figure 3.6, la matrice du haut représente l'allocation des

sous-porteuses pour un scénario impliquant 12 utilisateurs, tandis que la matrice du bas correspond à 8 utilisateurs. Chaque matrice fournit une représentation visuelle de la manière dont les sous-porteuses sont distribuées parmi les utilisateurs du système. Ce processus d'allocation garantit que les ressources sont distribuées efficacement entre les joueurs afin d'optimiser la performance globale du système. Un nombre plus élevé d'utilisateurs peut entraîner un plus grand nombre de sous-porteuses qui se chevauchent, ce qui peut augmenter le risque de conflits ou d'interférences. En comparant les deux matrices, nous pouvons observer des différences dans l'allocation des sous-porteuses et le degré de chevauchement entre les utilisateurs pour les cas de 12 et 8 utilisateurs.

3.5 Conclusion

Ce chapitre a proposé une approche innovante d'allocation des ressources dans les réseaux acoustiques sous-marins multi-utilisateurs, combinant l'apprentissage par renforcement [QL](#) et un mécanisme coopératif basé sur la solution de Nash. Les résultats de simulation ont confirmé l'efficacité de cette méthode, avec une amélioration notable de l'utilité individuelle et collective par rapport aux approches non coopératives, tout en favorisant une répartition plus équitable des ressources ce qui est expliqué aussi par notre article [[Gharsalli et al., 2025b](#)]. Toutefois, l'augmentation du nombre d'utilisateurs a révélé des limites liées à la complexité computationnelle et aux temps de convergence, particulièrement critiques dans des environnements sous-marins dynamiques et contraints par les retards de propagation.

Pour dépasser ces contraintes, il devient nécessaire d'explorer des modèles coopératifs plus adaptés et moins coûteux en calcul, tels que les jeux de coalition. C'est dans cette perspective que le chapitre suivant s'intéressera à l'intégration des jeux de coalition avec l'apprentissage par renforcement, afin de proposer une allocation des ressources plus flexible, robuste et adaptée aux réseaux acoustiques sous-marins denses.

Chapitre 4

Jeux coopératifs à base de coalitions avec apprentissage distribué

4.1 Introduction

L'étude des approches coopératives basées sur la négociation entre paires d'utilisateurs (jeux à deux joueurs), ainsi que leur extension à des scénarios multi-utilisateurs, souligne la pertinence d'adopter une forme de coopération plus structurée : les jeux coopératifs par coalitions. Dans le contexte des réseaux acoustiques sous-marins, comme mentionné dans le premier chapitre, les communications sont fortement limitées par une bande passante restreinte, des délais de propagation importants, une énergie disponible limitée et l'instabilité des liens. Ces contraintes rendent cruciale l'optimisation collective des ressources et la coopération entre nœuds pour garantir la robustesse et l'efficacité globale du réseau.

Contrairement aux jeux coopératifs classiques basés sur des interactions bilatérales, les jeux de coalition permettent la formation de groupes dynamiques d'agents, appelés coalitions, qui collaborent pour maximiser une utilité collective, cette utilité peut refléter : une économie d'énergie, une augmentation du débit global, une optimisation de ressources (fréquence, puissance), une réduction du taux d'interférences, ou encore une meilleure fiabilité de transmission sur un ensemble de liens. Ce type de modélisation est particulièrement pertinent dans les réseaux mobiles ad hoc (MANETs), et notamment dans leur application aux réseaux acoustiques sous-marins, où certains capteurs peuvent, relayer les informations d'autres nœuds ou partager intelligemment les ressources disponibles pour renforcer la qualité du réseau tout en optimisant l'allocation de ressources disponibles. La valeur d'une coalition devient alors une fonction clé qui mesure le bénéfice collectif que ses membres peuvent atteindre ensemble, en comparaison à une stratégie individuelle ou à des coalitions alternatives.

En fonction de l'objectif principal de l'étude, les jeux de coalition sont généralement classés en trois grandes catégories :

1. **Jeu de coalition canonique (Canonical Coalition Games)** : Les jeux canoniques représentent la forme la plus fondamentale des jeux coopératifs. Dans cette catégorie, l'objectif est de former une seule et unique coalition englobant l'ensemble des agents appelée grande coalition afin de maximiser une fonction de valeur globale. Cette modélisation repose sur l'hypothèse d'une utilité transférable (TU), c'est-à-dire que les gains générés par la coopération peuvent être librement

redistribués entre les membres. Ce cadre permet l'application de nombreuses solutions issues de la théorie des jeux coopératifs, telles que la valeur de Shapley, qui évalue la contribution marginale moyenne de chaque agent à l'ensemble des coalitions possibles [Winter, 2002], ou encore le noyau (core) [Banerjee et al., 2001], qui identifie les répartitions de la valeur où aucun sous-groupe ne trouve intérêt à se séparer. Ces jeux sont particulièrement adaptés aux contextes où la collaboration entre tous les participants est techniquement réalisable et économiquement avantageuse, comme dans certains réseaux de capteurs homogènes ou des systèmes de partage de ressources centralisés. En revanche, ils ne tiennent pas compte des limitations pratiques telles que les coûts de coordination, les contraintes topologiques ou l'hétérogénéité des agents, ce qui limite leur applicabilité dans des environnements dynamiques ou distribués, comme les réseaux acoustiques sous-marins [Saad et al., 2009].

2. **Jeu de formation de coalition (Coalition Formation Games)** : Contrairement aux jeux canoniques, les jeux de formation de coalitions reconnaissent que la grande coalition n'est pas toujours réalisable ni optimale. Dans ce cadre, la formation de coalitions se fait de manière partielle et dynamique : les agents décident de coopérer avec certains partenaires, en tenant compte de leurs préférences, des coûts de coopération, des incitations locales et du contexte réseau. Le jeu devient alors un processus évolutif, où les coalitions peuvent émerger, fusionner, ou se diviser selon des règles déterminées. Plusieurs notions de stabilité y sont introduites, comme la stabilité de Nash (aucun agent ne veut quitter sa coalition actuelle), la stabilité individuelle (aucun agent ne veut changer de coalition unilatéralement), ou la stabilité par groupe (aucun sous-ensemble d'agents ne souhaite former une nouvelle coalition). Les modèles de fusion/scission (merge-and-split) permettent d'analyser ces transitions de manière itérative. Cette approche est particulièrement pertinente pour les environnements contraints et décentralisés, comme les réseaux sous-marins, où les capteurs doivent gérer leur consommation énergétique, choisir avec qui partager leurs données, et s'adapter aux variations environnementales. Elle offre une flexibilité précieuse pour intégrer la dynamique du réseau, la mobilité des nœuds, ou encore les fluctuations de qualité de service [[Sun et al., 2019]-[Bardhan and Ghose, 2013b]].
3. **Jeu de graphe de coalition (Coalition Graph Games)** : Les jeux de coalition sur graphe introduisent une nouvelle dimension à la modélisation coopérative : la structure du réseau est explicitement prise en compte à travers un graphe, où les nœuds représentent les agents et les arêtes symbolisent les possibilités d'interaction ou de communication. Dans ce cadre, seuls les agents connectés directement ou indirectement via ce graphe peuvent coopérer. La valeur d'une coalition dépend alors non seulement de ses membres, mais aussi de la topologie induite par leurs connexions. Cette approche reflète fidèlement les réalités physiques des réseaux distribués à connectivité intermittente, notamment les réseaux acoustiques sous-marins, les réseaux mobiles ad hoc (MANET) ou encore les réseaux opportunistes (Opportunistic Networks), où la portée de communication est fortement limitée, la latence élevée, et les liens sujets à des interruptions fréquentes. Par conséquent, la coopération n'est possible que dans des sous-graphes localisés, souvent restreints. Les jeux de coalition sur graphe permettent d'analyser la coopération dans des

configurations où les communications sont contraintes par la distance, le bruit, ou la topologie dynamique du réseau. Ils sont également adaptés à l'optimisation de processus localisés tels que la synchronisation, le routage collaboratif (souvent basé sur le paradigme "store-carry-forward" dans le cas opportuniste) ou la sélection de relais dans des architectures en maillage. Intégrant la structure réseau dans la théorie du jeu coopératif, cette catégorie de modèles ouvre la voie à une prise de décision plus réaliste et plus efficace dans les systèmes distribués à contraintes fortes [Guan et al., 2015]-[Zhai et al., 2012].

Dans un environnement dynamique et incertain comme l'océan, les coalitions ne peuvent pas être figées : il est nécessaire de permettre une formation et reconfiguration dynamique des coalitions, ce qui justifie l'intégration de méthodes d'apprentissage adaptatif, telles que le renforcement learning multi-agent.

4.2 Outils de formation de coalitions

La formation de coalitions repose sur des outils issus de la théorie des jeux coopératifs en forme caractéristique dont l'objectif est de regrouper les utilisateurs d'un système en sous-ensembles (coalitions) capables de coopérer pour maximiser une fonction d'utilité collective. Cette section présente les principales notions théoriques permettant la modélisation de telles coalitions, ainsi que les approches d'apprentissage permettant leur formation dynamique dans un environnement distribué et incertain tel qu'un réseau acoustique sous-marin.

4.2.1 Jeux en forme caractéristique

La théorie des jeux coopératifs en forme caractéristique fournit un cadre formel pour modéliser la coopération entre agents dans un réseau. Un jeu coopératif est défini par un couple (N, v) , où $N = \{1, 2, \dots, n\}$ est l'ensemble des agents (ou joueurs), et $v : 2^N \rightarrow \mathbb{R}$ est une fonction de valeur caractéristique qui assigne à chaque coalition $S \subseteq N$ une utilité collective $v(S)$, représentant le gain que S peut obtenir en coopérant. Dans le contexte de la Communication acoustique sous-marine (CASM), $v(S)$ peut représenter, par exemple, le débit total atteignable, la robustesse de la transmission, ou encore l'économie d'énergie obtenue par coopération entre nœuds.

Plusieurs concepts permettent de décrire comment cette valeur peut être répartie entre les agents.

1. Le cœur (Core) : Le cœur d'un jeu est l'ensemble des répartitions (ou vecteurs d'imputation) $\mathbf{x} = (x_1, \dots, x_n)$ satisfaisant :

$$\begin{aligned} \sum_{i \in N} x_i &= v(N) && \text{(Efficacité)} \\ \sum_{i \in S} x_i &\geq v(S), \quad \forall S \subseteq N && \text{(Rationalité de coalition)} \end{aligned}$$

Une répartition est dans le cœur si aucune sous-coalition S ne peut obtenir une valeur $v(S)$ supérieure à la somme de ce que ses membres reçoivent individuellement. Cela garantit la ****stabilité**** de la coopération : aucun groupe n'a intérêt à se détacher.

2. La valeur de Shapley : La valeur de Shapley est une méthode de répartition équitable basée sur la contribution marginale moyenne d'un agent à toutes les coalitions possibles. Pour chaque agent $i \in N$, sa valeur de Shapley est donnée par :

$$\phi_i(v) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(n - |S| - 1)!}{n!} [v(S \cup \{i\}) - v(S)]. \quad (4.1)$$

Cette expression pondère la contribution marginale de i à chaque coalition S par la probabilité que i rejoigne S dans un ordre aléatoire. La valeur de Shapley vérifie plusieurs axiomes d'équité (symétrie, efficacité, nullité, additivité) et est unique. Dans notre cas, cela permet de récompenser équitablement les nœuds en fonction de leur rôle dans la coopération.

3. Le nucléole : Pour une coalition $S \subseteq N$ et une répartition des gains $\mathbf{x} = (x_1, \dots, x_n)$, on définit l'*excès* de la coalition S par :

$$e(S, \mathbf{x}) = v(S) - \sum_{i \in S} x_i, \quad (4.2)$$

où $e(S, \mathbf{x})$ mesure la différence entre la valeur que la coalition S peut obtenir en agissant seule et la somme des parts qui lui sont attribuées dans \mathbf{x} .

L'idée du nucléole est de trouver une allocation \mathbf{x} qui minimise le vecteur des excès $\{e(S, \mathbf{x})\}_{S \subseteq N, S \neq \emptyset}$. Autrement dit, on cherche à réduire d'abord le mécontentement de la coalition la plus désavantagée, puis celui de la suivante, et ainsi de suite.

Formellement, le nucléole est défini comme :

$$\text{Nucl}(v) = \arg \min_{\mathbf{x} \in \mathcal{I}} \text{lex-max} (\{e(S, \mathbf{x})\}_{S \subseteq N, S \neq \emptyset}), \quad (4.3)$$

où \mathcal{I} est l'ensemble des *imputations valides*, c'est-à-dire :

$$\sum_{i \in N} x_i = v(N), \quad x_i \geq v(\{i\}) \quad \forall i \in N. \quad (4.4)$$

Le nucléole est particulièrement pertinent lorsque le cœur est vide (c'est-à-dire lorsqu'aucune répartition parfaitement stable n'existe), ou lorsque l'on souhaite garantir une équité maximale dans des environnements très hétérogènes, par exemple dans des réseaux de capteurs avec des niveaux d'énergie différents.

Ces concepts fournissent des garanties d'équité et de stabilité mais supposent une connaissance complète de la fonction $v(S)$, ce qui est difficilement réalisable dans des environnements dynamiques et distribués. Cependant, ces outils analytiques doivent être complétés par des méthodes pratiques et adaptatives pour la formation réelle des coalitions.

4.2.2 Algorithmes de formation de coalitions

1. Algorithmes de fusion et de scission (Merge and Split) : Les règles de merge-and-split (fusion et scission) permettent une formation progressive et distribuée des coalitions [Saad et al., 2008]. Ce cadre repose sur des opérations élémentaires appliquées à des partitions de l'ensemble des agents N :

- **Merge** : un groupe de coalitions disjointes $\{S_1, S_2, \dots, S_k\}$ peut fusionner en une coalition unique $S = \bigcup_{i=1}^k S_i$ si cette opération est bénéfique, selon un critère comme :

$$v(S) > \sum_{i=1}^k v(S_i). \quad (4.5)$$

Si la répartition résultante est Pareto-dominante par rapport à l'état précédent.

- **Split** : une coalition S peut se scinder en sous-coalitions $\{S_1, \dots, S_k\}$ si cette opération génère un surplus de valeur :

$$\sum_{i=1}^k v(S_i) > v(S). \quad (4.6)$$

Pour être acceptée, cette scission doit s'accompagner d'une nouvelle répartition des gains qui soit Pareto-dominante par rapport à l'état précédent (c'est-à-dire, si au moins un agent améliore sa situation sans nuire aux autres).

Le processus itératif alterne ces opérations jusqu'à atteindre un état stable, c'est-à-dire un état où aucune opération supplémentaire n'est à la fois bénéfique et mutuellement acceptable. Il est prouvé que ce processus converge vers des partitions stables, par exemple des partitions appartenant au Cœur ou stables au sens de Nash, selon les règles de répartition spécifiques.

2. Algorithmes de clustering (K-means, Density-Based Spatial Clustering of Applications with Noise (DBSCAN), etc.) Le **clustering** constitue une étape de prétraitement utile pour initier ou contraindre la formation de coalitions.

- **K-means** : chaque agent i est représenté par un vecteur de caractéristiques $\mathbf{x}_i \in \mathbb{R}^d$. L'objectif est de partitionner les agents en K clusters $\{C_1, \dots, C_K\}$ en minimisant :

$$\sum_{k=1}^K \sum_{i \in C_k} \|\mathbf{x}_i - \boldsymbol{\mu}_k\|^2, \quad (4.7)$$

où $\boldsymbol{\mu}_k$ est le centroïde du cluster C_k . Cette méthode suppose une connaissance partielle des caractéristiques globales, mais peut être distribuée si chaque agent estime localement ses distances [Ju et al., 2024].

- **DBSCAN** : cette méthode forme des clusters à partir de régions densément peuplées en agents, sans nécessiter la spécification de K [Kazemi and Soleimani, 2025]. Elle définit un ε -voisinage et un seuil $MinPts$:

Un point i est un point de cœur si :

$$|\mathcal{N}_\varepsilon(i)| \geq MinPts \quad \text{avec} \quad \mathcal{N}_\varepsilon(i) = \{j \in N : d(i, j) \leq \varepsilon\} \quad (4.8)$$

Cette approche est robuste aux topologies irrégulières, fréquentes dans les réseaux [ASM](#).

Ces méthodes permettent de réduire l'espace de recherche pour l'apprentissage par renforcement et d'accélérer la convergence.

3. Coalition Formation with Subset Stability (CFSS), Formation de coalition avec stabilité du sous-ensemble : Repose sur le principe de *stabilité par sous-ensemble*. Une partition $\Pi = \{S_1, \dots, S_k\}$ est dite **subset-stable** si aucun sous-ensemble $T \subseteq N$ ne préfère former une coalition séparée. Formellement :

$$\forall T \subseteq N, \nexists T' \subseteq T, v(T') > \sum_{i \in T'} x_i, \quad (4.9)$$

où x est la répartition actuelle

Cette méthode nécessite l'exploration de toutes les sous-coalitions, ce qui entraîne une complexité exponentielle en $|N|$. Toutefois, elle garantit des structures robustes vis-à-vis de la déviation.

Avantage : CFSS peut être utilisé comme référence pour évaluer la performance d'approches heuristiques ou d'apprentissage plus légères.

4. Jeux de coalition sur graphes : Dans les réseaux sous-marins, la coopération est limitée par la portée acoustique. Les *jeux de coalition sur graphes* modélisent cette contrainte structurelle en restreignant les coalitions admissibles aux sous-graphes connexes.

Soit un graphe $G = (N, E)$ où chaque agent est un nœud et une arête $(i, j) \in E$ signifie que les agents i et j peuvent coopérer (portée, affinité, liaison acoustique, etc.). Une coalition $S \subseteq N$ est admissible si le sous-graphe $G[S]$ est connexe.

$$\mathcal{C}_{\text{admissible}} = \{S \subseteq N \mid G[S] \text{ est connexe}\} \quad (4.10)$$

De plus, on peut pondérer les arêtes avec des valeurs w_{ij} (qualité de canal, coût énergétique) et définir une valeur de coalition dépendante du sous-graphe :

$$v(S) = \sum_{(i,j) \in E_S} w_{ij} - \text{Coût}(S). \quad (4.11)$$

Cette modélisation permet de refléter la réalité physique des réseaux ASM tout en limitant la complexité du problème de coalition formation.

Definition 3 (Sous-graphe connexe). *Soit $G = (N, E)$ un graphe. Pour un sous-ensemble de nœuds $S \subseteq N$, on note*

$$G[S] = (S, E_S) \text{ le sous-graphe induit par } S, \text{ où } E_S = \{(i, j) \in E \mid i, j \in S\}.$$

On dit que le sous-graphe $G[S]$ est connexe si, pour tout couple de nœuds $i, j \in S$, il existe un chemin reliant i à j à l'intérieur de $G[S]$. Autrement dit, tous les nœuds de S appartiennent à une seule et même composante connexe.

4.2.3 Apprentissage adaptatif pour la formation de coalitions

Dans un environnement incertain, dynamique et à forte latence tel qu'un réseau acoustique sous-marin, la formation de coalitions statiques est inefficace. Il devient alors essentiel de recourir à des approches adaptatives où les agents apprennent par l'expérience les structures de coalition les plus avantageuses. Plusieurs méthodes d'apprentissage par renforcement sont adaptées à ce contexte :

1. Coalitional Q-Learning : Cette méthode repose sur une extension du Q-learning classique à l'espace des coalitions. Chaque agent apprend une fonction de valeur $Q(S, a)$ représentant l'utilité attendue d'appartenir à une coalition S tout en exécutant une action a . La mise à jour suit la règle [Wang et al., 2019] :

$$Q(S, a) \leftarrow Q(S, a) + \alpha \left[r(S, a) + \gamma \max_{a'} Q(S', a') - Q(S, a) \right]. \quad (4.12)$$

où : $r(S, a)$ est la récompense collective de la coalition S , S' est la nouvelle coalition après action a , α est le taux d'apprentissage, γ est le facteur de discount.

Les agents sélectionnent progressivement les coalitions les plus rentables à long terme.

2. RL avec partage d'utilité (Payoff Sharing) : Dans ce cadre, les agents reçoivent une récompense globale $R(S)$ issue de la performance collective de la coalition S , mais cette récompense est redistribuée entre les membres selon une règle de partage :

$$r_i = \phi_i(R(S)) \quad \text{où} \quad \sum_{i \in S} \phi_i(R(S)) = R(S). \quad (4.13)$$

Les fonctions ϕ_i peuvent s'inspirer : de la **valeur de Shapley** (contribution marginale moyenne) [Wang et al., 2020], de politiques équitables (égalité, mérite, besoin), ou encore être apprises dynamiquement.

Ce partage d'utilité influence le comportement d'adhésion ou de départ des agents vis-à-vis d'une coalition.

3. Distributed Actor-Critic : Cette approche hybride repose sur deux entités : un *acteur* $\pi_\theta(a|s)$ qui propose des décisions (coalition à rejoindre), et un *critique* $V_\omega(s)$ qui évalue la qualité de l'état (coalition actuelle).

Chaque agent met à jour sa politique π_θ selon le gradient de politique [Ji et al., 2025] :

$$\nabla_\theta J(\theta) = \mathbb{E}_{\pi_\theta} [\nabla_\theta \log \pi_\theta(a|s) \cdot A(s, a)], \quad (4.14)$$

et la fonction de valeur V_ω par descente de gradient sur l'erreur TD :

$$\delta = r + \gamma V_\omega(s') - V_\omega(s). \quad (4.15)$$

Cette méthode permet aux agents de sélectionner dynamiquement des partenaires et d'adapter leur comportement stratégique.

4. Apprentissage basé sur l'historique des interactions : Les agents construisent une mémoire locale H_i enregistrant les coalitions formées, les récompenses obtenues, et les performances observées. Cette mémoire est utilisée pour :

- évaluer la fiabilité d'un partenaire potentiel.
- prédire les gains futurs en se basant sur l'expérience.
- éviter les coalitions passées peu performantes.

On peut intégrer une fonction de score basée sur un historique pondéré :

$$\text{score}_{i,j}(t) = \sum_{k=1}^t \beta^{t-k} \cdot r_{i,j}(k), \quad (4.16)$$

où $r_{i,j}(k)$ est la récompense issue de la coopération avec l'agent j à l'instant k , et β un facteur d'oubli.

Ces approches permettent de garantir que la formation de coalitions soit :

- **Dynamique** : les coalitions s'adaptent aux changements de topologie, de conditions de canal ou d'objectifs.
- **Décentralisée** : chaque agent prend ses décisions localement sans coordination globale, ce qui est essentiel dans un réseau [ASM](#).
- **Robuste** : l'apprentissage compense les pertes de messages, les délais, et les incertitudes de l'environnement sous-marin.

Dans la section suivante, nous passons en revue des travaux de recherche récents mettant en évidence l'utilisation de stratégies de coalition dans des environnements tels que la communication D2D, les réseaux femtocellulaires, les systèmes radar et de drones, ainsi que les réseaux acoustiques sous marins. L'accent est mis sur leur potentiel à relever les défis liés à l'efficacité et à l'optimisation de l'allocation des ressources dans des scénarios dynamiques.

4.3 État de l'art sur la formation de coalitions dans les réseaux sans fil

4.3.1 Communications [D2D](#)

Dans le domaine des communications [D2D](#), la formation de coalitions a été largement explorée comme une stratégie efficace pour améliorer l'allocation des ressources, réduire les interférences inter-utilisateurs, et accroître l'efficacité spectrale des réseaux. Contrairement aux approches centralisées classiques, les jeux de coalition permettent une prise de décision distribuée, particulièrement adaptée aux réseaux denses et hétérogènes.

L'étude présentée dans [\[Khan et al., 2010\]](#) fournit une analyse rigoureuse de la dynamique et de la stabilité des structures de coalition, en combinant un cadre de jeu de coalition avec une modélisation probabiliste à l'aide de chaînes de Markov. Cette approche permet d'observer l'évolution des coalitions dans un canal d'interférence et de déterminer les conditions d'équilibre sous contraintes dynamiques. L'intégration des communications [D2D](#) dans ce contexte favorise une communication directe entre dispositifs sans passer par l'infrastructure, ce qui contribue à une réduction significative de la latence, à une amélioration de la connectivité, ainsi qu'à une meilleure gestion du spectre disponible.

Dans [\[Ao et al., 2023\]](#), un algorithme d'allocation de ressources basé sur la théorie des jeux de coalition est proposé pour gérer le problème du changement de mode des dispositifs [D2D](#) (passage du mode cellulaire au mode direct, ou inversement). L'algorithme itère jusqu'à atteindre une structure de coalition stable au sens de l'équilibre de Nash, ce qui assure une convergence vers une solution sous-optimale mais stable. Cette méthode prend en compte à la fois le canal, la puissance de transmission, et le gain mutuel obtenu par les utilisateurs lorsqu'ils partagent leurs ressources de manière coopérative.

Une approche complémentaire est présentée dans [\[Sharma and Singh, 2020\]](#), où les

auteurs introduisent la théorie des coalitions superposées (Overlapping Coalition Formation (OCF) en anglais). Contrairement aux modèles de coalition disjoints, l'OCF permet à un utilisateur de participer simultanément à plusieurs coalitions, augmentant ainsi la flexibilité de l'allocation des ressources. Cette méthode vise à maximiser le débit agrégé du réseau, tout en respectant les contraintes d'interférence et de coordination multi-utilisateurs.

L'approche hédonique, explorée notamment dans [Wu et al., 2015], modélise de manière conjointe la sélection du mode de communication (cellulaire ou D2D) et l'allocation des ressources sous la forme d'un jeu de coalition hédonique. Dans ce type de jeu, les préférences des utilisateurs sont formulées localement, en fonction de leur gain individuel au sein d'une coalition donnée.

Chaque utilisateur évalue de manière autonome s'il doit rejoindre, quitter ou rester dans une coalition, selon la règle de changement (switch rule). Cette règle permet à un utilisateur de changer de coalition dès qu'il identifie une alternative plus avantageuse pour sa propre utilité. Le processus est entièrement distribué, ne nécessitant aucun contrôleur central, ce qui est particulièrement adapté aux réseaux dynamiques à forte densité. Ce mécanisme conduit généralement à une partition stable au sens de Nash, mais ne garantit pas l'atteinte d'un optimum global du système, car les décisions sont fondées uniquement sur des préférences individuelles et non sur des objectifs collectifs. Ainsi, l'approche hédonique favorise la simplicité et la scalabilité, mais au prix d'une possible perte d'efficacité système.

4.3.2 Réseaux Femtocellulaires

Dans les réseaux femtocellulaires, caractérisés par une forte densité de petites cellules et une gestion complexe des interférences intra-niveau (entre les Femtocell User Equipment (FUE)), la formation de coalitions s'est avérée être une solution prometteuse pour optimiser les performances globales tout en maintenant une faible charge computationnelle. L'étude de [Shi et al., 2015a] propose une approche coopérative dans laquelle les utilisateurs d'une même femtocellule forment des coalitions pour coordonner leurs transmissions en liaison montante. L'objectif est de réduire les interférences mutuelles tout en maximisant la capacité du système. L'algorithme de formation de coalition proposé est distribué, ce qui permet aux utilisateurs de prendre des décisions autonomes sans nécessiter de coordination centralisée. De plus, le modèle intègre les coûts énergétiques dans la fonction d'utilité, ce qui améliore sa pertinence dans des contextes à ressources limitées, notamment pour les déploiements à grande échelle.

Cependant, cette approche reste sensible à la structure du réseau et aux conditions de propagation, d'où l'intérêt d'explorer des alternatives plus globalement optimales. À ce titre, l'algorithme **RECORD**, introduit dans [Ma et al., 2013], propose une méthode exhaustive de formation de coalitions, dans laquelle toutes les configurations possibles sont explorées à l'aide des nombres de Bell. Cette technique vise à maximiser l'utilité totale du réseau en identifiant la meilleure structure de coalition possible, tout en optimisant la distribution de la puissance entre les membres. Bien que RECORD garantisse une solution quasi-optimale, son principal inconvénient réside dans sa complexité computationnelle exponentielle : en effet, le nombre de partitions possibles croît très rapidement avec le nombre d'utilisateurs, rendant son application impraticable dans des environnements de

grande taille.

Pour résoudre cette limitation, les auteurs de [Shi et al., 2015b] proposent une approche plus scalable basée sur le **Received Signal Strength Indicator (RSSI)**. Cette méthode repose sur une évaluation locale des conditions de canal, en utilisant la puissance des signaux reçus et les niveaux d'interférence observés pour guider la formation des coalitions. L'algorithme ainsi proposé permet une prise de décision rapide et décentralisée, ce qui le rend particulièrement adapté aux scénarios dynamiques à grand nombre de nœuds. Bien que cette méthode ne garantisse pas une optimalité globale comme RECORD, elle présente un bon compromis entre performance, complexité et adaptabilité.

4.3.3 Réseaux à base de drones et réseaux radar

Les avancées récentes en matière de drones à petites cellules (small-cell drones) et de réseaux radar introduisent de nouveaux défis pour la gestion des ressources, notamment dans des environnements distribués, dynamiques et soumis à des contraintes strictes d'énergie et d'interférence. La formation de coalitions s'y révèle une stratégie pertinente pour coordonner efficacement les transmissions et minimiser les conflits d'allocation.

Dans le contexte des réseaux de drones, l'étude menée dans [Bui et al., 2023] aborde le problème de l'évolutivité par le biais d'un mécanisme de regroupement des drones selon leur proximité géographique. Cette phase de clustering spatial est suivie d'une allocation optimisée des sous-canaux en utilisant la méthode hongroise, bien connue pour sa capacité à résoudre des problèmes d'affectation bipartite de manière optimale en temps polynomial. Ensuite, un jeu de Stackelberg est mis en œuvre pour réguler la puissance de transmission de façon hiérarchique : les drones leaders (ou maîtres) prennent d'abord leurs décisions, suivis des drones suiveurs, ce qui permet un contrôle de puissance distribué tout en réduisant les interférences mutuelles. Cette approche hiérarchique est particulièrement efficace dans des contextes où les rôles peuvent être dynamiquement assignés selon les capacités ou la position des nœuds.

Du côté des réseaux radar, [Dai et al., 2023] propose une approche coopérative fondée sur la théorie des jeux de coalition pour résoudre conjointement les problèmes d'allocation de puissance et de sélection de cibles dans un système multi-radar multi-cible. Le principal objectif est de minimiser la puissance émise par les radars tout en garantissant un certain niveau de qualité de service (QoS) dans la détection et le suivi des cibles. Pour cela, le TDMA est utilisé au sein des coalitions pour coordonner les transmissions radar, évitant ainsi les chevauchements temporels et limitant les interférences intra-coalition.

Ces travaux démontrent que la formation de coalitions, couplée à des stratégies d'optimisation distribuée (comme les jeux hiérarchiques ou les méthodes d'appariement), constitue une approche prometteuse pour améliorer la robustesse, l'efficacité énergétique et la qualité des transmissions dans des réseaux mobiles hétérogènes.

4.3.4 Limites des approches classiques

Malgré leur pertinence théorique et leurs résultats encourageants dans des cas spécifiques, les approches classiques basées sur la formation de coalitions, telles que les jeux hé-

doniques et l'algorithme RECORD, présentent plusieurs limitations majeures lorsqu'elles sont appliquées à des environnements complexes, dynamiques, ou de grande échelle.

Les jeux hédoniques, bien qu'attrayants par leur nature distribuée et leur faible complexité, reposent sur des préférences individuelles locales, ce qui limite leur capacité à prendre en compte les effets externes globaux du réseau. En conséquence, ils conduisent souvent à des équilibres localement stables, mais sous-optimaux du point de vue de la performance globale. De plus, l'absence de coordination entre les coalitions limite la possibilité de partage efficace des ressources au niveau système.

D'un autre côté, l'algorithme RECORD offre une exploration exhaustive de toutes les structures de coalitions possibles en utilisant la formule des nombres de Bell. Cette méthode garantit, en théorie, une solution globalement optimale en matière d'allocation de ressources et de répartition de puissance. Toutefois, son principal inconvénient réside dans sa complexité exponentielle, qui devient rapidement inapplicable dans des réseaux à grande échelle, en raison du coût computationnel très élevé induit par le nombre de partitions croissant de manière explosive avec le nombre de participants.

Ces constats mettent en lumière les limites des approches traditionnelles, qu'elles soient purement centralisées ou complètement décentralisées. En effet, la complexité croissante des environnements, notamment dans les réseaux sous-marins où les conditions changent rapidement et où les ressources sont rares, exige des solutions hybrides, à la fois adaptatives et intelligentes. Ces approches doivent savoir tirer parti des forces des méthodes distribuées notamment la flexibilité, la robustesse et l'autonomie des agents tout en garantissant une performance globale élevée et cohérente à l'échelle du réseau.

Dans cette perspective, l'intégration de l'intelligence artificielle, en particulier de l'apprentissage par renforcement, dans les jeux de coalition apparaît comme une solution efficace. Elle permet de concevoir des mécanismes de coopération dynamiques, capables de s'adapter aux incertitudes de l'environnement tout en optimisant la performance collective.

4.3.5 Intelligence Artificielle dans les Jeux de Coalition

Avec la progression des nouvelles technologies, l'intégration de l'intelligence artificielle dans les réseaux sans fil fait l'objet d'études approfondies en raison de son potentiel à améliorer la gestion des ressources et les performances globales du réseau. En particulier, la combinaison des techniques de RL avec la théorie des jeux de coalition ouvre un cadre innovant pour résoudre les problèmes multi-agents. Cette intégration permet aux agents de développer des stratégies optimales qui maximisent non seulement leur utilité individuelle, mais encouragent également la coopération et la formation dynamique de coalitions.

Par exemple, dans l'étude [Chalkiadakis and Boutilier, 2004], l'apprentissage bayésien par renforcement est utilisé pour améliorer la formation de coalitions. Les agents peuvent ainsi construire et mettre à jour leurs croyances concernant les capacités des autres agents dans des situations d'incertitude, ce qui améliore progressivement le processus de négociation et la constitution des coalitions. Cette approche est particulièrement adaptée aux

systèmes multi-agents décentralisés, c'est-à-dire des systèmes constitués d'agents autonomes qui prennent des décisions indépendantes tout en ayant une perception locale de leur environnement. Historiquement issue de l'intelligence artificielle, la notion de système multi-agents considère chaque agent comme un élément capable d'agir, de percevoir et de raisonner localement, tout en interagissant avec d'autres agents. Dans des environnements complexes, chaque agent peut ainsi prendre des décisions basées sur ses préférences et sur les bénéfices partagés, comme c'est le cas dans l'allocation de tâches ou la gestion collaborative des ressources.

D'autres recherches récentes explorent également l'utilisation de [MARL](#) pour étudier comment les agents peuvent négocier efficacement et former des coalitions solides. Les travaux [[Taywade, 2021](#)] et [[Chalkiadakis and Boutilier, 2004](#)] sont basés sur les fondements théoriques qui permettent d'appliquer le [MARL](#) à ces contextes, mettant en lumière les avantages tels que l'adaptabilité et la robustesse ainsi que les limites par rapport aux méthodes classiques basées sur la théorie des jeux. En particulier, l'étude [[Taywade, 2021](#)] analyse les conditions sous lesquelles le [MARL](#) apporte une réelle valeur ajoutée en termes de performances et de convergence des stratégies.

Par ailleurs, des méthodes plus avancées comme l'apprentissage profond par renforcement double ont été appliquées dans des cadres coopératifs multi-agents, notamment pour optimiser l'allocation du spectre et de la puissance dans les réseaux sans fil. Ces approches ont démontré une amélioration significative des débits de transmission [[Huang et al., 2021](#)]. Cependant, malgré leurs avantages en termes de scalabilité et de capacité d'adaptation à des environnements dynamiques, ces méthodes présentent aussi des défis majeurs, notamment des coûts computationnels élevés et des besoins importants en données pour l'entraînement, ce qui limite leur déploiement dans des dispositifs sous-marins aux ressources restreintes.

Dans ce qui suit, nous examinons un domaine particulièrement représentatif de ces enjeux : les réseaux acoustiques sous-marins, où les communications sont limitées par des contraintes physiques spécifiques et où la coopération entre nœuds devient essentielle. La formation de coalitions dans ce contexte permet non seulement d'assurer une communication efficace, mais aussi d'optimiser l'utilisation des ressources rares et de répondre aux défis posés par l'environnement marin.

4.3.6 Réseaux Acoustiques Sous-Marins et Jeux de Coalition

Dans les environnements acoustiques sous-marins, les communications sont limitées par des contraintes physiques spécifiques (faible bande passante, latence élevée, bruit, etc.), rendant les mécanismes classiques inefficaces voire inapplicables comme cité précédemment. Dans ce cadre, la formation de coalitions devient une stratégie nécessaire. En permettant à des agents comme des capteurs ou des véhicules autonomes sous-marins ([AUVs](#)) de coopérer pour effectuer des tâches de communication et d'échange d'informations, ces coalitions visent à optimiser l'utilisation des ressources et améliorer la performance globale du réseau.

Les auteurs de [[Muhammed et al., 2018](#)] proposent un état de l'art des principales stratégies de coopération dans les réseaux sous marins, en analysant les défis spécifiques à

ces réseaux et en suggérant des pistes prometteuses pour la recherche future. Ils mettent notamment en évidence les limites des approches non coopératives, et défendent l’usage de la théorie des jeux comme cadre pertinent pour modéliser les comportements collaboratifs dans des environnements aussi incertains. Dans cette optique, certaines recherches récentes se sont appuyées sur une combinaison entre théorie des jeux et apprentissage par renforcement pour proposer des solutions plus robustes. Par exemple, le travail présenté dans [Cui et al., 2023] s’intéresse à l’utilisation de l’apprentissage profond par renforcement pour adapter dynamiquement la modulation dans les communications acoustiques sous-marines, en fonction des variations de l’environnement. L’objectif est d’augmenter le débit, réduire le taux d’erreur (bit error rate (BER)) et maintenir une bonne qualité de transmission dans des conditions très variables.

En parallèle, l’étude de [Han et al., 2018] propose une approche d’allocation des ressources combinant théorie des jeux et apprentissage hiérarchique distribué. Le modèle vise à résoudre le problème conjoint d’allocation de canal et de puissance dans un scénario multi-utilisateur. Toutefois, dans ces modèles, les agents prennent souvent des décisions de manière autonome, avec parfois peu d’informations, ou en s’appuyant sur des intermédiaires comme des relais ou entités centrales. Cette architecture peut engendrer des pertes d’efficacité, surtout lorsque les conditions de propagation du canal varient rapidement et que la coordination directe entre agents est essentielle.

Malgré les progrès réalisés dans la formation de coalitions et l’allocation intelligente des ressources, plusieurs défis persistent : la complexité algorithmique, la dépendance vis-à-vis d’entités centralisées, et une adaptabilité encore limitée aux contextes très dynamiques. De plus, certaines approches prennent en compte les préférences individuelles des agents, mais manquent de la flexibilité nécessaire pour s’ajuster efficacement aux conditions changeantes. C’est précisément dans cette optique que s’inscrit notre contribution. Nous proposons un algorithme innovant, spécifiquement conçu pour les réseaux de communication ASM, qui combine la théorie des jeux de coalition avec l’apprentissage par renforcement, en établissant une analogie avec un modèle multi-agent, où chaque acteur peut être considéré comme une entité autonome de décision. Cette solution hybride vise à optimiser la formation des coalitions et l’allocation des ressources, tout en s’adaptant dynamiquement aux variations de l’environnement et en prenant en compte les capacités et préférences spécifiques des agents. Grâce à cette démarche, nous cherchons à concilier autonomie locale des agents et coordination globale efficace, un équilibre essentiel pour répondre aux enjeux des réseaux complexes et hétérogènes.

Dans ce qui suit, nous appliquons ces concepts à un cas d’étude concret : l’allocation efficace des ressources dans un réseau multi-utilisateurs. Pour simplifier l’analyse initiale, nous considérons un scénario où l’information sur l’état des canaux CSI est parfaitement connue de tous les utilisateurs. Cette hypothèse permet de se concentrer sur la dynamique de coalition et l’optimisation des ressources sans ajouter la complexité de l’incertitude sur le canal, qui sera abordée dans la section 4.6.

4.4 Jeux coopératifs de coalition pour l'allocation des ressources (Connaissance parfaite de CSI)

4.4.1 Introduction aux jeux de coalition et hypothèses appliquées

Dans la continuité des approches non coopérative et coopérative par négociation explorées précédemment, nous conservons le même modèle physique de système : il s'agit toujours d'un réseau acoustique sous-marin multi-utilisateurs, avec des transmissions multi-porteuses dans un environnement statique ou faiblement dynamique. Toutefois, l'approche que nous développons ici repose sur la théorie des jeux en coalition, dans laquelle les utilisateurs peuvent former des groupes pour coopérer plus efficacement. Ce changement de cadre implique certaines hypothèses supplémentaires par rapport aux modèles précédents notamment sur la capacité à échanger de l'information et à se coordonner, mais les contraintes physiques et le cadre de transmission de base restent inchangés, assurant ainsi la cohérence de l'analyse globale.

Lorsque le nombre d'utilisateurs K augmente et que la complexité du système s'accroît, une structuration naturelle du réseau peut émerger sous forme de coalitions, comme illustré dans la Figure 4.1. L'ensemble des utilisateurs \mathcal{K} est alors divisé en sous-ensembles formant une structure de coalition $\mathcal{C} = \{C_1, \dots, C_{N_c}\}$, représentant les groupes coopératifs issus de cette partition.

Chaque utilisateur $i \in \mathcal{K}$ peut rejoindre une coalition $C_{N_c} \subseteq \mathcal{C}$ selon des critères définis dans la section sur la formation des coalitions. Cette approche permet de mutualiser les ressources, réduire les interférences, et améliorer les performances globales du réseau, en particulier dans des environnements dynamiques ou fortement peuplés.

Afin d'assurer la validité du modèle dans un contexte réaliste de communication sous-marine, les hypothèses suivantes sont adoptées :

Hypothèses physiques communes aux modèles étudiés :

- Le canal acoustique est supposé quasi-stationnaire pendant la transmission d'un bloc (temps de cohérence long), avec absence d'effet Doppler significatif (nœuds statiques ou faiblement mobiles).
- L'environnement acoustique reste soumis à un étalement temporel important, compensé par un temps de garde T_g supérieur à la durée de l'impulsion de canal, assurant l'absence d'ISI.
- Le schéma de transmission repose sur une architecture OFDM multi-porteuse orthogonales.

Hypothèses spécifiques au modèle en coalition :

- Les utilisateurs sont capables de coopérer activement en formant des coalitions et en partageant certaines informations.
- L'information d'état du canal (CSI) est supposée parfaitement connue pour tous les membres d'une coalition, ce qui permet d'optimiser collectivement les décisions de transmission.

- Des contraintes de puissance sont appliquées à la fois au niveau individuel (chaque utilisateur) et au niveau de chaque coalition, pour refléter les limitations réalistes dans un réseau acoustique sous-marin.

4.4.2 Formulation mathématique du problème

Considérons un ensemble $\mathcal{K} = \{1, 2, \dots, K\}$ de K utilisateurs (ou joueurs), chacun étant une paire émetteur-récepteur sous-marine équipée d'un modem OFDM. Ces utilisateurs doivent partager de manière efficace les ressources disponibles (sous-porteuses et puissance), tout en minimisant les interférences et en maximisant leur utilité.

Le problème est formulé comme un jeu de formation de coalition \mathcal{G}_c défini par les éléments suivants :

Definition 4 (Coalition). *Une coalition est un sous-ensemble $A_c \subseteq \mathcal{K}$ d'utilisateurs qui décident de coopérer entre eux. Une partition en coalitions est une famille $\mathcal{C} = \{A_1, A_2, \dots, A_{N_c}\}$ telle que :*

$$(\mathcal{H}) : \begin{cases} \bigcup_{c=1}^{N_c} A_c = \mathcal{K}, \\ A_c \cap A_{c'} = \emptyset, \quad \forall c \neq c'. \end{cases} \quad (4.17)$$

Chaque joueur appartient à une et une seule coalition.

Definition 5 (Jeu de coalition). *Un jeu de formation de coalition est défini par :*

$$\mathcal{G}_c = \{\mathcal{K}, \Omega, \mathcal{P}, U\}, \quad (4.18)$$

où :

- \mathcal{K} : ensemble des joueurs (utilisateurs),
- Ω : ensemble de toutes les partitions (coalitions) possibles satisfaisant la condition (\mathcal{H}) ,
- \mathcal{P} : ensemble des stratégies de puissance possibles $P = \{p_i^n\}_{i \in \mathcal{K}, n \in \mathcal{N}}$,
- $U(P, \mathcal{C})$: fonction d'utilité globale (ou valeur de la coalition), dépendant de la stratégie de puissance P et de la partition en coalition \mathcal{C} .

Chaque coalition A_c peut décider collectivement d'une stratégie d'allocation de puissance et de sous-porteuses optimisée pour ses membres. L'objectif est alors de :

- Maximiser l'utilité globale de la coalition $U_c(P)$,
- Répartir équitablement les gains entre les membres,
- Maintenir une structure de coalition stable, c'est-à-dire résistante aux défections.

Dans le cadre des jeux coopératifs à formation de coalitions, il est fondamental de caractériser les configurations de coalitions qui sont **viables et soutenables à long terme**. Cela implique de définir des notions de *stabilité* et d'*optimalité collective*, tout en tenant compte des incitations individuelles des joueurs.

Stabilité d'une coalition

Definition 6 (Stabilité d'une coalition). *Pour une stratégie de puissance P fixée, une partition $\mathcal{C} = \{A_1, \dots, A_{N_c}\}$ de l'ensemble des joueurs \mathcal{K} est dite **stable** si :*

1. Aucun sous-ensemble $S' \subseteq A_c$ n'a intérêt à quitter sa coalition actuelle A_c pour former une nouvelle coalition S' ,
2. Aucun joueur $i \in \mathcal{K}$ n'a intérêt à migrer de sa coalition actuelle A_c vers une autre coalition $A_{c'}$,

c'est-à-dire si :

$$u_i(P, A_c) \geq u_i(P, S'), \quad \forall i \in A_c, \quad \forall S' \subseteq A_c, \quad (4.19)$$

et

$$u_i(P, A_c) \geq u_i(P, A_{c'} \cup \{i\}), \quad \forall i \in \mathcal{K}, \quad \forall A_{c'} \neq A_c. \quad (4.20)$$

Intuition. Une coalition stable est une coalition « résistante aux défections », dans laquelle tous les membres sont suffisamment satisfaits pour ne pas chercher à quitter le groupe. Cela permet d'assurer la *pérennité de la structure coopérative*. Cette notion est essentielle pour garantir que les coalitions soient réalisables dans un réseau distribué sans autorité centrale.

Équilibre de Nash stable (S-NE)

Definition 7 (Équilibre de Nash stable (S-NE)). Une partition $\mathcal{C}^* = \{A_1^*, A_2^*, \dots, A_{N_c}^*\}$ est un **équilibre de Nash stable** si, pour une stratégie de puissance P^* (optimisée localement ou globalement), aucun joueur i ne peut améliorer son utilité en quittant sa coalition actuelle A_c^* pour rejoindre une autre coalition $A_{c'}$, c'est-à-dire :

$$u_i(P^*, A_c^*) \geq u_i(P^*, A_{c'} \cup \{i\}), \quad \forall i \in \mathcal{K}, \quad \forall A_{c'} \neq A_c^*. \quad (4.21)$$

Différence avec un équilibre de Nash classique : Dans un jeu non coopératif, l'équilibre de Nash est atteint lorsque chaque joueur choisit une stratégie optimale en réponse aux autres. Dans un jeu de coalition, l'équilibre de Nash *stable* implique non seulement que les stratégies sont optimales, mais aussi que les coalitions sont stables face aux défections individuelles.

Optimalité de Pareto

Definition 8 (Optimalité de Pareto). Une partition \mathcal{C}^* est dite *Pareto-optimale* (ou *Pareto-efficace*) s'il n'existe aucune autre partition \mathcal{C}' telle que :

$$u_i(P, \mathcal{C}') \geq u_i(P, \mathcal{C}^*), \quad \forall i \in \mathcal{K}, \quad (4.22)$$

avec une inégalité stricte pour au moins un joueur j :

$$\exists j \in \mathcal{K} \text{ tel que } u_j(P, \mathcal{C}') > u_j(P, \mathcal{C}^*). \quad (4.23)$$

Une structure de coalition est Pareto-optimale si on ne peut pas améliorer l'utilité d'un joueur sans détériorer celle d'au moins un autre. C'est une condition de *justice et d'efficacité* au niveau global. Elle ne garantit cependant pas l'absence d'incitation à quitter une coalition : une structure Pareto-optimale peut être instable si elle ne respecte pas les conditions de stabilité.

Definition 9. — Une structure **stable** est toujours un **équilibre de Nash stable** si les joueurs n'ont aucune incitation individuelle à migrer.

- Un équilibre de Nash stable **n'est pas nécessairement Pareto-optimal** : il peut être stable mais sous-optimal en termes d'utilité globale.
- Une structure **Pareto-optimale** peut être instable si des joueurs ont intérêt à former une nouvelle coalition plus avantageuse.

Exemple illustratif détaillé :

Considérons un petit réseau composé de trois utilisateurs $\mathcal{K} = \{1, 2, 3\}$. Chaque utilisateur peut choisir de coopérer avec d'autres pour améliorer son utilité (par exemple, en réduisant les interférences, en partageant des ressources, ou en échangeant des informations sur les canaux).

On étudie ici deux structures possibles de coalition :

- **Structure 1 (coalitions \mathcal{C}_1)** : $\{\{1, 2\}, \{3\}\}$
- **Structure 2 (coalitions \mathcal{C}_2)** : $\{\{1\}, \{2, 3\}\}$

On suppose que chaque coalition maximise une fonction d'utilité conjointe, puis répartit le gain entre ses membres. L'utilité individuelle des joueurs dans chaque structure est donnée par les tableaux suivants (unités arbitraires) :

Utilisateur	$u_i(\mathcal{C}_1)$	$u_i(\mathcal{C}_2)$
1	5	4
2	4	6
3	3	5

Analyse de Pareto-optimalité :

Comparons les deux structures :

- Passage de \mathcal{C}_1 à \mathcal{C}_2 :
 - u_1 passe de 5 à 4 \Rightarrow perte
 - u_2 passe de 4 à 6 \Rightarrow gain
 - u_3 passe de 3 à 5 \Rightarrow gain
- \mathcal{C}_2 améliore l'utilité de 2 et 3, mais au détriment de 1. Il s'agit donc d'une amélioration **non-Pareto** puisque 1 est désavantagé.
- \mathcal{C}_1 est donc **Pareto-optimale**, car aucune autre structure ne permet d'améliorer l'utilité d'au moins un joueur sans nuire à un autre.

Analyse de stabilité :

Dans \mathcal{C}_1 , l'utilisateur 2 reçoit une utilité de 4. Il peut envisager de quitter la coalition $\{1, 2\}$ pour former une nouvelle coalition avec 3. Cette transition vers \mathcal{C}_2 lui permettrait d'obtenir 6 (gain), et 3 aussi passerait de 3 à 5 (gain), donc tous deux y gagnent.

- Cette déviation est **avantageuse** pour les deux membres de la nouvelle coalition $\{2, 3\}$.
- Donc, \mathcal{C}_1 est **instable**, même si elle est Pareto-optimale.

Analyse de l'équilibre de Nash stable (S-NE) :

Dans \mathcal{C}_2 , chaque utilisateur ne peut pas augmenter son utilité en migrant seul :

- 1 est seul avec une utilité de 4. Il ne peut pas rejoindre $\{2, 3\}$ sans leur accord, et ne gagne rien à rejoindre 2 ou 3 seuls.
- 2 est avec 3, et obtient 6. Il n'a aucun intérêt à quitter cette coalition.
- 3 est avec 2, obtient 5. S'il quittait seul, il tomberait à une utilité inférieure.

Ainsi, C_2 constitue un **équilibre de Nash stable**, car aucune migration individuelle n'est incitative. Cependant, comme mentionné précédemment, C_2 n'est pas Pareto-optimale.

Cet exemple illustre bien :

- l'**optimalité de Pareto**, qui évalue l'efficacité collective ;
- la **stabilité**, qui évalue la viabilité d'une coalition à travers les déviations ;
- l'**Équilibre de Nash stable**, qui examine les déviations *individuelles*.

4.4.3 Représentation schématique

La figure 4.1 illustre le principe de formation de coalitions dans un réseau ASM, où les utilisateurs choisissent de coopérer afin d'optimiser l'allocation des sous-porteuses et la puissance de transmission. En collaborant, les joueurs peuvent mieux coordonner l'utilisation des ressources spectrales disponibles. En effet, plutôt que d'agir individuellement et de manière égoïste ce qui crée des interférences mutuelles massives, la coalition permet une gestion concertée. Cette coordination permet d'allouer les ressources (puissance et sous-porteuses) de manière à minimiser les conflits, ce qui permet de réduire les interférences et d'améliorer les performances globales du réseau, notamment en termes de débit et d'efficacité spectrale en se regroupant en coalitions.

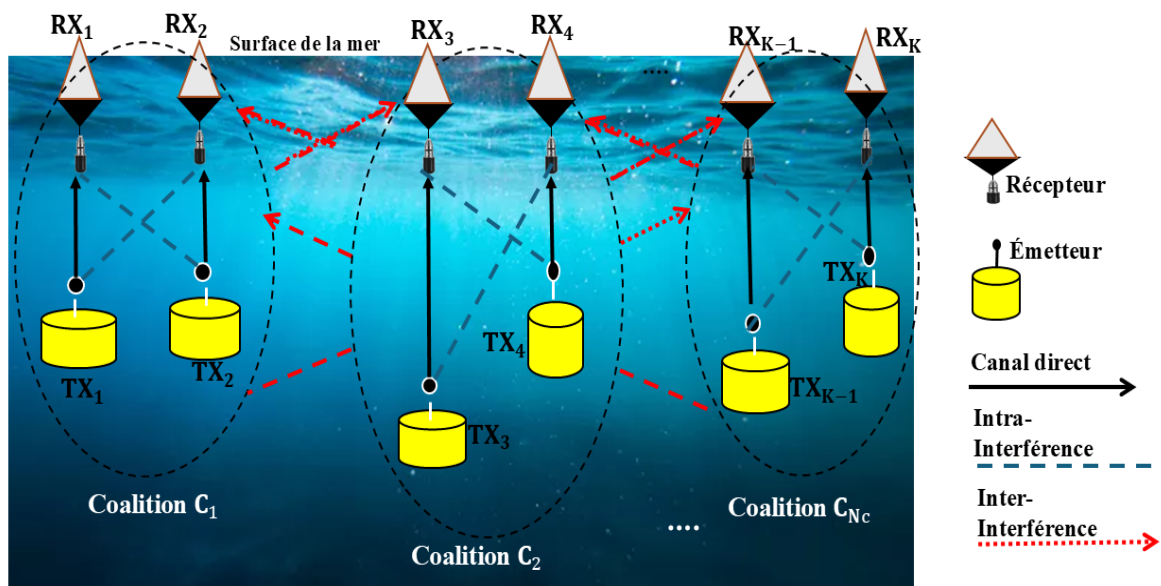


FIGURE 4.1 – Exemple de formation de coalitions dans un réseau multi-utilisateurs sous-marin.

Pour les jeux de coalitions, on distingue deux niveaux de coopération :

- communication coopérative intra-coalition : entre membres d’une même coalition, favorisant la coordination locale pour une gestion efficace des ressources.
- communication coopérative inter-coalition : entre différentes coalitions, visant à réduire les conflits d’allocation et améliorer l’équité globale.

Afin d’approfondir l’analyse de l’allocation des ressources via les jeux de coalitions et d’optimiser les performances du système, la section suivante introduit la représentation mathématique formelle du problème.

4.4.4 Représentation mathématique

Le but de ce type de jeu de coalition est donc de trouver une partition $\mathcal{C}^* \in \Omega$ et une stratégie de puissance P^* telles que :

$$\mathcal{C}^*, P^* = \arg \max_{\mathcal{C} \in \Omega, P \in \mathcal{P}} U(P, \mathcal{C}), \quad (4.24)$$

tout en assurant la stabilité des coalitions et une certaine équité entre les joueurs.

Une formulation mathématique détaillée du problème d’optimisation associé à ce jeu de coalition, sera expliquée suivie d’une solution algorithmique distribuée.

Afin de modéliser ces interactions coopératives et de formaliser le partage des ressources, nous introduisons une fonction d’utilité globale $U(P, \mathcal{C})$, qui dépend de :

- la répartition des puissances de transmission $P = \{p_{i,n}\}$;
- la structure de coalition $\mathcal{C} = \{A_1, A_2, \dots, A_{N_c}\}$.

Pour simplifier le problème d’optimisation et assurer sa faisabilité, les hypothèses suivantes sont posées :

Hypothèses :

1. **Puissance maximale par utilisateur** : Chaque utilisateur i appartenant à une coalition A_c dispose d’une puissance maximale de transmission P^{\max} , qu’il ne doit pas dépasser.
2. **Taille maximale de coalition** : Pour des raisons de gestion de ressources et de complexité de calcul, le nombre d’utilisateurs par coalition est borné par une valeur maximale N_c^{\max} .
3. **Connaissance du canal** : CSI est supposée parfaitement connue du côté récepteur, ce qui permet un calcul fiable du SINR pour chaque sous-porteuse.

Définition de la fonction d’utilité $U(P, \mathcal{C})$

Le critère de performance utilisé ici repose sur SINR, qui mesure la qualité du lien pour une sous-porteuse donnée. Pour un utilisateur $i \in A_c$ sur la sous-porteuse $n \in N_i$, le SINR est défini par :

$$\gamma_{i,n}(P, A_c) = \frac{|h_{ii}(n)|^2 p_{i,n}}{\sum_{j \in A_c, j \neq i} \xi |h_{ji}(n)|^2 p_{j,n} + \sigma_{i,n}^2}, \quad (4.25)$$

où :

- $|h_{ii}(n)|^2$ est le gain du canal direct entre l'émetteur i et son récepteur sur la sous-porteuse n ,
- $|h_{ji}(n)|^2$ représente le gain du canal entre un utilisateur j (interfèrent) et le récepteur de i ,
- $\sigma_{i,n}^2$ est la puissance du bruit,
- ξ est un facteur d'ajustement de l'impact des interférences (par exemple lié à l'atténuation acoustique).

Le débit moyen (ou utilité) de l'utilisateur i est alors :

$$u_i(P, A_c) = \frac{1}{N_i} \sum_{n \in N_i} \log_2 (1 + \gamma_{i,n}(P, A_c)), \quad (4.26)$$

ce qui reflète le débit moyen par sous-porteuse allouée.

L'utilité globale d'une coalition A_c est simplement la somme des utilités de ses membres :

$$U(P, A_c) = \sum_{i \in A_c} u_i(P, A_c). \quad (4.27)$$

L'objectif global est donc de maximiser la somme des utilités de toutes les coalitions, soit :

$$U(P, C) = \sum_{A_c \in C} U(P, A_c). \quad (4.28)$$

Le problème consiste à déterminer :

- la puissance optimale P^* allouée à chaque sous-porteuse,
- la structure de coalition optimale C^* ,

telles que l'utilité globale $U(P, C)$ soit maximale, tout en respectant un ensemble de contraintes.

$$\left\{ \begin{array}{l} \max_{P, C \in \Omega} U(P, C) = \max_{P, C \in \Omega} \sum_{A_c \in C} U(P, A_c) \\ \text{s.t.} \left\{ \begin{array}{l} \sum_{n \in N_i} p_{i,n} \leq P^{\max}, \forall i \in A_c \quad (4.29.a) \\ \sum_{i \in A_c} \sum_{n \in N_i} p_{i,n} \leq P_{C, \max} \quad (4.29.b) \\ I_{C, \min} \leq \sum_{i \in A_c} \sum_{n \in N_i} I_{i,n} \leq I_{C, \max} \quad (4.29.c) \\ |C| \leq N_C^{\max} \quad (4.29.d) \\ u_i(P, A_{c'}) > u_{i, \min}, \forall i \in A_c, \forall C' \neq C \quad (4.29.e) \end{array} \right. \end{array} \right. \quad (4.29)$$

Interprétation des contraintes :

- **(a)** : Puissance maximale par utilisateur.
- **(b)** : Puissance totale maximale autorisée pour chaque coalition.
- **(c)** : L'interférence produite par une coalition est bornée pour garantir la qualité de service globale.
- **(d)** : Limite sur la taille des coalitions pour éviter la surcharge.
- **(e)** : Un utilisateur ne rejoint une coalition que si cela améliore son utilité.

Existence d'un équilibre stable de Nash (S-NE)

L'objectif est d'atteindre une solution stable où aucun utilisateur n'a intérêt à modifier sa puissance ou sa coalition actuelle. On montre que cette solution existe et est unique sous les hypothèses suivantes :

Proposition 2. *Il existe un équilibre stable (P^*, C^*) maximisant l'utilité globale $U(P, C)$ sous les contraintes (4.29).a–(4.29).d.*

Démonstration. — L'espace des stratégies admissibles est convexe, fermé et non vide, donc il existe au moins une solution faisable.

— La fonction $U(P, C)$ est concave par rapport à P . En effet, la dérivée première est :

$$\frac{\partial U(P, C)}{\partial p_{i,n}} = \frac{1}{N_i \ln 2} \cdot \frac{|h_{ii}(n)|^2}{|h_{ii}(n)|^2 p_{i,n} + \Delta_{i,n}}, \quad (4.30)$$

avec $\Delta_{i,n} = \sum_{j \in A_c, j \neq i} |h_{ji}(n)|^2 p_{j,n} + \sigma_{i,n}^2$.

La dérivée seconde étant négative :

$$\frac{\partial^2 U(P, C)}{\partial p_{i,n}^2} = -\frac{1}{N_i \ln 2} \cdot \frac{|h_{ii}(n)|^4}{(|h_{ii}(n)|^2 p_{i,n} + \Delta_{i,n})^2} \leq 0, \quad (4.31)$$

cela prouve la concavité de $U(P, C)$ et donc l'existence d'un maximum global. \square

Notion d'optimalité de Pareto : Une fois la stabilité atteinte, la configuration (P^*, C^*) est dite **Pareto-optimale** si aucun utilisateur ne peut améliorer sa propre utilité sans détériorer celle d'un autre. Ce critère est essentiel pour valider l'équité et l'efficacité de l'allocation proposée.

La partie suivante présente l'algorithme distribué proposé pour atteindre cette solution, ainsi qu'une comparaison avec des stratégies existantes.

4.5 Solution proposée et intégration du Q-Learning pour l'optimisation des coalitions

La solution proposée **U-CoCRL** vise à réconcilier la satisfaction individuelle des utilisateurs avec l'efficacité globale du réseau, tout en maintenant une scalabilité acceptable. Il repose sur trois piliers fondamentaux : L'utilisation de K-means pour la formation initiale de coalitions. L'intégration de la valeur de Shapley, pour assurer une répartition équitable des gains au sein des coalitions avec un algorithme de négociation de puissance pour réduire les interférences intra- et inter-coalitions. L'utilisation de l'apprentissage par renforcement **QL**, permettant aux agents d'apprendre dynamiquement à former des coalitions et à ajuster leurs décisions en fonction de l'évolution de l'environnement. La Figure 4.2 illustre le mécanisme de cette solution.

L'algorithme adopte une approche dynamique et adaptative pour la formation de coalitions et l'allocation des ressources dans les systèmes **OFDM**, comme illustré à la Figure 4.2. Le processus s'initie par une phase de clustering (Étape 1) utilisant l'algorithme

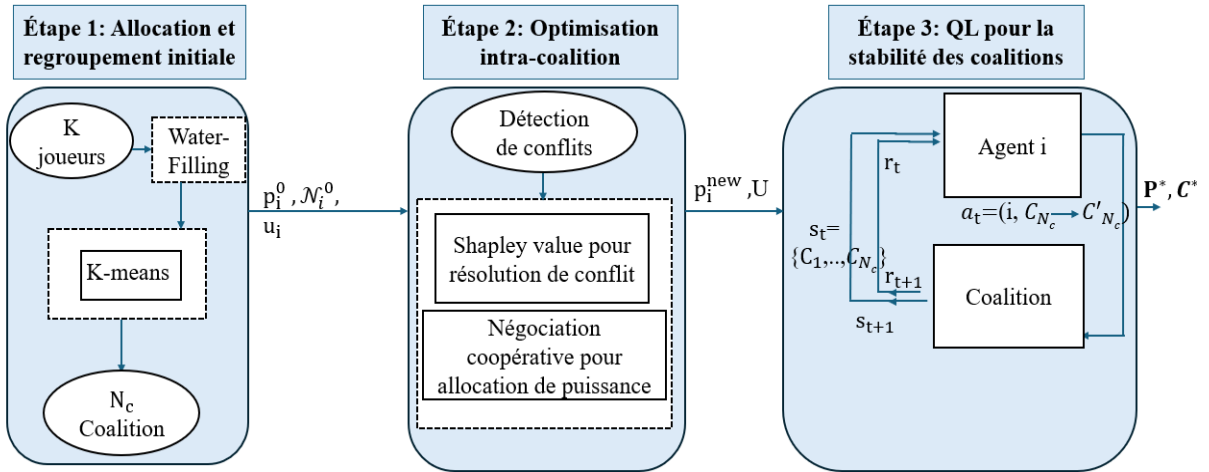


FIGURE 4.2 – Processus de la solution proposée.

K-means. Ce choix a été privilégié pour sa faible complexité calculatoire et son efficacité pour partitionner les utilisateurs en N_c coalitions initiales, par rapport à d'autres approches comme le clustering hiérarchique ou [DBSCAN](#), qui sont plus coûteuses ou moins adaptées à une définition a priori du nombre de coalitions.

La solution [U-CoCRL](#) proposée combine l'équité, la robustesse et l'adaptabilité. En effet, la valeur de Shapley (Étape 2) est largement reconnue en théorie des jeux coopératifs pour son efficacité à garantir une répartition équitable des gains entre les joueurs basée sur la contribution marginale de chaque joueur. Plusieurs travaux l'ont exploitée dans des contextes de communication collaborative, mais sans l'associer à un mécanisme d'adaptation dynamique, ce qui limite son efficacité dans des environnements très fluctuants tels que l'[ASM](#). L'algorithme de négociation de puissance répond quant à lui aux limites des approches centralisées de contrôle de puissance, souvent coûteuses en signalisation et peu adaptées aux réseaux distribués, il permet une gestion optimale des interférences tout en maintenant l'autonomie décisionnelle des utilisateurs, un aspect rarement pris en compte dans les solutions existantes.

Enfin, l'apprentissage par renforcement (Étape 3), et en particulier le [QL](#), a été retenu car il permet aux joueurs d'explorer et d'exploiter leur environnement de manière autonome. Bien que des méta-heuristiques, telles que les algorithmes génétiques (GA) ou l'optimisation par essaims particulaires ([PSO](#)), puissent être utilisées pour l'optimisation des coalitions, elles convergent souvent plus lentement et sont moins adaptées à une prise de décision séquentielle en temps réel. Contrairement aux approches statiques comme les jeux hédoniques ou le modèle [RECORD](#), le [QL](#) offre une adaptabilité en temps réel aux conditions changeantes du réseau, répondant ainsi à une lacune identifiée dans plusieurs études récentes.

L'explication de chaque phase sera détaillée dans la partie suivante.

4.5.1 Allocation et formation de coalition initiale

Étape 1 : Allocation de puissance par la méthode du Water-Filling (WF) : Dans une première phase, l'algorithme WF est utilisé pour effectuer l'allocation de puis-

sance entre les utilisateurs. Cette méthode permet de distribuer la puissance disponible de manière optimale, en tenant compte des contraintes de puissance et des conditions des canaux, afin d'assurer que chaque utilisateur atteigne le niveau de **SINR** requis.

Ainsi, pour chaque utilisateur i et pour chaque sous-porteuse n , l'allocation de puissance est donnée par :

$$p_{i,n} = \left[\frac{1}{\lambda_i} - \frac{1}{\gamma_{i,n}(P, A_c)} \right]^+, \quad (4.32)$$

où λ_i est un multiplicateur de Lagrange spécifique à l'utilisateur i , garantissant le respect de la contrainte de puissance maximale.

Étape 2 : Calcul du **SINR et du débit de transmission** : Une fois la puissance allouée, le **SINR** ainsi que le débit de transmission sont calculés pour chaque utilisateur à partir des équations (4.25) et (4.26). Ces mesures permettent d'évaluer l'efficacité de l'allocation et d'orienter la suite du processus de formation des coalitions.

Étape 3 : Formation initiale des coalitions via l'algorithme K-means : Pour former les coalitions initiales, l'algorithme de classification non supervisée K-means est appliqué afin de regrouper les utilisateurs en \mathcal{N}_c coalitions, en se basant sur leur utilité individuelle. Après avoir initialisé les centroïdes des \mathcal{N}_c clusters, chaque utilisateur est affecté à la coalition dont le centroïde est le plus proche, minimisant ainsi la distance entre ses caractéristiques et le centre du groupe.

L'algorithme est exécuté plusieurs fois avec des initialisations différentes pour améliorer la convergence et retenir la meilleure configuration obtenue. Une fois les affectations terminées, les coalitions sont formées en regroupant les utilisateurs partageant le même indice de cluster. Cette procédure permet d'obtenir une répartition initiale cohérente des utilisateurs dans les coalitions, comme détaillé dans l'**Algorithme 3** présenté dans l'annexe.

Une fois cette étape de formation initiale des coalitions complétée à l'aide du K-means, la phase suivante consistera à résoudre les conflits d'intérêts et à optimiser l'allocation de puissance au sein de chaque coalition, comme présenté dans l'**Algorithme 4**.

Optimisation intra-coalition : Pour garantir une allocation optimale des ressources au sein de chaque coalition, nous introduisons dans un premier temps la notion de valeur de Shapley.

Valeur de Shapley : La valeur de Shapley est un concept fondamental de la théorie des jeux coopératifs, tel que présentée dans [Ribeiro Junior et al., 2022]. Elle permet d'évaluer la contribution marginale moyenne attendue de chaque joueur au sein d'une coalition, en tenant compte de toutes les combinaisons possibles avec les autres membres. Elle est définie dans notre problème comme suit [Abe and Nakada, 2023] :

$$\phi_i(A_c) = \sum_{S \subseteq A_c \setminus i} \frac{|S|!(|A_c| - |S| - 1)!}{|A_c|!} (U_S(S \cup i) - U_S(S)). \quad (4.33)$$

Dans le contexte de l'allocation des ressources, dire qu'une méthode est équitable revient à vérifier que la valeur de Shapley respecte plusieurs axiomes garantissant une répartition rationnelle et juste :

1. **Symétrie** : Si deux joueurs ont une contribution équivalente dans toutes les coalitions possibles, ils doivent recevoir la même valeur de Shapley :

$$\phi_i(A_c) = \phi_j(A_c) \quad \text{si} \quad U_S(S \cup i) = U_S(S \cup j). \quad (4.34)$$

2. **Joueur nul** : Si un joueur n'apporte aucune contribution à la valeur totale de la coalition, sa valeur de Shapley doit être nulle :

$$\phi_i(A_c) = 0 \quad \text{si} \quad U_S(S \cup \{i\}) = U_S(S). \quad (4.35)$$

3. **Additivité** : Si deux jeux de coalition indépendants sont combinés, la valeur de Shapley du jeu résultant est la somme des valeurs de Shapley de chaque jeu :

$$\phi_i(U^1 + U^2) = \phi_i(U^1) + \phi_i(U^2), \quad \forall i \in A_c \quad (4.36)$$

4. **Efficacité** : La totalité de la valeur générée par une coalition doit être entièrement répartie entre ses membres :

$$\sum_{i \in A_c} \phi_i(A_c) = U(A_c). \quad (4.37)$$

En s'appuyant sur la valeur de Shapley comme cadre équitable et efficace d'allocation des ressources, nous décrivons maintenant son application dans notre stratégie de résolution des conflits, d'optimisation dynamique de la puissance et d'amélioration de la performance globale de la coalition.

4.5.2 Résolution des conflits et allocation de puissance basée sur la valeur de Shapley

Étape 1 : Résolution des conflits : Une fois les coalitions formées, les valeurs de Shapley sont approximées via un échantillonnage Monte-Carlo selon (4.33). Cette approche permet de déterminer la contribution marginale équitable de chaque utilisateur en calculant la moyenne de leurs contributions sur plusieurs permutations aléatoires. Elle réduit considérablement la complexité de calcul tout en maintenant une bonne précision, ce qui la rend adaptée aux systèmes de grande taille [Štrumbelj and Kononenko, 2014]. Les conflits de sous-porteuses sont ensuite identifiés. Un conflit est défini lorsque deux utilisateurs $i, j \in A_c$ utilisent simultanément la même sous-porteuse n . L'ensemble des conflits est noté \mathcal{A} . Une fois les conflits détectés, l'allocation de puissance est effectuée proportionnellement à la valeur de Shapley approximée de chaque utilisateur :

$$p_{i,n}^{\text{new}} = \frac{\hat{\phi}_i(A_c)}{\sum_{j \in A_c} \hat{\phi}_j(A_c)} \cdot P^{\text{max}}. \quad (4.38)$$

Étape 2 : Négociation de puissance basée sur la valeur de Shapley : Afin d'affiner davantage l'allocation de puissance et de résoudre les déséquilibres restants, un mécanisme de négociation est mis en œuvre. Les utilisateurs ayant une valeur de Shapley élevée sont incités à redistribuer une partie de leur puissance à ceux ayant une contribution moindre. Cela garantit une allocation plus équitable, évite la domination d'utilisateurs forts et permet à ceux avec de faibles contributions d'améliorer leur impact collectif.

Cette redistribution contribue à une meilleure gestion de l'interférence et à l'amélioration globale de la satisfaction des utilisateurs.

Étape 3 : Satisfaction des utilisateurs : Le niveau de Individual satisfaction (IS) de chaque utilisateur i est défini comme suit :

$$IS_i = u_i^{\text{new}} - u_i^{\text{old}}, \quad (4.39)$$

où u_i^{new} est l'utilité après ajustement de puissance, et u_i^{old} est l'utilité initiale. Un $IS_i > 0$ indique une amélioration de performance pour l'utilisateur i , tandis qu'une valeur négative suggère la nécessité d'ajustements supplémentaires pour garantir l'équité et la stabilité de la coalition.

Afin de permettre au système d'évoluer vers une configuration optimale de manière dynamique, une phase d'apprentissage par renforcement est envisagée, notamment à travers le QL. Celui-ci permet d'ajuster les stratégies d'allocation en continu, maximisant la satisfaction des utilisateurs tout en maintenant la stabilité des coalitions ce qui est présenté par l'Algorithme 5.

4.5.3 Apprentissage par renforcement pour optimiser la répartition des utilisateurs au sein des coalitions finales

Definition 10. *Principe de l'Apprentissage par Q-Learning*

Le Q-Learning est un algorithme d'apprentissage par renforcement, permettant à un agent d'apprendre une politique optimale de prise de décision dans un environnement dynamique, sans connaître préalablement les transitions d'états. L'objectif principal du Q-Learning est de maximiser la récompense cumulée à long terme en apprenant progressivement la meilleure action à entreprendre dans chaque situation.

Concrètement, un agent (ici un utilisateur) explore un environnement représenté par un ensemble d'états $s \in \mathcal{S}$ et choisit des actions $a \in \mathcal{A}$ selon une stratégie d'exploration-exploitation (par exemple ϵ -greedy). Après chaque action, l'agent reçoit une récompense r et observe un nouvel état s' , l'information reçue est utilisée pour mettre à jour la fonction de valeur d'action, appelée fonction Q, notée $Q(s, a)$, selon la règle de mise à jour suivante :

$$Q(s, a) \leftarrow (1 - \eta) \cdot Q(s, a) + \eta \cdot \left(r + \gamma \cdot \max_{a'} Q(s', a') \right), \quad (4.40)$$

où :

- η est le taux d'apprentissage (*learning rate*), contrôlant l'importance de la nouvelle information par rapport à l'ancienne ;
- γ est le facteur d'actualisation (*discount factor*), qui pondère l'importance des récompenses futures ;
- $\max_{a'} Q(s', a')$ représente l'estimation de la meilleure récompense future possible à partir de l'état suivant.

L'algorithme Q-Learning est particulièrement adapté aux environnements où les décisions doivent être prises localement et de manière décentralisée, comme dans le cas de la formation de coalitions dynamiques dans les réseaux distribués. Il permet à chaque utilisateur d'adapter son comportement en fonction de son expérience individuelle, tout en

tenant compte des effets collectifs à travers une récompense composée de facteurs locaux (utilité individuelle) et globaux (satisfaction collective).

Dans ce travail, le Q-Learning est utilisé pour optimiser dynamiquement la répartition des utilisateurs entre les coalitions, en maximisant simultanément leur satisfaction individuelle **IS**, individual satisfaction en anglais et celle du groupe Collective satisfaction (**CS**) collective satisfaction, tout en respectant la nature décentralisée du système. Cette phase joue un rôle crucial en répondant au défi de l'optimisation simultanée de **IS** et **CS** au sein des coalitions. Ces deux objectifs sont souvent conflictuels : maximiser l'utilité individuelle peut se faire au détriment de l'utilité collective, et vice versa. Pour concilier ces deux aspects, nous adoptons une approche décentralisée basée sur le **QL**. Bien que les algorithmes de **QL** multi-agents puissent présenter des défis de convergence (notamment en termes de vitesse et de stabilité dans un environnement non-stationnaire), cette approche a été retenue. En effet, elle est la plus à même de permettre à chaque utilisateur de prendre des décisions locales fondées sur l'échange d'informations avec ses voisins, répondant ainsi à notre contrainte principale d'une solution purement décentralisée.

L'algorithme proposé intègre ces deux formes de satisfaction dans une fonction de récompense unifiée, favorisant l'apprentissage de stratégies permettant d'atteindre un équilibre optimal entre performance individuelle et performance collective ce qui est illustrer par les étapes suivantes :

Étape 1 : Chaque utilisateur i est représenté par un état s , reflétant sa coalition actuelle et l'utilité associée. Les actions a disponibles consistent à rester dans sa coalition actuelle ou à rejoindre une autre coalition C' .

Étape 2 : À chaque itération, l'utilisateur choisit une action selon une stratégie ϵ -gloutonne (*epsilon-greedy*), combinant exploration de nouvelles coalitions et exploitation des décisions les plus rentables connues. Cette décision est guidée par l'échange d'informations utilitaires avec les utilisateurs voisins, assurant un fonctionnement entièrement décentralisé.

Étape 3 : La récompense r_i est calculée comme une combinaison pondérée de **IS** et de **CS** :

$$r_i = \beta \cdot \text{IS}_i + (1 - \beta) \cdot \text{CS}. \quad (4.41)$$

où le paramètre $0 \leq \beta \leq 1$ permet d'ajuster l'importance relative de chaque composante.

Étape 4 : La valeur Q de chaque utilisateur est mise à jour selon l'équation :

$$Q_i(s, a) \leftarrow (1 - \eta) \cdot Q_i(s, a) + \eta \cdot \left(r_i + \gamma \cdot \max_{a'} Q_i(s', a') \right), \quad (4.42)$$

où η est le taux d'apprentissage, γ est le facteur d'actualisation, et $Q_i(s', a')$ représente la récompense future maximale attendue à partir du nouvel état s' .

Étape 5 : L'utilisateur peut changer de coalition si ce changement améliore soit sa satisfaction personnelle, soit l'utilité globale de la nouvelle coalition. Ce processus d'adaptation continue jusqu'à atteindre une configuration stable.

Étape 6 : La stabilité du système est assurée en évaluant si une sous-coalition C' offre à un utilisateur une utilité supérieure à celle de sa coalition actuelle C . Le transfert est effectué uniquement si cela améliore l'utilité collective, évitant ainsi des divisions non nécessaires.

Afin de mieux comprendre le fonctionnement du processus, un organigramme est proposé dans la figure suivante 4.3. Son efficacité sera évaluée à travers deux scénarios distincts.

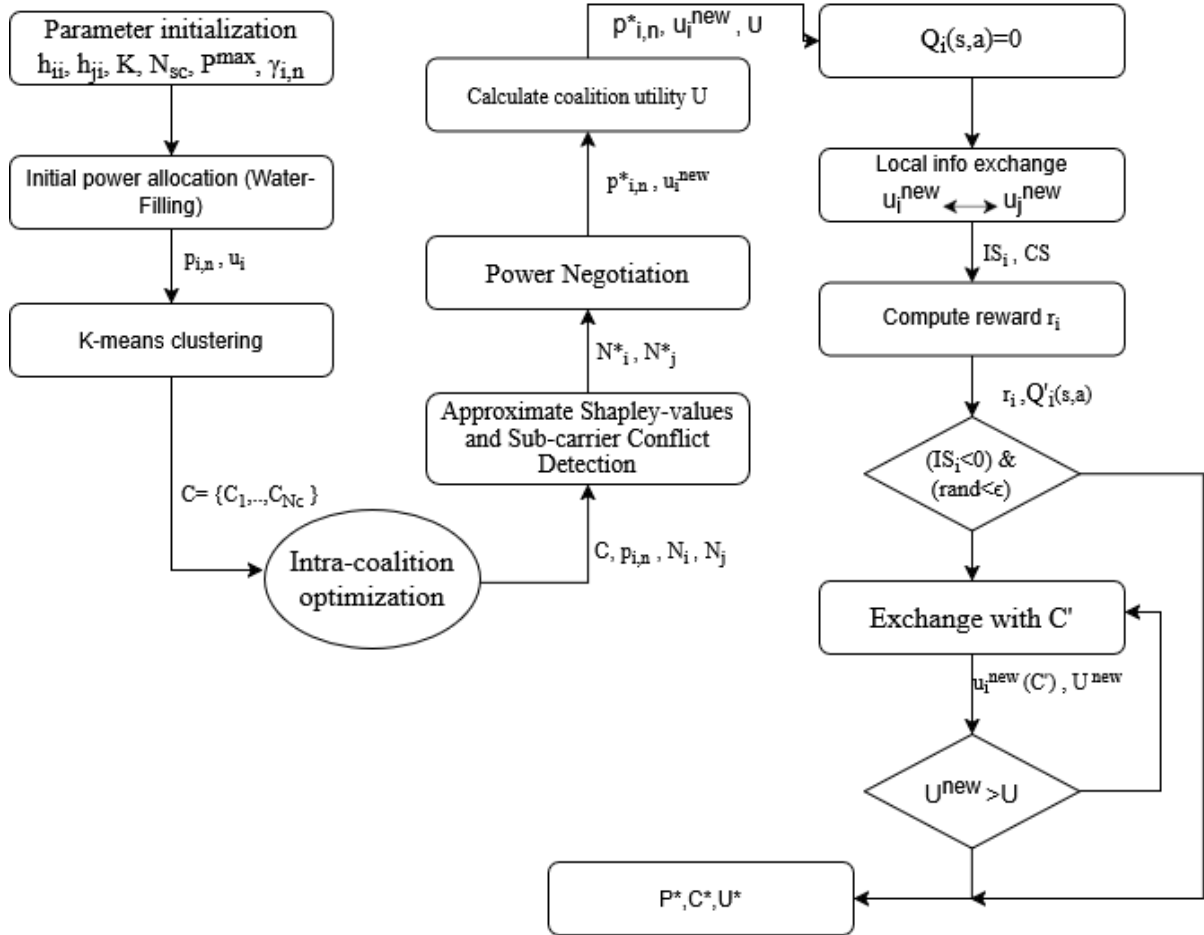


FIGURE 4.3 – Organigramme de la solution proposée.

Dans un premier scénario, nous considérons un contexte idéal où CSI est parfaitement connue, avec un nombre fixe de 14 utilisateurs. Ce cas sert de référence pour mesurer les performances optimales de l'algorithme dans des conditions idéales. Dans un second scénario, plus réaliste, l'algorithme est testé sous l'hypothèse d'une connaissance partielle et incertaine du canal (CSI partielle). Ce scénario est étudié pour deux configurations d'utilisateurs : 14 et 20, afin d'évaluer la robustesse et la scalabilité de la méthode face à un nombre plus grand des joueurs et à des environnements plus complexes et dynamiques, caractérisés par des incertitudes liées à la mesure du canal et aux variations dues à l'effet Doppler.

Exemple illustratif d'allocation de ressources en OFDM :

Pour clarifier le mécanisme d'allocation et de résolution des conflits au sein du schéma U-CoCRL, considérons une coalition A_c composée de trois utilisateurs : $i = 1$, $i = 2$ et $i = 3$, opérant sur un canal OFDM avec un ensemble limité de sous-porteuses (n_a, n_b, \dots) . La puissance maximale totale de la coalition est P^{\max} .

Phase 1 : Allocation Initiale par Water-Filling (WF) : Supposons que l'algorithme WF (Équation (4.32)) alloue initialement la puissance. L'objectif de WF est de maximiser le débit total :

- L'utilisateur $i = 1$ a un très bon canal sur la sous-porteuse n_a . WF lui alloue une puissance p_{1,n_a} élevée.
- L'utilisateur $i = 2$ a un très bon canal sur la sous-porteuse n_b . WF lui alloue une puissance p_{2,n_b} .
- L'utilisateur $i = 3$ a un canal moyen sur la sous-porteuse n_a . WF lui alloue une puissance p_{3,n_a} .

Identification d'un Conflit : Un conflit se produit lorsque deux utilisateurs demandent la même ressource. Ici, les utilisateurs $\mathbf{i} = \mathbf{1}$ et $\mathbf{i} = \mathbf{3}$ utilisent simultanément la sous-porteuse n_a . L'ensemble des conflits \mathcal{A} inclut la paire $(i = 1, i = 3)$ sur n_a . Leur allocation initiale cause une forte interférence mutuelle, réduisant leur SINR et leur débit effectif.

Phase 2 : Résolution du Conflit par la Valeur de Shapley : Pour résoudre le conflit sur n_a , l'allocation de puissance doit être revue en fonction de la contribution marginale de chaque utilisateur à l'utilité globale de la coalition A_c .

1. Calcul de la Contribution (Valeur de Shapley, $\hat{\phi}_i(A_c)$) : L'approximation de Shapley (via Monte-Carlo, Équation (4.33)) est effectuée. Supposons les résultats suivants, reflétant la qualité de canal globale et l'impact de chaque joueur sur l'utilité collective :

- $\hat{\phi}_1(A_c) = 0.5$ (Joueur fort)
- $\hat{\phi}_2(A_c) = 0.4$
- $\hat{\phi}_3(A_c) = 0.1$ (Joueur faible)

Somme totale : $0.5 + 0.4 + 0.1 = 1.0$ (Respect de l'Efficacité).

2. Nouvelle Allocation de Puissance (Équation (4.38)) : La puissance P^{\max} est répartie sur la sous-porteuse n_a proportionnellement à la valeur de Shapley :

$$p_{1,n_a}^{\text{new}} = \frac{0.5}{1.0} \cdot P^{\max}$$

$$p_{3,n_a}^{\text{new}} = \frac{0.1}{1.0} \cdot P^{\max}$$

L'utilisateur $i = 1$ (fort) reçoit une allocation beaucoup plus importante que $i = 3$ (faible) sur la sous-porteuse en conflit, réduisant ainsi l'interférence subie par $i = 1$ et stabilisant le gain.

3. Négociation de Puissance (Étape 2 de l'Optimisation Intra-Coalition) : Pour éviter la domination totale du joueur fort ($i = 1$), le mécanisme de négociation incite

$i = 1$ (qui a une $\hat{\phi}$ élevée) à redistribuer une petite portion de sa puissance à $i = 3$. Ce petit ajustement garantit que $i = 3$ améliore son utilité ($IS_3 > 0$) et évite de quitter la coalition, renforçant ainsi la stabilité et l'équité de la coalition, même si l'allocation reste majoritairement guidée par la contribution (Shapley).

Phase 3 : Optimisation Dynamique par Q-Learning : Suite à l'ajustement de puissance, si l'utilisateur $i = 3$ a toujours un IS_3 faible ou négatif, il entre dans la phase Q-Learning.

- **État s :** $i = 3$ est dans A_c avec une faible IS.
- **Action a :** Rejoindre une autre coalition A'_c (exploration) ou rester dans A_c (exploitation).
- **Récompense r_3 :** Si rejoindre A'_c maximise la récompense $r_3 = \beta \cdot IS_3 + (1 - \beta) \cdot CS$, l'utilisateur $\mathbf{i} = \mathbf{3}$ change de coalition, conduisant à une nouvelle configuration globale plus stable et plus efficace.

Cet exemple illustre comment [U-CoCRL](#) combine l'équité (**Shapley**) pour l'allocation de puissance avec l'adaptabilité dynamique (**Q-Learning**) pour la formation des coalitions finales stables et efficaces.

La section suivante est dédiée à l'étude détaillée du cas où la [CSI](#) est imparfaite, en mettant en lumière la modélisation de cette incertitude ainsi que son impact sur la formation des coalitions. Un chapitre sera consacré aux simulations et aux résultats présentera une analyse approfondie des performances de la solution proposée, comparant les différents scénarios afin de démontrer la robustesse et l'efficacité de l'algorithme dans des conditions variées.

4.6 Allocation des ressources avec variation temporelle (Connaissance partielle de [CSI](#))

Dans les réseaux de communication sans fil, et plus particulièrement dans les environnements acoustiques sous-marins, la connaissance de l'état du canal [CSI](#) joue un rôle fondamental dans l'optimisation de l'allocation des ressources. Toutefois, en pratique, il est souvent difficile, d'obtenir une [CSI](#) complète et exacte, en raison des effets de propagation acoustique, de la mobilité des nœuds, du bruit ambiant, et des délais de retour d'information. En conséquence, dans de nombreux cas pratiques, seules des informations partielles ou imprécises sur l'état du canal sont disponibles. Ce type de connaissance est désigné sous le terme de **Partial CSI** (informations partielles sur l'état du canal), et nécessite le développement de stratégies de communication robustes capables de fonctionner efficacement même en présence d'incertitude sur le canal. Dans la partie suivante, nous proposons un cadre de modélisation adapté à ce contexte incertain, en nous appuyant sur la théorie des jeux coopératifs. Nous décrivons comment la partialité de la [CSI](#) est intégrée dans la formation des coalitions, et comment elle influence les interactions stratégiques entre utilisateurs et les performances globales du système.

4.6.1 Modélisation du jeu de coalition sous information partielle sur l'état du canal

Dans le même cadre d'un réseau de communication ASM présenté par la figure 4.1 dans la sous-section 4.4.3, nous considérons un ensemble de $K = 14$ et $K = 20$ utilisateurs évoluant dans un environnement caractérisé par une connaissance imparfaite de l'état du canal (CSI partielle) et soumis à des perturbations dynamiques, notamment l'effet Doppler.

Modélisation de la CSI partielle :

Dans un environnement sous-marin, l'acquisition précise de la CSI est particulièrement difficile à cause des caractéristiques du canal acoustique : forte atténuation, bruit ambiant élevé, propagation lente, et réflexions multiples. Ces conditions entraînent une estimation bruitée de l'état réel du canal.

Le canal entre deux utilisateurs i et j est donc modélisé par un gain $h_{i,j}$, qui n'est pas parfaitement connu. Seule une estimation partielle est disponible, notée $\hat{h}_{i,j}$, affectée par un bruit additif complexe gaussien :

$$\hat{h}_{i,j} = h_{i,j} + \eta_{i,j}, \quad \eta_{i,j} \sim \mathcal{CN}(0, \sigma^2), \quad (4.43)$$

où :

- $h_{i,j}$ est le gain (inconnu) du canal ;
- $\eta_{i,j}$ modélise l'erreur d'estimation ;
- $\hat{h}_{i,j}$ représente la CSI partielle.

Le paramètre σ^2 reflète le niveau d'incertitude dans l'estimation du canal. Plus sa valeur est élevée, plus la CSI est imprécise, ce qui rend la prise de décision coopérative plus difficile. Ce modèle permet de capturer de manière réaliste la nature imparfaite et fluctuante de la CSI dans un contexte sous-marin.

4.6.2 Modélisation de l'effet Doppler

À cette incertitude s'ajoute l'effet Doppler, causé par le mouvement relatif des joueurs (déplacements dus aux courants, ou à la navigation). Cet effet introduit une variation fréquentielle dans le canal, modélisée par le facteur :

$$\text{Dopp}(t) = \exp \left(j \cdot 2\pi f_{\text{vec}} \cdot \left(\frac{\text{cumsum}(v_{\text{td}} + v_{\text{rd}})}{c_w} \right) \cdot dt \right), \quad (4.44)$$

où :

- f_{vec} est le vecteur de fréquences porteuses utilisées ;
- v_{td} et v_{rd} sont les vitesses respectives du transmetteur et du récepteur ;
- c_w est la vitesse du son dans l'eau ;
- dt est le pas de temps utilisé dans la simulation.

Ce facteur traduit la variation temporelle de la phase du canal liée au déplacement des nœuds, ce qui contribue à sa dynamique non stationnaire.

Gain effectif du canal : En tenant compte à la fois de l'imprécision de la CSI et des effets Doppler, le gain effectif du canal utilisé dans les simulations est donné par :

$$\tilde{h}_{i,j}(t) = \hat{h}_{i,j}(t) \cdot \text{Dopp}_{i,j}(t). \quad (4.45)$$

Ce modèle global fournit une représentation réaliste du canal acoustique sous-marin, intégrant l'incertitude sur l'estimation et la variation temporelle. Il constitue la base de notre simulation pour évaluer la performance de notre stratégie proposée de formation de coalitions coopératives.

Impact de l'incertitude sur la formation des coalitions : Dans cette étude, nous utilisons un algorithme de formation de coalitions identique dans deux contextes distincts : (1) un scénario idéal dans lequel l'état du canal est parfaitement connu (CSI parfait), et (2) un scénario réaliste où seule une estimation bruitée du canal est disponible (CSI partielle). L'objectif de cette comparaison est de mesurer l'influence de l'incertitude sur les performances du système, tout en conservant une base algorithmique identique. Cette approche garantit une évaluation rigoureuse de la robustesse de l'algorithme face aux imperfections du canal.

Dans le cas d'une CSI parfaite, chaque utilisateur dispose d'une connaissance exacte des gains de canal, permettant une estimation fiable de sa propre utilité au sein d'une coalition donnée. Cette connaissance précise favorise des décisions de regroupement optimales, menant à la formation de coalitions efficaces sur le plan des performances globales.

À l'inverse, dans le scénario de CSI partielle, les gains de canal sont entachés d'erreurs d'estimation modélisées par un bruit gaussien complexe ainsi que par les distorsions introduites par l'effet Doppler, résultant de la mobilité relative entre les nœuds émetteurs et récepteurs. Ce contexte d'incertitude réduit la fiabilité de l'information disponible lors de la phase de décision, ce qui peut impacter la structure des coalitions formées et entraîner une baisse des performances.

4.7 Conclusion

Dans ce chapitre, nous avons exploré les fondements des jeux coopératifs par coalition appliqués à l'allocation de ressources dans les réseaux acoustiques sous-marins. Après avoir présenté les outils théoriques de formation de coalitions et les différents contextes d'application dans les réseaux sans fil, une attention particulière a été portée à l'intégration de mécanismes d'apprentissage adaptatif, notamment le Q-Learning, pour améliorer la stabilité et l'efficacité des coalitions formées.

La première partie du chapitre a abordé le cas où les informations d'état du canal sont parfaitement connues, permettant une formulation mathématique précise du problème et la mise en œuvre de stratégies optimales. Par la suite, le modèle a été étendu à des environnements plus réalistes où la CSI est partielle et soumise à des variations temporelles (effet Doppler, incertitude), afin de tester la robustesse du système. L'approche proposée, combinant jeux de coalition et apprentissage distribué, s'avère particulièrement bien adaptée aux réseaux sous-marins distribués, en raison de leur nature dynamique, bruitée et complexe. Elle permet de construire des stratégies coopératives robustes, efficaces, et

équitable dans des contextes à haute incertitude.

Le chapitre suivant est dédié à une analyse comparative approfondie, basée sur des simulations, entre notre solution coopérative par apprentissage et d'autres stratégies d'allocation, incluant les jeux non coopératifs et les approches classiques. Cette analyse portera sur plusieurs critères : utilité globale et individuelle, équité, robustesse face aux interférences, et complexité algorithmique.

Chapitre 5

Évaluation des contributions apportées au problème d'allocation de ressources en ASM : étude comparative

5.1 Introduction

Ce chapitre présente les résultats de simulation obtenus à partir de l'approche proposée basée sur les jeux coopératifs par coalition avec apprentissage distribué. L'objectif est de valider expérimentalement l'efficacité de cette méthode dans un contexte de communication [ASM](#), et de la comparer à d'autres approches existantes, notamment les jeux non coopératifs et certaines stratégies classiques d'allocation de ressources.

Les simulations ont été menées dans divers scénarios représentatifs du milieu sous-marin, avec des niveaux variables d'interférences, de dynamique du canal ([CSI](#) partielle), et de densité de utilisateurs. Plusieurs indicateurs de performance ont été analysés :

1. L'utilité globale du système (efficacité collective).
2. L'utilité individuelle (satisfaction des joueurs)
3. L'équité entre les joueurs.
4. La complexité algorithmique.

Une attention particulière est accordée à la comparaison avec les modèles non coopératifs, qui reposent sur des décisions décentralisées sans coordination. Bien que ces modèles puissent converger vers les équilibres de Nash, ils ne garantissent pas l'optimalité ou l'équité collective. À travers cette évaluation comparative, nous cherchons à démontrer que l'intégration de la coopération intelligente, guidée par l'apprentissage, permet d'obtenir des performances significativement plus élevées, tout en assurant une allocation plus stable et équitable des ressources dans les environnements acoustiques sous-marins complexes.

5.2 Résultats de simulation et analyse

Afin d'évaluer les performances de l'algorithme proposé de formation de coalitions pour l'allocation conjointe de la bande passante et de la puissance dans un environnement acoustique sous-marin, nous avons mis en œuvre une simulation complète dans l'environnement MATLAB. Cette simulation repose sur le modèle de canal acoustique décrit dans [Qarabaqi and Stojanovic, 2013], reproduisant des conditions réalistes de communication sous-marine, incluant la limitation de bande passante, le fading du canal, ainsi que les interférences inter-utilisateurs. Dans ce cas, le réseau étudié est constitué initialement de 14 paires émetteur-récepteur disposées aléatoirement dans une zone circulaire de rayon 3000 mètres et à une profondeur fixe de 100 mètres. Tous les utilisateurs partagent la même bande de fréquence, ce qui introduit des interférences potentielles dues à la nature non orthogonale de la transmission. Ce contexte reflète la compétition pour les ressources limitées dans un réseau acoustique dense.

L'objectif principal est d'évaluer la capacité de l'algorithme de formation de coalitions à optimiser l'allocation des ressources (puissance et sous-porteuses), tout en limitant les conflits d'usage et en maximisant la satisfaction globale des utilisateurs au sein des coalitions formées sous la conditions de forte interférence.

L'interférence reçue par l'utilisateur i sur la sous-porteuse n est modélisée comme suit :

$$I_i(n) = \sum_{j \neq i} \xi \cdot |h_{ji}(n)|^2 p_{j,n}, \quad (5.1)$$

où $h_{ji}(n)$ est le gain de canal entre l'utilisateur j et l'utilisateur i sur la sous-porteuse n , $p_{j,n}$ est la puissance allouée par j sur n , et ξ est un facteur d'échelle introduit pour simuler différents niveaux d'interférences.

D'autre part, pour justifier le choix du nombre optimal de coalitions, nous avons utilisé K-means, guidée par le critère du coude [Cui, 2020] en analysant les valeurs d'inertie dérivées de différentes valeurs de K. Notre simulation montre que la coalition $k = 4$ est le choix optimal. À ce stade, l'inertie diminue considérablement au-delà de ce nombre, ce qui implique que l'ajout de plus de coalitions présente un bénéfice minimal en termes d'amélioration des performances ou de la cohérence. Cela renforce l'hypothèse selon laquelle le nombre de coalitions est un facteur important dans l'amélioration des performances du réseau.

La simulation a été effectuée en exécutant plusieurs itérations Monte-Carlo, afin de garantir la convergence statistique des résultats. Les paramètres retenus pour cette évaluation sont résumés dans le tableau 5.1.

Le paramètre N_C^{\max} (Taille maximale d'une coalition) a été fixé à 4. Ce choix représente un compromis crucial entre la performance et la complexité. En effet, la complexité de l'étape de négociation, notamment le calcul de la valeur de Shapley qui est exponentielle par rapport au nombre de membres. Limiter la taille à 4 permet de maintenir cette complexité à un niveau gérable pour une solution adaptative, tout en offrant des gains coopératifs significatifs.

TABLE 5.1 – Paramètres de simulation.

Désignation	Symbole	Valeur
Nombre d'utilisateurs	K	14
Nombre de sous-porteuses	N_{sc}	128
Bande passante totale	B	14 kHz
Bande de fréquences	f_{span}	16 kHz – 30 kHz
Fréquence centrale	f_0	23 kHz
Espacement entre sous-porteuses	Δf	109.37 Hz
Préfixe cyclique	T_g	9 ms
Puissance maximale par utilisateur	P^{max}	2 W
Rapport signal/bruit	SNR	15 dB
Itérations Monte-Carlo	M	100
Itérations d'apprentissage (QL)	T	50
Nombre de coalitions formées	N_c	4
Taille maximale d'une coalition	N_C^{max}	4
Taux d'apprentissage	η	0.1
Facteur d'actualisation	γ	0.9
Taux d'exploration	ϵ	0.1
Facteur d'interférence	ξ	{0.01, 0.1, 1}

Après avoir défini les paramètres de simulation et les scénarios proposés, nous présentons dans la section suivante une analyse détaillée des résultats obtenus.

5.2.1 Analyse des résultats : Cas de CSI parfait

Dans cette section, nous analysons l'évolution des fonctions d'utilités en utilisant notre solution par rapport aux méthodes mentionnées précédemment (la stratégie RECORD et jeux Hédoniques). Nous évaluons également l'utilité globale du système et l'équité en fonction de l'indice d'équité de Jain [Khan et al., 2016]. De plus, une autre métrique est évaluée à l'aide de la mesure de performance PI, qui est définie par l'équation 5.3, ce qui nous permet d'examiner les compromis entre efficacité et équité dans des scénarios multi-utilisateurs.

Afin d'évaluer les performances des algorithmes de formation de coalitions présentés précédemment, plusieurs indicateurs quantitatifs ont été utilisés. Ces métriques permettent non seulement d'estimer l'efficacité globale de l'algorithme, mais aussi d'évaluer l'équité dans la répartition des ressources entre les utilisateurs.

Indice d'équité (Fairness Index) : L'indice d'équité mesure à quel point les utilités individuelles u_i sont réparties de manière équilibrée entre les utilisateurs. Une valeur élevée de cet indice indique une distribution plus équitable. L'un des indices d'équité les plus utilisés est l'indice de Jain, défini comme suit [Khan et al., 2016] :

$$F = \frac{\left(\sum_{i=1}^K u_i\right)^2}{K \cdot \sum_{i=1}^K u_i^2}. \quad (5.2)$$

Cet indice varie entre $\frac{1}{K}$ (distribution totalement inéquitable) et 1 (distribution parfaitement équitable), où K est le nombre total d'utilisateurs. Il permet donc d'évaluer si certaines coalitions ou utilisateurs bénéficient de manière disproportionnée de l'allocation des ressources.

Indice de performance global PI : C'est une mesure qui intègre à la fois l'efficacité (via l'utilité moyenne) et l'équité (via la variance). Il est défini par l'expression suivante :

$$PI = \frac{\mathbb{E}[u_i]}{1 + \text{Var}(u_i)}. \quad (5.3)$$

Dans le cadre de cette étude, le **PI** joue un rôle central dans l'évaluation des algorithmes. Il permet de capturer le compromis entre la performance globale du système (efficacité moyenne mesurée par $\mathbb{E}[u_i]$) et la régularité de cette performance entre les utilisateurs (équité, représentée par la variance $\text{Var}(u_i)$). Ainsi, un **PI** élevé indique que le système parvient à fournir une utilité moyenne importante tout en maintenant une répartition relativement homogène entre les utilisateurs. À l'inverse, une forte variabilité dans les utilités individuelles, même si la moyenne est élevée, se traduit par un **PI** plus faible, puisque la métrique est intrinsèquement conçue pour favoriser les politiques équitables (à faible variance).

Ces deux métriques complémentaires sont essentielles pour comparer différentes stratégies de formation de coalitions, notamment en présence d'incertitudes (comme une **CSI** partielle), où l'équilibre entre performance et équité devient plus difficile à atteindre.

5.2.2 Étude de la fonction d'utilité globale du système

Nous commençons tout d'abord par étudier l'évolution de la fonction d'utilité globale du système par la figure 5.1. La méthode **U-CoCRL** proposée, qui combine Q-Learning avec la théorie des jeux de coalition, surpasse en moyenne les autres méthodes, atteignant une utilité globale de 9.8 bits/s/Hz, surpassant notablement l'algorithme hédonique. Cette valeur correspond à la moyenne obtenue sur plusieurs simulations Monte Carlo. Cet avantage de performance est dû à sa capacité à ajuster dynamiquement les structures de coalition et les allocations de puissance en temps réel via **QL**, permettant une adaptation continue aux conditions variables du réseau, une caractéristique cruciale dans des environnements dynamiques comme les réseaux de communication sous-marins.

En revanche, la méthode hédonique, bien qu'efficace pour optimiser la satisfaction des utilisateurs individuels, est moins performante dans des conditions de forte interférence. Elle privilégie les préférences des utilisateurs locaux par rapport à l'utilité globale du système, ce qui conduit à une distribution sous-optimale des ressources.

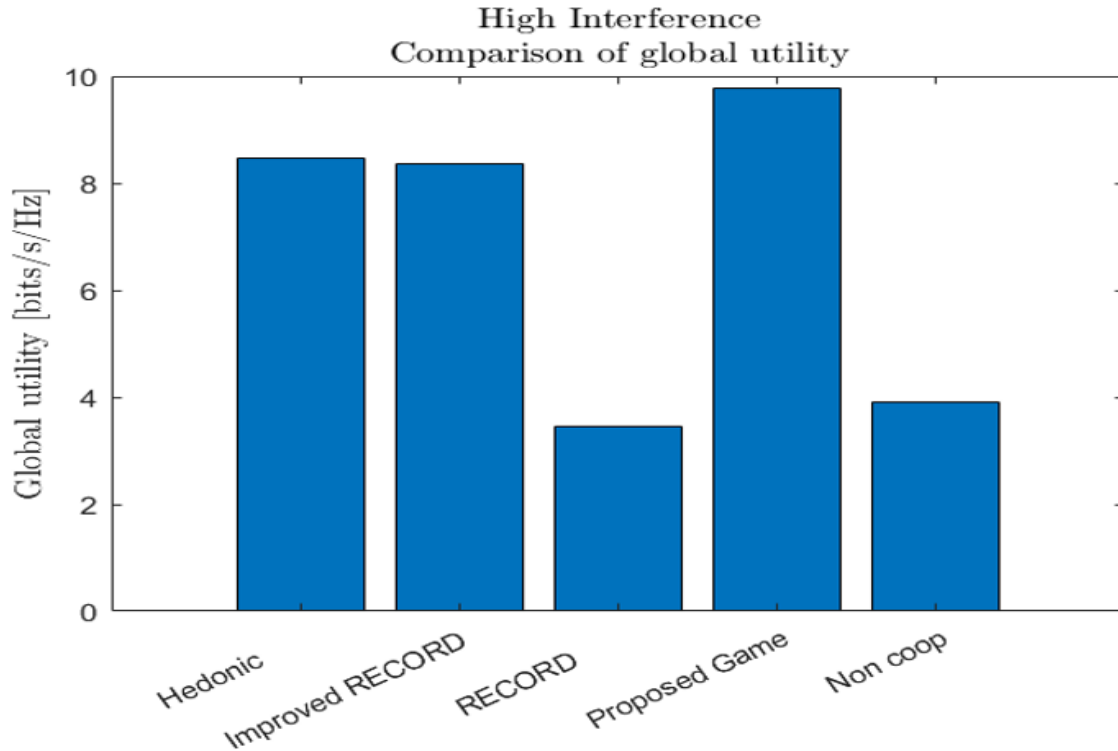


FIGURE 5.1 – Évolution de la fonction d'utilité globale selon les différentes approches.

De même, l'algorithme RECORD, bien qu'il soit conçu pour former des coalitions stables, souffre d'une grande complexité computationnelle en raison de sa recherche exhaustive. Contrairement à notre approche [U-CoCRL](#) qui utilise le Q-Learning pour converger vers une solution efficace sans nécessiter une telle recherche, RECORD est inadapté aux environnements dynamiques. Cette nature statique explique sa faible performance (3.5 bits/s/Hz) dans notre scénario aux conditions changeantes. Pour évaluer la différence de performance, nous avons intégré la combinaison RECORD avec la valeur de Shapley (version RECORD amélioré). Malgré cette amélioration, son utilité globale reste nettement inférieure à celle de la méthode proposée.

Concernant l'algorithme non coopératif, qui ne permet pas la formation de coalition, il atteint les performances les plus faibles avec une utilité globale de seulement 3.9 (bits/s/Hz). Cela s'explique par les interférences mutuelles massives que les utilisateurs génèrent en choisissant tous les meilleures ressources de manière égoïste. Cela met en évidence l'inefficacité des approches non coopératives dans la gestion des interférences. Sans coopération ou échange d'informations entre les utilisateurs, ces algorithmes ne parviennent pas à atténuer efficacement les interférences, ce qui entraîne une gestion des ressources sous-optimale.

Ces résultats soulignent l'importance des stratégies coopératives, telles que la méthode proposée, qui s'adaptent dynamiquement aux interférences et assurent une allocation optimale des ressources dans les environnements à fortes interférences.

Bien que l'utilité globale mesure l'efficacité globale, la distribution des utilités individuelles reflète l'efficacité du réseau à équilibrer les performances collectives et la satisfaction individuelle des utilisateurs ce qui est illustrée dans la partie suivante.

5.2.3 Étude de la fonction d'utilité individuelle

Il est essentiel d'étudier comment diverses stratégies affectent l'utilité individuelle (figure 5.2). L'analyse de cette figure met en évidence le compromis fondamental entre l'optimisation individuelle égoïste et l'équité globale.

On observe que les méthodes alternatives (Hedonic, Improved RECORD) génèrent des pics d'utilité extrêmes pour des utilisateurs spécifiques (par exemple, l'utilisateur 13), mais cela se fait au détriment de sacrifier d'autres utilisateurs (comme les 10 et 11) qui n'obtiennent presque rien. Cette variation illustre la difficulté à concilier l'utilité globale avec la satisfaction individuelle.

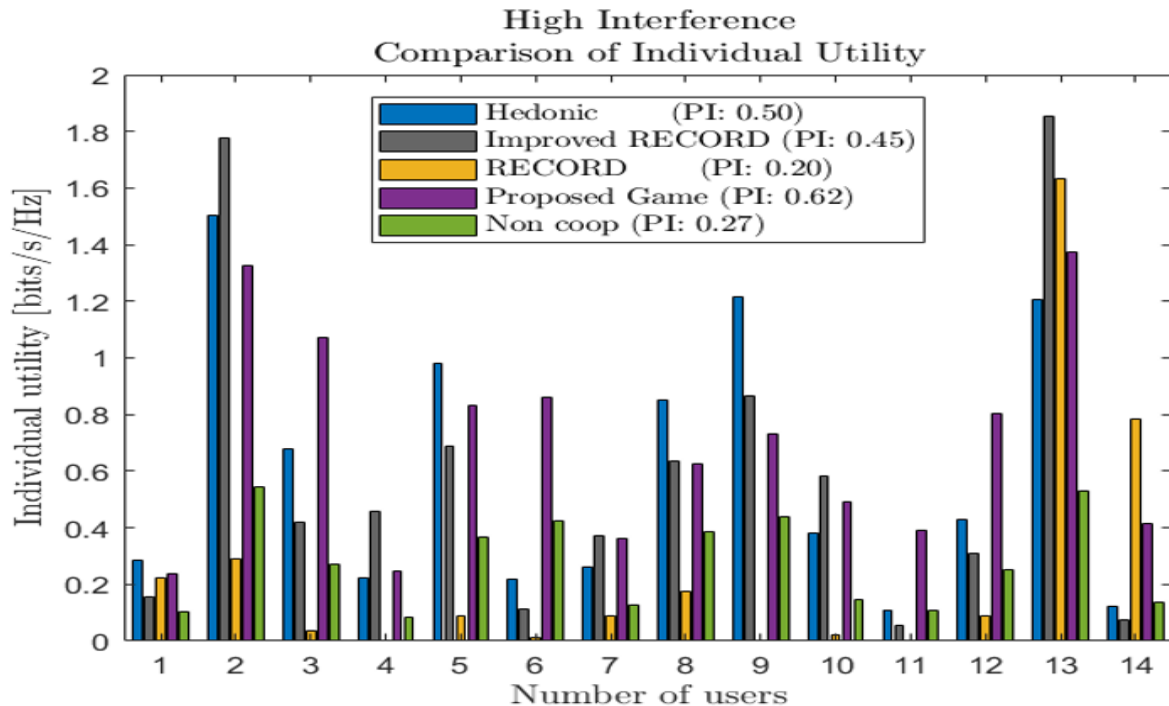


FIGURE 5.2 – Évolution de la fonction d'utilité individuelle selon les différentes approches.

L'algorithme [U-CoCRL](#) (en violet) se distingue par sa capacité d'adaptation en temps réel. Cette approche ne cherche pas seulement à maximiser l'utilité globale (vue en Fig. 5.1), mais aussi à garantir une meilleure équité. Visuellement, bien que [U-CoCRL](#) n'atteigne pas les pics extrêmes de l'utilisateur 13, ses barres sont distribuées de manière bien plus homogène et équitable sur l'ensemble des utilisateurs. Cette supériorité en termes de compromis est prouvée quantitativement par l'Indice de [PI](#) de notre solution (0.62), le plus élevé de tous. Cette adaptabilité s'avère précieuse dans les environnements à forte interférence, où les conditions de communication évoluent rapidement.

En revanche, les approches hédoniques [[Wu et al., 2015](#)] illustrent parfaitement ce manque d'équité. Bien qu'elles permettent à certains utilisateurs (comme le 13) d'obtenir une utilité très élevée, elles pénalisent fortement d'autres utilisateurs, ce qui résulte en un [PI](#) global de seulement 0.50.

De son côté, l'algorithme [RECORD](#) amélioré ([PI](#) de 0.45) présente le même défaut, peinant à concilier utilité et équité, comme le montre le pic extrême pour l'utilisateur 13. Les algorithmes [RECORD](#) classiques [[Ma et al., 2013](#)] et non-coopératifs affichent les indices de performance les plus faibles (0.20 et 0.27), confirmant leur incapacité à gérer les ressources de manière équitable dans ces environnements.

Dans l'ensemble, tandis que les autres méthodes favorisent injustement quelques utilisateurs au détriment de la majorité, l'approche [U-CoCRL](#) offre le meilleur compromis entre l'utilité individuelle et l'équité globale, ce qui est validé par son score [PI](#) supérieur.

5.2.4 Analyse d'équité

L'analyse des indices d'équité, présentée dans le [Tableau 5.2](#), met en évidence des différences significatives entre les algorithmes en termes de répartition des utilisateurs au sein des coalitions. Ce tableau est donné comme exemple illustratif issu de la dernière itération d'une série de 100 simulations de type Monte Carlo. l'indice d'équité moyen est calculé en utilisant l'Indice d'Équité de Jain (Jain's Fairness Index). Cet indice est appliqué à la taille des coalitions formées lors de chaque itération. La valeur présentée (ex : 0.9189) est la moyenne de cet indice sur l'ensemble des 100 simulations Monte Carlo.

TABLE 5.2 – Comparaison de l'indice d'équité moyen : approche [U-CoCRL](#) vs méthodes classiques.

Algorithme	Coalition 1	Coalition 2	Coalition 3	Coalition 4	Indice d'équité moyen
Jeu proposé U-CoCRL	{14, 10, 7, 4}	{13, 2, 5, 1}	{6, 8, 9, 11}	{3, 12}	0.9189
Jeu Hédonique [Wu et al., 2015]	{4, 7, 12, 14, 11, 10, 8}	{2, 3}	{5, 9, 13}	{1, 6}	0.8439
RECORD amélioré	{4}	{2, 3, 5, 6, 8, 9, 11, 12, 14, 1}	{7, 13}	{10}	0.8049
Jeu RECORD [Ma et al., 2013]	{1, 4}	{2, 3, 5, 6, 8, 9, 11, 12}	{7}	{10, 13, 14}	0.7084

Le jeu proposé [U-CoCRL](#) forme des coalitions structurellement plus équilibrées. Il atteint ainsi l'indice de Jain moyen le plus élevé (0.9189), ce qui reflète une meilleure homogénéité en évitant les déséquilibres extrêmes. Par exemple, la répartition de l'exemple [4, 4, 4, 2] est bien plus homogène que les répartitions [1, 10, 2, 1] de RECORD amélioré ou [7, 2, 3, 2] du Jeu Hédonique, qui incluent des coalitions singleton ou des coalitions géantes.

Concernant la coalition 3, 12, sa formation n'est pas une anomalie mais une démonstration de l'adaptabilité de l'algorithme. Elle illustre une configuration ciblée où ces deux utilisateurs, en raison de l'état de leur canal, maximisent leur utilité en s'isolant plutôt qu'en rejoignant un groupe plus large où ils subiraient plus d'interférences. Les autres algorithmes, comme le Jeu Hédonique (formant 1, 6), créent aussi de petites coalitions, mais sans atteindre le même niveau d'équité globale, comme le prouve leur indice de Jain moyen inférieur.

En revanche, les méthodes alternatives telles que le jeu Hédonique [[Wu et al., 2015](#)] ou RECORD [[Ma et al., 2013](#)] et RECORD amélioré présentent des répartitions plus déséquilibrées. Leurs indices d'équité moyens (entre 0.70 et 0.84) confirment cette instabilité structurelle. Une autre analyse de l'équité : celle des utilités individuelles est représentée

par la Figure 5.3, qui illustre la comparaison de l'équité dans la distribution de l'utilité reçue par chaque utilisateur selon différentes approches, dans un environnement à forte interférence.

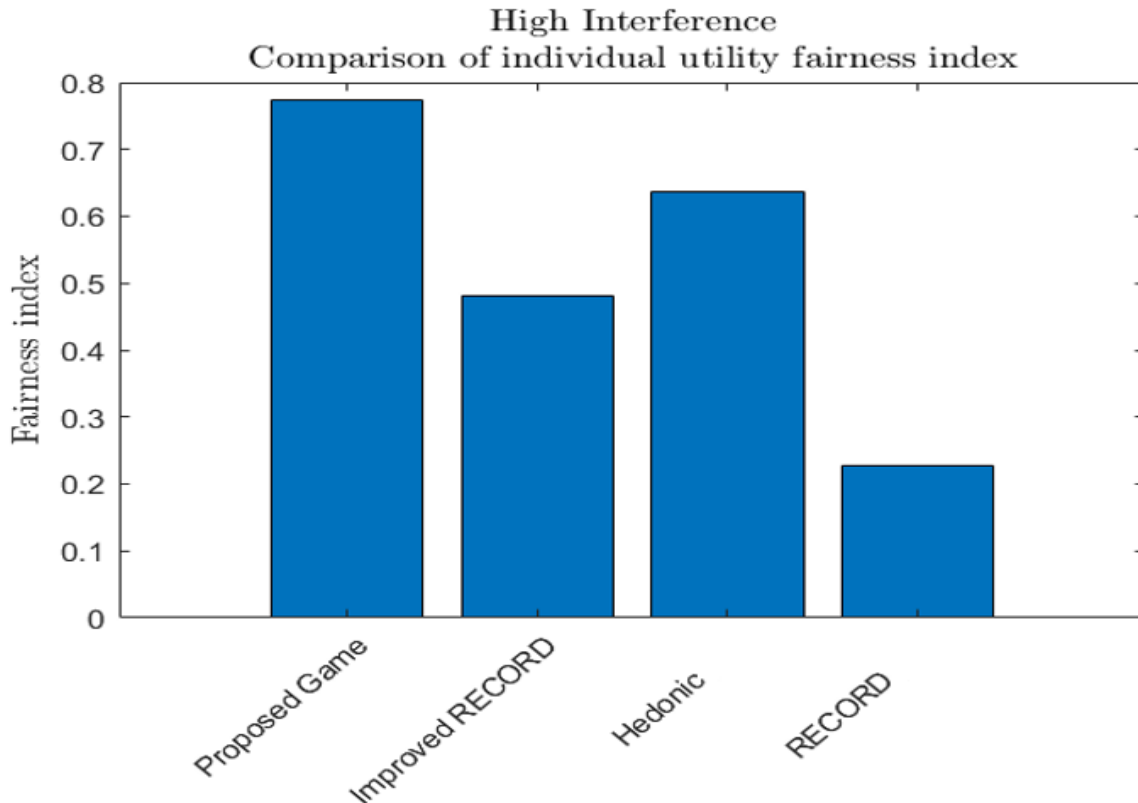


FIGURE 5.3 – Comparaison de l'indice d'équité des utilités individuelles.

L'approche [U-CoCRL](#) affiche le meilleur score, avec un indice proche de 0.78, soulignant sa capacité à garantir une répartition équitable des bénéfices entre utilisateurs. Cela confirme que sa structure de coalition plus stable (vue au [Tableau 5.2](#)) se traduit par une répartition plus équitable des gains. À l'inverse, la méthode RECORD améliorée et RECORD classique présentent une forte inégalité (indices autour de 0.48 et 0.23 respectivement), ce qui témoigne d'une satisfaction utilisateur plus hétérogène. Le jeu Hédonique, bien qu'axé sur la satisfaction individuelle, n'atteint qu'un équilibre partiel (indice d'environ 0.63), suggérant une répartition inégale dans les gains.

Ces résultats confirment qu'une approche intégrant à la fois l'équité structurelle des coalitions et l'équité perçue des utilités individuelles est essentielle pour renforcer la stabilité, limiter les conflits d'intérêt et améliorer la performance globale du système.

5.2.5 Analyse de satisfaction

Il est essentiel d'analyser à la fois la stabilité des coalitions formées et la satisfaction des utilisateurs. Une coalition stable témoigne de la capacité des utilisateurs à y rester sans volonté de réorganisation, ce qui est crucial dans des environnements dynamiques. La satisfaction reflète l'adéquation entre les attentes des utilisateurs et les utilités qu'ils obtiennent au sein de leur coalition. Les [Figures 5.4](#) et [5.5](#) mettent en évidence ces deux

dimensions dans un contexte de forte interférence.

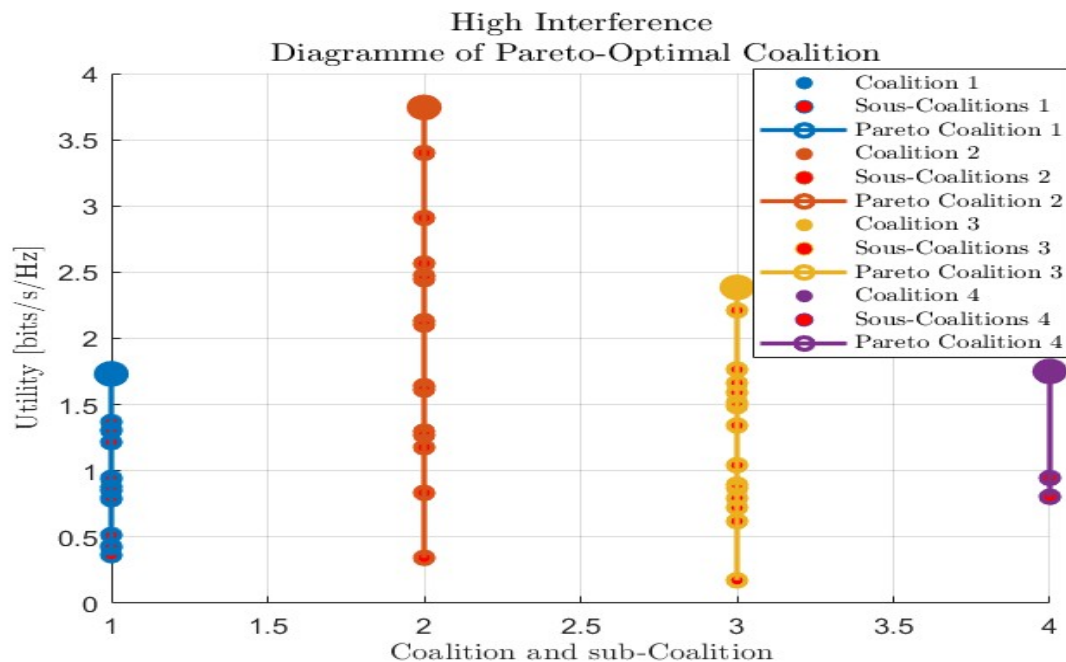


FIGURE 5.4 – Histogramme des coalitions Pareto-optimales.

La Figure 5.4 montre que toutes les coalitions formées atteignent un état Pareto-optimal, c'est-à-dire qu'aucun utilisateur ne peut améliorer son utilité sans nuire à celle d'un autre. Toutefois, elle révèle aussi des disparités, notamment la grande distribution (forte variance) des utilités au sein de la Coalition 2, où les utilités s'étalent de 0.5 à 3.8. Cela reflète une organisation efficace mais qui impose des compromis sévères : l'algorithme [U-CoCRL](#) a trouvé un équilibre stable où, en raison des fortes interférences, les utilités individuelles doivent varier pour préserver l'équilibre collectif.

La Figure 5.5 complète cette analyse en illustrant le niveau de satisfaction des utilisateurs.

Plutôt qu'un échec, cela démontre la capacité de l'algorithme à gérer le conflit inévitable entre l'utilité individuelle et la performance globale. La majorité d'entre eux sont satisfaits (marqueurs verts), ce qui illustre la pertinence de l'assignation. Quelques utilisateurs, notamment dans les coalitions 2 et 3, expriment toutefois une insatisfaction (marqueurs rouges). Ces utilisateurs sont insatisfaits car leur utilité individuelle (proche de 0.5) est faible. Il s'agit du compromis nécessaire qu'ils doivent accepter pour que l'ensemble de la coalition reste Pareto-optimal dans cet environnement à forte interférence. Malgré cela, le taux de satisfaction reste élevé, confirmant la capacité de l'algorithme à former des coalitions à la fois efficaces, justes et adaptées aux préférences des utilisateurs.

Enfin, bien que ces résultats soient performants en situation de forte interférence, une analyse plus large sous différents niveaux d'interférence permettrait de consolider l'évaluation de la robustesse et de l'adaptabilité de l'algorithme.

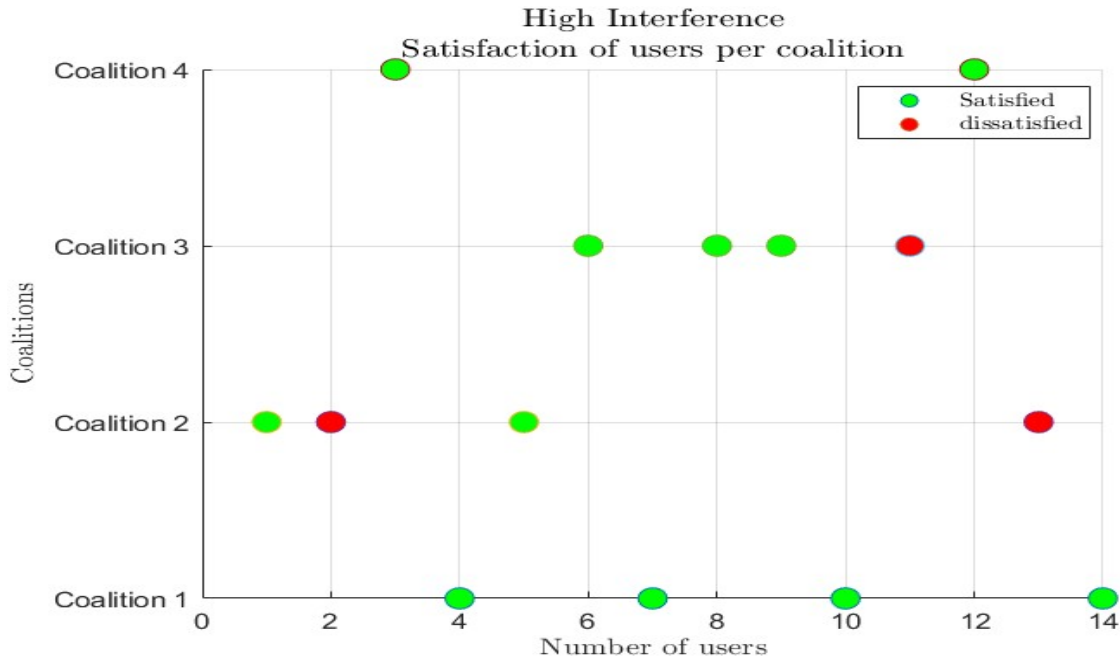


FIGURE 5.5 – Satisfaction des utilisateurs au sein des coalitions.

5.2.6 Comparaison des algorithmes selon différents niveaux d’interférence

Afin d’évaluer les performances des algorithmes, le Tableau 5.3 présente une comparaison directe entre l’algorithme proposé **U-CoCRL** et les méthodes de référence, en fonction de différents niveaux d’interférence.

TABLE 5.3 – Comparaison de l’utilité globale (bits/s/Hz) et de l’indice de performance selon différents niveaux d’interférence.

Algorithme	Interférence faible		Interférence moyenne		Interférence élevée	
	Utilité	PI	Utilité	PI	Utilité	PI
U-CoCRL	14.0	0.83	12.5	0.79	9.8	0.62
RECORD amélioré	13.2	0.65	11.0	0.59	8.0	0.45
Hédonique [Wu et al., 2015]	11.8	0.62	10.9	0.60	8.3	0.50
RECORD [Ma et al., 2013]	5.1	0.31	4.6	0.27	3.6	0.20
Jeu Non-coopératif	12.9	0.79	9.6	0.65	3.9	0.27

Cette comparaison met en évidence la valeur croissante de la coopération à mesure que l’environnement se dégrade.

Interférence faible : Dans ce scénario, l'impact des interférences est minime. La coopération est moins cruciale, comme le montre l'écart de performance serré : l'approche non-coopérative (12.9 bits/s/Hz) est performante et atteint presque notre solution (14.0 bits/s/Hz). La coopération n'offre qu'un gain de 8.5%.

Dans ce contexte peu contraint, la formation de coalitions ou la coopération entre utilisateurs est moins cruciale, car chacun peut allouer les ressources de manière autonome et efficace. Les algorithmes non coopératifs offrent des performances satisfaisantes, ce qui indique que l'apport de la coopération dans ces conditions reste limité.

Interférence moyenne : Lorsque l'interférence devient modérée, les approches non-coopératives commencent à souffrir. L'utilité du Jeu Non-coopératif chute à 9.6 bits/s/Hz et son **PI** se dégrade (de 0.79 à 0.65) aussi que l'algorithme RECORD enregistre une chute significative à 4.6 bits/s/Hz. L'écart se creuse : notre approche **U-CoCRL** (12.5 bits/s/Hz) est désormais 30% plus performante.

dans un environnement d'interférence moyenne, les approches coopératives comme **U-CoCRL** ou le jeu Hédonique se révèlent plus aptes à s'adapter. Grâce à la mise en place de coalitions et au partage des ressources, elles parviennent à mieux gérer les interférences et à préserver des performances élevées, surpassant largement les modèles non coopératifs.

Interférence élevée : C'est dans ce scénario réaliste et fortement contraint que la coopération devient essentielle. Sous de fortes interférences, les effets négatifs dus aux collisions de signaux entre utilisateurs deviennent plus importants, ce qui compromet la qualité de la communication si aucune stratégie d'adaptation n'est mise en œuvre. Dans ce contexte, la coopération et la formation de coalitions deviennent essentielles afin d'optimiser la répartition des ressources et de limiter les interférences croisées.

- L'algorithme proposé atteint une utilité globale de 9.8 bits/s/Hz, démontrant sa capacité à réduire efficacement l'impact des interférences. D'autres algorithmes coopératifs, comme le jeu Hédonique, obtiennent 8.3 bits/s/Hz, un score correct mais inférieur à leur performance dans un environnement moins perturbé.
- L'indice de performance de **U-CoCRL** reste le plus robuste avec une valeur de 0.62, ce qui témoigne de sa stabilité et de sa résilience face à des conditions extrêmes. À l'inverse, les modèles non coopératifs subissent une nette dégradation : l'approche non coopérative descend à 0.27, et l'algorithme RECORD atteint un seuil critique de 0.20, illustrant leur incapacité à gérer efficacement les interférences sans mécanismes de coordination.

Ces résultats démontrent que la coopération n'est pas un simple plus : elle est la clé de la robustesse. En situation de forte interférence, **U-CoCRL** est 151% plus performant que l'approche non-coopérative, prouvant la supériorité d'une gestion intelligente et coordonnée des ressources.

Afin de mieux comprendre l'importance de chaque composant de notre méthode, nous présentons dans la section suivante une analyse d'ablation (Tableau 5.4) qui évalue l'impact individuel du K-Means et du Q-learning sur les performances globales.

5.2.7 Scénario sans K-Means vs Scénario sans Q-learning

Analyse du scénario sans K-Means et avec Q-learning : Dans ce scénario, les coalitions sont formées selon un ordre fixe, sans application de l’algorithme K-Means. Cependant, les utilisateurs utilisent le Q-learning pour adapter dynamiquement leurs décisions en fonction de leur satisfaction locale et globale, cette approche permet au système d’atteindre une utilité globale de 4,9 bits/s/Hz, tout en maintenant un indice de fairness élevé de 0,71. Ces résultats montrent que, malgré l’absence de regroupement initial optimisé, le Q-learning permet de réagir efficacement aux dynamiques du réseau et de réduire les déséquilibres entre utilisateurs. La fonction de récompense utilisée, qui combine à la fois la satisfaction individuelle et collective, favorise un comportement coopératif équilibré, capable d’optimiser la répartition des ressources sans sacrifier l’équité.

Afin de compléter l’analyse, nous considérons maintenant le scénario intégrant K-Means pour la formation des coalitions mais sans Q-learning. L’objectif est d’évaluer l’efficacité de la structuration initiale en l’absence d’apprentissage adaptatif.

Analyse du scénario avec K-Means et sans Q-learning : Dans ce second scénario, les coalitions sont formées via K-Means, ce qui assure un regroupement initial pertinent basé sur la similarité des utilisateurs. Cette structuration initiale permet au système d’atteindre une utilité globale légèrement supérieure de 5,4 bits/s/Hz, mettant en évidence l’efficacité du clustering pour améliorer la coordination locale entre utilisateurs. Cependant, en l’absence de Q-learning, aucune adaptation dynamique n’est possible après la formation des groupes. Le système reste donc rigide face aux changements de l’environnement, et les déséquilibres initiaux ne peuvent être corrigés. Cela se reflète dans une fairness plus faible (0,655), signalant une répartition inégale des ressources. Certains utilisateurs bénéficient de leur position initiale favorable dans le clustering, tandis que d’autres restent durablement pénalisés, faute de mécanisme adaptatif.

À l’inverse, dans les scénarios intégrant le Q-learning, même en l’absence de K-Means, les agents sont capables d’ajuster leurs décisions au fil du temps pour mieux gérer les conflits, compenser les désavantages structurels, et améliorer l’équilibre général. Cela s’explique par la conception de la fonction de récompense, qui intègre à la fois la satisfaction individuelle et la satisfaction globale du système. Grâce à ce double objectif, l’algorithme oriente les joueurs vers des décisions coopératives et équilibrées, évitant les comportements égoïstes. Ainsi, le Q-learning agit comme un mécanisme de régulation dynamique, capable de redistribuer intelligemment les ressources selon l’évolution des besoins et des conditions du réseau, et assurant ainsi un compromis efficace entre performance et équité.

Le tableau 5.4 ci-dessous présente une synthèse comparative des résultats obtenus pour les différents scénarios, mettant en évidence l’impact respectif du regroupement par K-Means et de l’apprentissage via Q-learning.

TABLE 5.4 – Résumé des performances selon les scénarios d’ablation

Scénario	K-Means	Q-learning	Utilité globale (bits/s/Hz)	Fairness
Sans K-Means, avec QL	Non	Oui	4.9	0.71
Avec K-Means, sans QL	Oui	Non	5.4	0.655
Avec K-Means et QL (U-CoCRL)	Oui	Oui	9.8	0.78

Les résultats présentés dans le tableau 5.4 mettent clairement en évidence la complémentarité entre le regroupement initial par K-Means et l’adaptation dynamique assurée par le Q-learning. Chaque approche apporte un gain : K-Means permet une meilleure structuration des coalitions et améliore l’utilité globale, tandis que le Q-learning favorise une répartition plus équitable des ressources grâce à sa capacité d’adaptation en temps réel. Cependant, c’est leur combinaison qui conduit aux meilleures performances globales, en atteignant à la fois une utilité élevée (9.8 bits/s/Hz) et une fairness renforcée (0.78). Ces résultats soulignent l’importance d’un équilibre entre structuration initiale et intelligence adaptative dans les environnements dynamiques et incertains. L’approche complète permet non seulement d’optimiser l’efficacité du système, mais aussi d’assurer une allocation équitable et durable des ressources entre les utilisateurs. Cette allocation, nous nous référons à l’analyse de la stabilité Pareto-optimale (Figure 5.4). C’est cette figure qui représente visuellement la durabilité de nos coalitions. En effet, elle montre que les coalitions finales formées sont stables (Pareto-optimales), signifiant qu’aucun utilisateur n’a d’incitation rationnelle à quitter sa coalition. C’est cette convergence vers un état stable, atteinte grâce à l’apprentissage du Q-Learning et validée statistiquement sur 100 simulations Monte-Carlo, qui nous permet de qualifier cette allocation de durable.

Dans la partie suivante, nous présentons et analysons les résultats obtenus dans un contexte plus réaliste, c’est-à-dire avec une CSI partielle, mais également lorsque le nombre d’utilisateurs dans le réseau est plus important. L’objectif est de mettre en évidence l’impact de ces conditions pratiques sur l’allocation des ressources et sur les performances globales du système.

5.2.8 Analyse de la robustesse avec CSI partielle et augmentation du nombre d’utilisateurs

Dans les sections précédentes, notre analyse reposait sur une connaissance parfaite du canal (CSI Parfaite), où notre solution (pour 14 utilisateurs) atteignait une utilité globale de 9.8 bits/s/Hz et un PI de 0.62.

Nous passons maintenant à un contexte plus réaliste, en introduisant l’hypothèse de la CSI partielle et en évaluant la scalabilité du réseau.

1. Impact de l’incertitude du canal (CSI Parfaite vs Partielle)

Pour évaluer la robustesse, il faut d’abord mesurer le coût de l’incertitude. En comparant les performances pour 14 utilisateurs, l’introduction de la CSI partielle a un impact

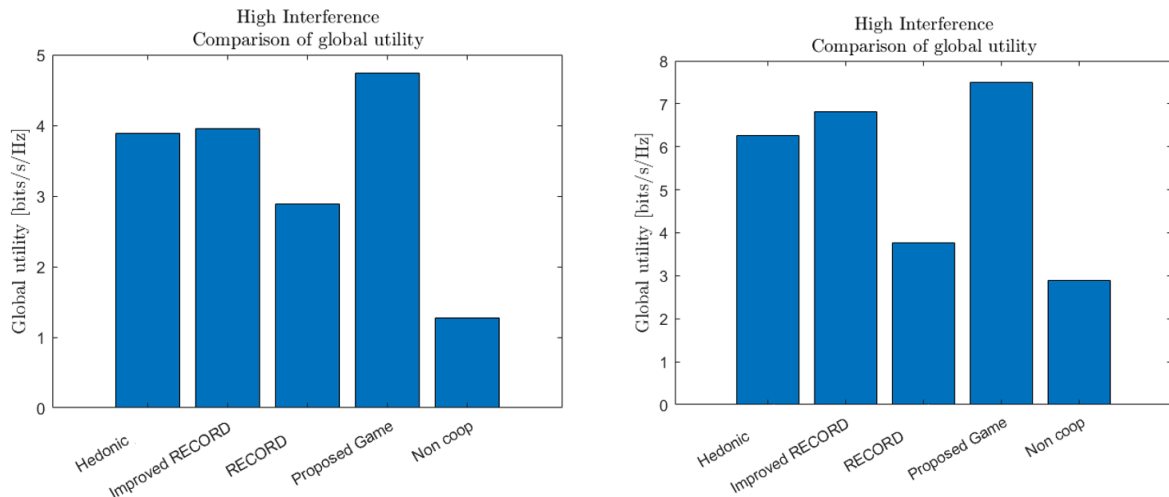
significatif :

- L'utilité globale de **U-CoCRL** chute de 9.8 (**CSI** Parfaite) à 4.7 bits/s/Hz (**CSI** Partielle, voir Figure 5.6a). Cette chute de 52% n'est pas un échec de l'algorithme ; c'est le coût de l'incertitude, car l'agent Q-Learning prend désormais des décisions basées sur des informations imparfaites.
- le **PI** de **U-CoCRL** chute de 0.62 à 0.32.
- Les autres algorithmes comparés souffrent d'une dégradation similaire de leur **PI** en passant en **CSI** partielle (par exemple, Hédonique passe de 0.50 à 0.26).

Ce qui est crucial, c'est que malgré cette chute, l'algorithme **U-CoCRL** (4.7 bits/s/Hz) reste 19% supérieur à son concurrent le plus proche (RECORD amélioré, 3.94) dans ce scénario dégradé, prouvant sa robustesse relative.

2. Analyse de la scalabilité (14 vs 20 utilisateurs) sous **CSI** Partielle

Nous évaluons maintenant le comportement de l'algorithme sous **CSI** partielle lorsque le nombre d'utilisateurs augmente de 14 à 20 (Figure 5.6).



(a) Évaluation de l'utilité globale pour 14 utilisateurs (**CSI** Partielle).

(b) Évaluation de l'utilité globale pour 20 utilisateurs (**CSI** Partielle).

FIGURE 5.6 – Évaluation des performances en fonction du nombre d'utilisateurs (**CSI** Partielle).

Lorsque le nombre d'utilisateurs passe de 14 à 20, l'utilité globale de **U-CoCRL** augmente de 4.7 à 7.3 bits/s/Hz (une hausse de 55%). Ce résultat, est le point clé de notre approche. Dans un jeu non-coopératif, plus d'utilisateurs signifie plus d'interférences. Dans notre jeu coopératif, plus d'utilisateurs signifie plus d'opportunités de coopération et une plus grande diversité multi-utilisateurs. Notre algorithme est le seul assez adaptatif pour exploiter cette diversité : il réorganise les coalitions pour trouver de meilleurs partenaires, ce qui augmente l'efficacité spectrale globale et surpasse l'effet négatif des interférences.

L'Indice de Performance **PI** (Tableau 5.5) suit une tendance similaire, progressant de 0.32 à 0.36.

TABLE 5.5 – Comparaison de **PI** sous haute interférence (**CSI** Partielle).

Stratégies	PI (14 utilisateurs)	PI (20 utilisateurs)
U-CoCRL	0.32	0.36
Hédonique [Wu et al., 2015]	0.26	0.27
RECORD amélioré	0.26	0.26
RECORD [Ma et al., 2013]	0.18	0.15
Jeu non coopératif	0.09	0.14

L’analyse de ce tableau (14 vs 20 utilisateurs) montre que notre solution est la seule à s’améliorer significativement en scalabilité. Les performances des concurrents sont mitigées : si le RECORD statique se dégrade (0.18 \rightarrow 0.15), les approches Hédonique (0.26 \rightarrow 0.27) et non-coopérative (0.09 \rightarrow 0.14) s’améliorent légèrement. Cependant, cet effet est minime et leur performance reste nettement inférieure à celle de **U-CoCRL** (0.36). En résumé, ces résultats démontrent que **U-CoCRL** parvient à dépasser les effets d’une **CSI** imparfaite et à exploiter l’augmentation du nombre d’utilisateurs pour optimiser les performances, là où les autres méthodes stagnent ou échouent.

Un tableau 5.6 supplémentaire présente une comparaison globale des performances des différents algorithmes évalués selon plusieurs critères clés dans un scénario comportant 14 utilisateurs, notamment l’utilité globale, **PI**, la satisfaction des utilisateurs, la stabilité (au sens de l’équilibre de Pareto), la robustesse ainsi que la scalabilité. Cette comparaison met en évidence la supériorité et l’efficacité de l’algorithme **U-CoCRL** proposé, qui se distingue nettement face aux autres solutions testées dans des conditions de forte interférence.

TABLE 5.6 – Comparaison des performances (14 utilisateurs, **CSI** Partielle, Interférence élevée).

Algorithmes	Utilité globale	PI	Stabilité (Pareto)	Robustesse	Scalabilité
U-CoCRL (proposé)	4.7	0.32	Stable	Excellente	Oui
Hédonique [Wu et al., 2015]	3.89	0.26	Modérée	Faible	Modérée
RECORD amélioré	3.94	0.26	Modérée	Moyenne	Modérée
RECORD [Ma et al., 2013]	2.89	0.18	Faible	Faible	Faible
Non coopératif	1.27	0.09	Très faible	Très faible	Non

L’approche **U-CoCRL** réussit à concilier efficacement deux objectifs souvent contradictoires : maximiser l’efficacité globale du système tout en garantissant une équité dans la répartition des ressources entre les utilisateurs, même dans ce scénario dégradé. Cette capacité repose notamment sur l’intégration de la valeur de Shapley, qui permet d’attribuer de manière juste et transparente la contribution de chaque utilisateur au sein des coalitions formées.

La formation dynamique des coalitions combinée à un mécanisme d'apprentissage par renforcement offre une adaptabilité remarquable face à la variabilité du réseau. Cette dynamique permet à chaque utilisateur d'ajuster son allocation en fonction des changements environnementaux et des comportements des autres joueurs, assurant ainsi une performance globale élevée tout en améliorant l'équité.

Enfin, la capacité de l'algorithme **U-CoCRL** à maintenir ses performances dans des conditions de réseau fluctuantes illustre non seulement son efficacité, mais aussi sa robustesse et sa scalabilité. Ces qualités essentielles nous amènent à poursuivre notre étude en évaluant la complexité algorithmique de la solution proposée, en comparaison avec les autres stratégies, dans la section suivante.

5.3 Discussion sur la complexité de l'algorithme proposé

Afin d'évaluer la faisabilité pratique et l'efficacité de l'algorithme proposé, il est essentiel d'analyser sa complexité computationnelle. Cette étape permet de mesurer les ressources nécessaires en termes de temps de calcul et de mémoire, et d'estimer sa capacité à être déployé dans des environnements réels, souvent contraints par des limitations matérielles et temporelles. Dans cette section, nous détaillons les principaux facteurs influençant la complexité de notre approche, en mettant en lumière les compromis entre performance optimale et coût algorithmique. Les variables utilisées dans cette analyse (Tableau 5.7) sont : K le nombre d'utilisateurs, N_{sc} le nombre de sous-porteuses, T le nombre d'itérations d'apprentissage, N_c le nombre de coalitions, M le nombre d'itérations Monte-Carlo pour l'approximation de Shapley, et N_{cmax} la taille maximale d'une coalition.

Dans cette partie nous évaluons la complexité et l'efficacité de trois approches de formation de coalitions et d'allocation de ressources, telles que présentées dans le tableau 5.7.

TABLE 5.7 – Comparaison de la complexité et de l'efficacité des algorithmes de formation de coalitions.

Méthode	Complexité
Méthode proposée U-CoCRL	$O(MN_{sc}N_c^{\max}2^{N_c^{\max}})$
Hedonic	$O(TKN_c)$
RECORD	$O(2^K)$

La méthode **U-CoCRL** que nous proposons offre une allocation des ressources proche de l'optimal tout en assurant une équité entre les utilisateurs grâce à l'utilisation de la valeur de Shapley. Bien que le calcul exact de cette valeur soit coûteux, son approximation via des techniques de Monte-Carlo [Štrumbelj and Kononenko, 2014] permet une réduction significative de la complexité, rendant le processus plus scalable dans des réseaux de grande taille.

En comparaison, l'algorithme Hédonique repose sur des échanges itératifs entre utilisateurs pour former des coalitions selon leurs préférences individuelles. Bien qu'il présente une complexité algorithmique plus faible, il montre ses limites dans les réseaux à grande

échelle, car l'itération constante entre utilisateurs induit une croissance importante du trafic de contrôle. Concernant l'approche RECORD, elle garantit une solution optimale en matière de formation de coalition, mais son coût computationnel exponentiel en fonction du nombre de nœuds ($O(2^K)$) la rend difficilement applicable dès que l'échelle du réseau augmente.

Ces comparaisons mettent en lumière les compromis inévitables entre optimalité et efficacité : il est souvent nécessaire de sacrifier une partie de l'optimalité pour garantir une meilleure scalabilité et une complexité acceptable dans les systèmes réels.

Avantages face à la complexité accrue : Malgré cette complexité accrue, la méthode [U-CoCRL](#) présente des avantages considérables :

1. **Une allocation des ressources proche de l'optimal** (prouvé à la Figure [5.1](#)).
2. **Une formation des coalitions équitable** (prouvé par le [PI](#) de 0.62 et l'indice de Jain de 0.91).
3. **Une forte adaptabilité** (prouvée par la supériorité en [CSI](#) partielle).

Ces caractéristiques sont particulièrement cruciales dans les réseaux sous-marins où les ressources sont limitées et les communications sujettes à de fortes perturbations. Ainsi, le compromis en termes de complexité est pleinement justifié dans des scénarios où la qualité de service, la résilience et l'équité sont prioritaires.

Scalabilité et performance pratique : En termes de scalabilité, l'algorithme [U-CoCRL](#) reste performant à mesure que le nombre d'utilisateurs et de sous-porteuses augmente. L'utilisation de techniques de clustering, telles que le K-means, permet de réduire considérablement le coût de formation des coalitions en limitant l'espace de recherche. De plus, l'approximation de la valeur de Shapley via Monte-Carlo permet de contenir la croissance de la complexité computationnelle.

Même si l'utilisation de tables Q locales et le suivi des utilités augmentent la consommation de mémoire proportionnellement au nombre d'utilisateurs et de sous-porteuses, cette croissance reste gérable dans des applications pratiques. Des simulations menées avec jusqu'à 20 utilisateurs ont montré un temps d'exécution acceptable et des performances stables, validant ainsi la faisabilité de l'approche pour des réseaux sous-marins de taille moyenne (toutes les simulations ont été réalisées sous MATLAB R2024b sur une machine équipée d'un processeur Intel Core i7 et de 16 Go de RAM).

Enfin, l'efficacité du système a été mesurée en termes de latence computationnelle moyenne par itération, estimée à 163,02 **millisecondes**. Dans nos simulations, l'apprentissage convergeait en $T = 50$ **itérations**. Le temps de convergence total pour une re-configuration est donc d'environ **8,15 secondes** ($163,02\text{ms} \times 50$). Cela permet une adaptation dynamique face à un changement majeur de l'environnement (ex : arrivée de nouveaux utilisateurs), ce qui est adapté aux dynamiques des réseaux sous-marins, sans être du temps réel strict à la milliseconde.

5.4 Conclusion

Dans ce chapitre, nous avons exploré en profondeur les problématiques liées à l'allocation des ressources et à la formation de coalitions dans les réseaux acoustiques sous-

marins sans fil, particulièrement dans des environnements caractérisés par un fort niveau d'interférences. En prenant l'hypothèse d'un canal acoustique statique avec un temps de cohérence suffisamment long et une connaissance parfaite de l'état du canal (CSI), nous avons pu mobiliser la théorie des jeux coopératifs pour concevoir des stratégies efficaces de formation de coalitions. Cette hypothèse, bien que simplificatrice, permet d'établir une base solide pour la modélisation et l'analyse du système dans un premier temps.

L'algorithme proposé, nommé [U-CoCRL](#) constitue la contribution centrale de notre article publié dans IEEE Access [[Gharsalli et al., 2025a](#)]. Il combine plusieurs méthodes innovantes : l'algorithme de clustering K-means pour regrouper efficacement les utilisateurs, l'analyse des contributions individuelles via la valeur de Shapley afin d'assurer une répartition équitable des ressources, et l'apprentissage par renforcement (Q-learning) qui permet une adaptation dynamique des décisions en fonction de l'évolution du réseau. Cette approche hybride vise à concilier la nécessité d'optimisation globale avec les contraintes locales propres à chaque utilisateur, ce qui est fondamental dans des réseaux où la collaboration est essentielle pour améliorer la performance collective.

Les résultats de simulation obtenus confirment l'efficacité de cette méthode, avec des gains significatifs en termes d'utilité globale et d'indice d'équité, même dans des conditions d'interférence élevée. En particulier, l'augmentation de plus de 18% de l'utilité globale et de près de 34% de l'indice de performance met en lumière l'intérêt d'intégrer la théorie des jeux coopératifs avec des techniques d'apprentissage automatique. De plus, l'indice moyen de fairness proche de 0,92 atteste d'une répartition équilibrée des ressources entre les utilisateurs, garantissant ainsi la stabilité et la durabilité des coalitions formées.

Cependant, malgré ces résultats importants, ce travail met également en évidence plusieurs limites, la complexité algorithmique, en particulier liée au calcul de la valeur de Shapley qui est exponentielle en fonction du nombre des utilisateurs, ainsi que le temps nécessaire pour la convergence du Q-learning, constituent des difficultés majeures à la scalabilité du système lorsque le nombre d'utilisateurs augmente. Ces contraintes ouvrent la voie pour à la recherche de solutions plus efficaces et à l'intégration de techniques avancées telles que l'apprentissage profond (ex : DQN), pourraient permettre de gérer des espaces d'états beaucoup plus grands, bien que cela se fasse au prix d'une complexité de calcul accrue.

Conclusion générale et perspectives

5.5 Conclusion générale

Dans ce travail, nous nous sommes intéressés à l'optimisation de l'allocation des ressources dans les réseaux de communication acoustiques sous-marins, un domaine à la fois exigeant et stratégique en raison des nombreuses contraintes physiques et environnementales qui le caractérisent. L'environnement sous-marin impose des défis uniques : bande passante limitée, délais de propagation importants, instabilité du canal acoustique, et forte sensibilité aux interférences, tels sont les caractéristiques étudiées dans le premier chapitre. Face à ces obstacles, notre objectif a été de concevoir des approches auto-organisées, flexibles et robustes, capables d'adapter dynamiquement les stratégies de communication tout en garantissant des performances élevées et une certaine équité entre utilisateurs. Pour cela, nous avons suivi une démarche progressive à travers quatre chapitres complémentaires, chacun apportant une réponse à des limitations identifiées dans le première chapitre.

Dans le deuxième chapitre, nous avons étudié l'interaction entre deux utilisateurs partageant une bande acoustique dans un cadre théorique basé sur la théorie des jeux non coopératifs. En s'inspirant du travail de [Pottier, 2018], consistant à combiner cette approche avec la modulation OFDM et la stratégie de répartition de puissance de type Water-Filling, il est vérifié que chaque utilisateur pouvait, en l'absence de coordination, optimiser sa propre fonction d'utilité face aux interférences générées par l'autre. Cette approche décentralisée a permis de caractériser la région de capacité du système et de mieux comprendre les mécanismes d'influence mutuelle entre les agents. Toutefois, cette première analyse a aussi révélé une limite fondamentale : en l'absence d'un objectif commun et sans partage d'information, les utilisateurs se retrouvent dans des situations d'équilibre sous-optimales, où les performances globales du réseau sont compromises malgré les gains individuels. Ce constat a motivé une évolution du modèle vers une forme de coopération contrôlée entre utilisateurs.

Pour répondre à cette problématique, nous avons introduit dans le troisième chapitre une approche coopérative fondée sur l'apprentissage par renforcement, en particulier le Q-learning, combinée à un mécanisme de négociation inspiré de la solution de Nash. Ce cadre permet aux utilisateurs de prendre des décisions d'allocation de puissance de manière autonome tout en intégrant les préférences des autres agents. L'algorithme ainsi proposé apprend à adapter ses stratégies à mesure que l'environnement évolue, en tenant compte des conditions du canal, du niveau d'interférence et des objectifs des différents utilisateurs. Les résultats obtenus ont montré des améliorations notables en termes d'utilité collective et d'équité par rapport au modèle non coopératif. Toutefois, cette coopération active nécessite des échanges d'information fréquents et engendre un coût algorithmique croissant,

notamment lorsque le nombre d'utilisateurs augmente. Dans des environnements fortement dynamiques comme ceux des communications sous-marines, ces exigences peuvent nuire à la stabilité et à l'efficacité des solutions apprises.

C'est pourquoi, dans le quatrième chapitre, nous avons approfondi la stratégie coopérative en introduisant la formation de coalitions entre utilisateurs. En regroupant les agents selon leurs caractéristiques via l'algorithme de clustering K-means, nous avons pu structurer la coopération de manière plus efficace. La répartition des ressources au sein de chaque coalition a été assurée à l'aide de la valeur de Shapley, garantissant une allocation équitable fondée sur la contribution réelle de chaque utilisateur. Pour maintenir la capacité d'adaptation du système face aux variations du réseau, nous avons de nouveau intégré l'apprentissage par renforcement, permettant à chaque coalition de prendre des décisions optimales dans son propre contexte. L'algorithme hybride [U-CoCRL](#) ainsi conçu a montré des gains significatifs tant en termes de performance globale que d'équité entre utilisateurs, avec un niveau de fairness élevé, preuve de la stabilité et de l'efficacité des coalitions formées, ce qui est démontré par les résultats de simulation présentés dans le dernier chapitre. Néanmoins, certaines limites subsistent, notamment la complexité du calcul de la valeur de Shapley et la dépendance à des hypothèses idéalisées, telles qu'un canal statique et une connaissance parfaite de son état.

Ainsi, à travers l'évolution des modèles présentés dans ce manuscrit, nous avons mis en lumière l'importance croissante de la coopération, de l'adaptabilité et de l'intelligence collective dans les systèmes de communication sous-marins. La transition d'un cadre non coopératif vers des approches de plus en plus collaboratives, enrichies par l'apprentissage automatique, montre que les gains potentiels en performance et en stabilité sont substantiels lorsque les utilisateurs sont incités à partager des informations et à coordonner leurs stratégies. En intégrant des techniques avancées d'optimisation, de négociation et d'apprentissage, ce travail ouvre la voie à la conception de modems acoustiques autonomes, capables de d'auto-adapter et d'auto-organiser aux conditions dynamiques de l'environnement sous-marin.

5.6 Perspectives

Ce travail offre un cadre théorique et algorithmique solide pour l'allocation efficace des ressources dans les réseaux acoustiques sous-marins. Il trace des perspectives prometteuses pour le développement de modems acoustiques intelligents et de réseaux marins auto-organisés. Pour aller plus loin, plusieurs axes de recherche peuvent être explorés.

Amélioration de la scalabilité et apprentissage profond : La complexité algorithmique, en particulier liée au calcul de la valeur de Shapley (exponentielle) et à la taille des tables Q-Learning, constitue un défi majeur à la scalabilité. Pour répondre à cette limite, l'intégration de techniques d'apprentissage profond (Deep Reinforcement Learning), telles que le Deep Q-Network (DQN), est une piste prometteuse. Contrairement à ce qui est parfois suggéré, le DRL n'augmente pas la complexité de calcul, mais il la gère différemment : en utilisant des réseaux de neurones pour approximer la fonction de valeur (au lieu de tables Q), il permet de gérer des espaces d'états de très grande dimension, ce qui est impossible pour le Q-Learning tabulaire [[Malhotra et al., 2025](#)]. L'hybridation de [U-CoCRL](#) avec le DRL permettrait donc de passer à l'échelle de grands réseaux.

Robustesse face aux comportements non-coopératifs : Un deuxième axe essentiel concerne la robustesse du système. Notre modèle [U-CoCRL](#) repose sur une hypothèse de coopération où les nœuds sont rationnels et honnêtes. Il est donc crucial d'étudier la robustesse du système face à des nœuds malveillants ou tricheurs (qui mentiraient sur leur [CSI](#) ou violeraient leur allocation de puissance). Cette problématique diffère des comportements hétérogènes (nœuds gourmands), que la valeur de Shapley gère déjà. Pour contrer la triche, on pourrait intégrer des mécanismes de réputation au sein des coalitions. Les nœuds identifiés comme non-coopératifs seraient alors exclus des coalitions futures, les forçant à un état non-coopératif (seul) où leur performance est minimale.

Positionnement architectural et hybridation SDN (Software-Defined Networking) : Sur le plan architectural, notre approche a été volontairement conçue comme décentralisée pour répondre aux contraintes des réseaux sous-marins : elle évite la latence et le point de défaillance unique d'un contrôleur central. Cependant, les architectures centralisées de type SDN, comme SEANET, offrent une meilleure optimalité théorique grâce à leur vue globale. Une perspective puissante serait de combiner les deux approches. On pourrait imaginer une architecture hybride où les clusters (formés par K-Means) éliraient un "chef" agissant comme un contrôleur SDN local, optimisant les ressources au sein de sa propre coalition. Cela combinerait la robustesse du décentralisé avec l'optimalité locale du centralisé.

Au-delà de ces perspectives théoriques, une ouverture pratique essentielle consiste à positionner l'approche proposée par rapport aux standards de communication sous-marine. Alors que le standard actuel JANUS assure aujourd'hui une interopérabilité minimale (basée sur FSK), il n'est pas conçu pour l'allocation de ressources haut débit en mode partagé. L'émergence future de standards haut débit sous-marins, qui seront très probablement basés sur l'[OFDM](#) pour atteindre des débits élevés, nécessitera des mécanismes de coopération avancés pour gérer les interférences. L'algorithme développé dans ce travail [U-CoCRL](#) pourrait servir de base méthodologique pour cette future normalisation, garantissant une gestion dynamique et coopérative des ressources [OFDM](#). L'évaluation expérimentale est également indispensable. Des tests sur canaux enregistrés [[Stojanovic and Preisig, 2009](#)], puis des expérimentations en temps réel (en bassin puis en mer), permettraient de vérifier la robustesse de notre approche face aux environnements acoustiques réels.

Enfin, il serait intéressant d'étendre les approches développées, notamment en matière de coopération multi-agent et d'allocation de ressources, à d'autres contextes de communication présentant des contraintes similaires. Les réseaux radio ad hoc déployés lors d'opérations de secours, dans des zones complexes ou dans des missions militaires et spatiales partagent en effet des défis tels que la mobilité élevée des nœuds, la limitation des ressources et la nécessité d'une allocation dynamique et équitable. Les principes de coopération multi-agent et d'allocation équitable des ressources, tels que développés pour les réseaux acoustiques sous-marins, constituent des solutions méthodologiques transférables, pouvant inspirer la conception de systèmes robustes et adaptatifs dans ces environnements. Par exemple, dans les réseaux de véhicules, l'apprentissage par renforcement profond multi-agent combiné à l'apprentissage fédéré asynchrone a été utilisé avec succès pour optimiser l'accès au spectre, la puissance de transmission et la bande passante [[Liu and Ma, 2025](#)]. Ces travaux illustrent que des méthodes similaires peuvent être

adaptées aux réseaux acoustiques sous-marins pour améliorer la résilience et l'efficacité de l'allocation des ressources, tout en tenant compte des contraintes propres à cet environnement.

Publications

Articles de conférences internationales et nationale avec comité de lecture

1. K. Gharsalli, P. -J. Bouvet, S. Najeh, H. Besbes, T. L. Pors and H. Gazzah, "Non-cooperative Resource Allocation Game-Theoretic Approach in Underwater Acoustics," 2023 International Wireless Communications and Mobile Computing (IWCMC), Marrakesh, Morocco, 2023, pp. 632-637, doi : 10.1109/IWCMC58020.2023.10182682.
2. K. Gharsalli, T. L. Pors, S. Najeh, P. -J. Bouvet and H. Besbes, "Nash Bargaining Game for Cooperative Resource Allocation in Underwater Acoustic Communication," 2024 International Wireless Communications and Mobile Computing (IWCMC), Ayia Napa, Cyprus, 2024, pp. 1016-1021, doi : 10.1109/IWCMC61514.2024.10592489.
3. K. Gharsalli, T. L. Pors, S. Najeh, L. NAJJAR, P. -J. Bouvet and H. Besbes, "Optimizing Underwater Acoustic Communication with Game Theory and Machine Learning," OCEANS 2025 conference, BREST, FRANCE.
4. K. Gharsalli, S. Najeh, T. L. Pors, P. -J. Bouvet and H. Besbes, "Cooperative Repeated Game with Q-Learning for Fair Multi-User Underwater Acoustic Networks," NTRICS 2026.

Article de journal

1. K. Gharsalli, S. Najeh, T. L. Pors, L. Najjar, P. -J. Bouvet and H. Besbes, "Coalition Game-Based Reinforcement Learning in Underwater Acoustic Networks," in IEEE Access 2025, doi : 10.1109/ACCESS.2025.3582918.

Communication scientifique

1. Présentation en séminaire informel à l'ISEN 2023.
2. Présentation aux journées des Doctoriales PHC-Utique 2024.

Annexes

Algorithm 1 Nash Bargaining Algorithm – Two-User Case : Inspiré de la solution de Nash, ce mécanisme itératif permet à chaque utilisateur d’ajuster progressivement sa répartition de puissance sur les sous-porteuses afin d’améliorer l’utilité individuelle et les performances globales du système.

Inputs : $h_{11}, h_{22}, h_{12}, h_{21}, N_{sc}, P^{\max}, SNR_{i,n}, R_i^{\min}, N_{\text{offer}}, L$

- 1: **Initialize :**
- 2: **for** each user $i = 1, 2$ **do**
- 3: Set power allocation : $P_i \leftarrow \frac{P^{\max}}{N_{sc}}$
- 4: **end for**
- 5: **for** $\ell = 1$ to L **do**
- 6: **for** each user $i = 1, 2$ **do**
- 7: Propose N_{offer} subcarriers to the other user based on local SNR
- 8: **end for**
- 9: Identify the set of common subcarriers with high SNR : N_{Csc}
- 10: Share N_{Csc} based on weights α_i derived from channel gains :

$$\alpha_i = \frac{h_{ii,n}}{h_{ii,n} + h_{jj,n}}, \quad \alpha_j = 1 - \alpha_i$$

- 11: Allocate the remaining subcarriers using round-robin or priority-based rule
- 12: Update power allocations on each subcarrier, considering mutual interference
- 13: **for** each user $i = 1, 2$ **do**
- 14: Compute data rate :

$$R_i = \sum_n \log_2(1 + \text{SINR}_{i,n})$$

- 15: **end for**
- 16: Evaluate global utility :

$$U(\mathbf{P}) = \prod_{i=1}^2 (R_i - R_i^{\min})$$

- 17: **end for**
 - 18: **Output :** Power allocation \mathbf{P} , utility $U(\mathbf{P})$
-

Algorithm 2 Nash Q-learning-based Power Allocation – Multi-Users Case

Inputs : $h_{ii}, h_{ji}, K, N_{sc}, P^{\max}, P^{\min}, R_i^{\min}, \alpha, \beta, \tau, \epsilon.$

```

1: for each user  $i = 1$  to  $K$  do
2:   Set initial power allocation :  $p_i \leftarrow \frac{P^{\max}}{N_{sc}} \cdot \text{ones}(N_{sc})$ 
3:   Initialize Q-table :  $Q_i(s, a) \leftarrow 0$ 
4:   Estimate disagreement point :  $d_i \leftarrow \min(\mathbb{E}[R_i])$  using non-cooperative strategy
5: end for
6: while not converged do
7:   Proposal Stage :
8:   for each user  $i$  do : Select subcarriers where  $R_i(n) > R_i^{\min}$ 
9:   end for
10:  if multiple users propose the same subcarrier then
11:    Resolution Stage via Nash Q-learning :
12:    for each user  $i$  do : Compute utility :

```

$$r_i(\mathbf{P}) = \sum_{n \in \mathcal{N}_i} \log_2(1 + \gamma_{i,n}(\mathbf{P}))$$

```

13:      Use softmax action selection to choose new power allocation :

```

$$P_i^{\text{new}} \leftarrow P^{\min} + (P^{\max} - P^{\min}) \cdot \text{find}(\text{rand} < \text{cumsum}(\pi), 1)$$

```

14:      Calculate new utility :  $r_i^{\text{new}} \leftarrow r_i(\mathbf{P}^{\text{new}})$ 

```

```

15:      Compute reward :  $\text{re} \leftarrow r_i^{\text{new}} - r_i(\mathbf{P})$ 

```

```

16:      Update Q-table :

```

$$Q_i(s, a) \leftarrow (1 - \alpha)Q_i(s, a) + \alpha \left[\text{re} + \beta \max_a Q_i(s', a) \right]$$

```

17:      if  $\text{re} > 0$  then

```

```

18:        Update power allocation :  $P_i \leftarrow P_i^{\text{new}}$ 

```

```

19:      end if

```

```

20:    end for

```

```

21:  end if

```

```

22:  Fairness Optimization :

```

```

23:  if  $\exists i, j$  such that  $R_i < R_j$  then

```

```

24:    Evaluate subcarrier exchange to improve fairness

```

```

25:    if global fairness increases then

```

```

26:      Perform subcarrier swap

```

```

27:    end if

```

```

28:  end if

```

```

29:  Convergence Check : Compute Nash bargaining utility :

```

$$U(P) = \prod_{i=1}^K (R_i - R_i^{\min})^{1/K}$$

```

30:    if  $|Q^{\text{new}} - Q^{\text{old}}| < \epsilon$  then

```

```

31:      Converged  $\leftarrow$  True

```

```

32:    end if

```

```

33:  end while

```

```

34:  Output :  $P^*, U^*$ 

```

Cette stratégie conjointe repose à la fois sur les principes de la négociation de Nash pour encadrer les interactions entre agents, et sur le Q-learning pour permettre une adaptation progressive aux variations de l'environnement. Cette combinaison est proposée dans le but d'avoir une meilleure efficacité, robustesse et équité dans le partage des ressources.

À mesure que le nombre d'agents dans le réseau augmente, les approches basées sur la négociation de Nash deviennent coûteuses et difficiles à faire converger, en raison de la complexité croissante des interactions. Pour surmonter cette limite, il devient pertinent de se tourner vers les jeux de coalition, qui permettent aux agents de former des groupes coopératifs restreints. En combinant cette approche avec le Q-learning, chaque agent peut apprendre à coopérer efficacement au sein d'une coalition, réduisant ainsi l'interférence globale tout en améliorant l'équité et la performance réseau de manière décentralisée. Les algorithmes ci-dessous décrivent les mécanismes proposés pour cette approche.

Commençant par l'algorithme d'**Allocation et formation de coalition initiale** :

Algorithm 3 Initial coalition formation and resource allocation

```

1: Input :  $h_{ii}, h_{ji}, K, N_{sc}, P^{\max}, \gamma_{i,n}$ .
2: Step 1 : Power Allocation using Water-Filling
3: for each user  $i = 1$  to  $K$  do
4:   for each sub-carrier  $n \in \mathcal{N}_i$  do
5:     Compute power  $p_{i,n}$  using (4.32)
6:   end for
7: end for
8: Step 2 : SINR and Transmission Rate Calculation
9: for each user  $i = 1$  to  $K$  do
10:  for each sub-carrier  $n \in \mathcal{N}_i$  do
11:    Evaluate  $\gamma_{i,n}$  using (4.25)
12:    Compute transmission rate  $R_{i,n}$  using (4.26)
13:  end for
14: end for
15: Step 3 : Initial Coalition Formation using K-means
16: Initialize centroids  $\mu_c$  for  $c = 1$  to  $\mathcal{N}_c$ 
17: repeat
18:  for each user  $i = 1$  to  $K$  do
19:    Assign  $i$  to the coalition  $C_c$  such that :
    
$$c = \arg \min_{c' \in \{1, \dots, \mathcal{N}_c\}} \|u_i - \mu_{c'}\|^2$$

20:  end for
21:  for each coalition  $C_c$  do
22:    Update centroid  $\mu_c$  as the mean utility of users in  $C_c$ 
23:  end for
24: until centroids  $\mu_c$  converge
25: return  $P, N_i, \mathcal{C}$ 

```

Après la formation initiale de coalition, il s'agira de résoudre les conflits d'intérêts et d'optimiser l'allocation de puissance au sein de chaque coalition, en s'appuyant à la fois sur les valeurs de Shapley pour un partage équitable des ressources et sur un algorithme de négociation de puissance pour ajuster dynamiquement les stratégies des agents en appliquant l'algorithme suivant.

Optimisation intra-coalition :

Algorithm 4 Intra-Coalition Evaluation

- 1: **Input** : $p_{i,n}$, $\mathcal{C} = \{C_1, C_2, \dots, C_{N_c}\}$, N_i, N_j , $\mathcal{A} = \emptyset$
- 2: **Step 1 : Conflict Detection and Shapley Values Calculation**
- 3: **for** each coalition $C \in \mathcal{C}$ **do**
- 4: **for** each pair of users $i, j \in C$ **do**
- 5: Calculate $\mathcal{A} \leftarrow \{n \mid n \in \bigcap_{i,j \in C, i \neq j} (N_i \cap N_j)\}$
- 6: **end for**
- 7: **for** each user $i \in A_c$ **do**
- 8: Evaluate the marginal contribution of user i using Monte Carlo approximation :

$$\hat{\phi}_i(A_c) = \frac{1}{M} \sum_{m=1}^M [u(S_m \cup \{i\}) - u(S_m)] \quad (5.4)$$

- 9: **end for**
 - 10: **end for**
 - 11: **Step 2 : Power Negotiation and Interference Management**
 - 12: **for** $c = 1$ to N_c **do**
 - 13: **for** each user $i \in C_c$ **do**
 - 14: Adjust power allocation based on Shapley values using (4.38)
 - 15: Redistribute power among users through negotiation
 - 16: Evaluate interference levels and update utility functions
 - 17: **end for**
 - 18: **end for**
 - 19: **Step 3 : User Satisfaction Evaluation**
 - 20: **for** each user $i \in C_c$ **do**
 - 21: **if** $u_i^{\text{new}} \geq u_i^{\text{old}}$ **then**
 - 22: User i is satisfied
 - 23: **else**
 - 24: User i is not satisfied
 - 25: **end if**
 - 26: **end for**
 - 27: **Step 4 : Coalition Utility Computation**
 - 28: **for** each coalition C_c **do**
 - 29: Calculate U as defined in (4.27)
 - 30: **end for**
 - 31: **Output** : $p_{i,n}^{\text{new}}$, u_i^{new} , U
-

Après la résolution des conflits d'intérêts au sein de chaque coalition, une phase d'apprentissage par renforcement est introduite pour permettre une adaptation dynamique des stratégies d'allocation de puissance. Le Q-learning est ainsi utilisé afin d'optimiser les performances individuelles et collectives tout en maintenant l'équilibre et la stabilité des coalitions. Ce processus adaptatif est implémenté à travers l'algorithme détaillés ci-après.

Apprentissage par renforcement pour optimiser la répartition des utilisateurs au sein des coalitions finales :

β est un paramètre (avec $0 \leq \beta \leq 1$) qui ajuste l'importance relative de la satisfaction individuelle par rapport à la satisfaction collective. L'expression $\max_{a'} Q_i(s', a')$ représente l'utilité maximale attendue dans l'état suivant, en considérant toutes les actions possibles a' .

Algorithm 5 Decentralized Coalition Optimization via Q-Learning

Initialize : $Q_i(s, a) = 0, \forall i, s, a$

Input : $p_{i,n}^{\text{new}}, u_i^{\text{new}}, U, \eta, \gamma, \epsilon, N_C^{\text{max}}, T$

1: **for** $t = 1$ **to** T **do**

2: **for** each user $i \in A_c$ **do**

3: **Step 1 :** Exchange utility information with neighboring users in coalition C

4: $u_j^{\text{new}} \leftrightarrow u_i^{\text{new}}$ for each user $j \in C$

5: **Step 2 :** Compute IS using (4.39)

6: **Step 3 :** Compute CS

$$CS = \frac{1}{|C|} \sum_{j \in A_c} u_j^{\text{new}} \quad (5.5)$$

7: **Step 4 :** Compute reward

$$r_i = \beta \cdot IS_i + (1 - \beta) \cdot CS \quad (5.6)$$

8: **Step 5 :** Update Q-values

$$Q_i(s, a) = Q_i(s, a) + \eta \left[r_i + \gamma \max_{a'} Q_i(s', a') - Q_i(s, a) \right] \quad (5.7)$$

9: **Step 6 : Exploration**

10: **if** $IS_i < 0$ and $\text{rand} < \epsilon$ **then**

11: **for** each neighboring coalition $C' \in \mathcal{C}$ **do**

12: **Step 6a :** Exchange information with users in C'

13: $u_j^{\text{new}} \leftrightarrow u_i^{\text{new}}$ for each user $j \in C'$

14: **Step 6b :** Evaluate utility in the new coalition :

$$u_i^{\text{new}}(C') = f(C', i) \quad (5.8)$$

15: **Step 6c :** Compute new collective satisfaction in C'

$$U^{\text{new}} = \frac{1}{|C'|} \sum_{j \in C'} u_j^{\text{new}} \quad (5.9)$$

16: **if** $U^{\text{new}} > U$ **then**

17: **Switch coalition :** Move user i to coalition C'

18: **Break loop**

19: **end if**

20: **end for**

21: **end if**

22: **end for**

23: **end for**

24: **Output :** $P^*, U^*, C^*, Q_i(s, a)$

Bibliographie

- [Abe and Nakada, 2023] Abe, T. and Nakada, S. (2023). Core stability of the shapley value for cooperative games. *Social Choice and Welfare*, 60(4) :523–543.
- [Abramson, 2003] Abramson, N. (2003). Development of the alohanet. *IEEE transactions on Information Theory*, 31(2) :119–123.
- [Ahmad et al., 2018] Ahmad, A.-M., Barbeau, M., Garcia-Alfaro, J., Kassem, J., Krana-kis, E., and Porretta, S. (2018). Doppler effect in the underwater acoustic ultra low frequency band. In *Ad Hoc Networks : 9th International Conference, AdHocNets 2017, Niagara Falls, ON, Canada, September 28–29, 2017, Proceedings*, pages 3–12. Springer.
- [Akkarajitsakul et al., 2011] Akkarajitsakul, K., Hossain, E., and Niyato, D. (2011). Distributed resource allocation in wireless networks under uncertainty and application of bayesian game. *IEEE Communications Magazine*, 49(8) :120–127.
- [Akyildiz et al., 2004] Akyildiz, I. F., Pompili, D., and Melodia, T. (2004). Challenges for efficient communication in underwater acoustic sensor networks. *ACM Sigbed Review*, 1(2) :3–8.
- [Akyildiz et al., 2005] Akyildiz, I. F., Pompili, D., and Melodia, T. (2005). Underwater acoustic sensor networks : research challenges. *Ad hoc networks*, 3(3) :257–279.
- [Ali and Miraz, 2020] Ali, M. and Miraz, M. H. (2020). A review of underwater acoustic, electromagnetic and optical communications. In *International Conference for Emerging Technologies in Computing*, pages 86–97. Springer.
- [Ao et al., 2023] Ao, S., Niu, Y., Han, Z., Ai, B., Zhong, Z., Wang, N., and Qiao, Y. (2023). Resource allocation for ris-assisted device-to-device communications in heterogeneous cellular networks. *IEEE Transactions on Vehicular Technology*, 72(9) :11741–11755.
- [Ausubel et al., 2002] Ausubel, L. M., Cramton, P., and Deneckere, R. J. (2002). Bargaining with incomplete information. In Aumann, R. J. and Hart, S., editors, *Handbook of Game Theory with Economic Applications*, volume 3, pages 1897–1945. Elsevier.
- [Avni et al., 2016] Avni, G., Henzinger, T. A., and Kupferman, O. (2016). Dynamic resource allocation games. In *International Symposium on Algorithmic Game Theory (SAGT)*, volume 9928 of *Lecture Notes in Computer Science*, pages 153–166, Berlin, Heidelberg. Springer Berlin Heidelberg.
- [Banerjee et al., 2001] Banerjee, S., Konishi, H., and Sönmez, T. (2001). Core in a simple coalition formation game. *Social Choice and Welfare*, 18(1) :135–153.
- [Bardhan and Ghose, 2013a] Bardhan, R. and Ghose, D. (2013a). Resource allocation and coalition formation for uavs : A cooperative game approach. In *2013 IEEE International Conference on Control Applications (CCA)*, pages 1200–1205.
- [Bardhan and Ghose, 2013b] Bardhan, R. and Ghose, D. (2013b). Resource allocation and coalition formation for uavs : A cooperative game approach. In *2013 IEEE International Conference on Control Applications (CCA)*, pages 1200–1205. IEEE.

- [Bello, 1963] Bello, P. (1963). Characterization of randomly time-variant linear channels. *IEEE transactions on Communications Systems*, 11(4) :360–393.
- [Benchekroun and Van Long, 2011] Benchekroun, H. and Van Long, N. (2011). Game theory : Static and dynamic games. In de Zeeuw, A. and Batabyal, A. A., editors, *Research Tools in Natural Resource and Environmental Economics*, pages 89–139. Edward Elgar Publishing, Cheltenham, UK and Northampton, MA, USA.
- [Bui et al., 2023] Bui, T. T., Nguyen, L. D., Kha, H. H., Vo, N.-S., and Duong, T. Q. (2023). Joint clustering and resource allocation optimization in ultra-dense networks with multiple drones as small cells using game theory. *Sensors*, 23(8) :3899.
- [Burrowes and Khan, 2009] Burrowes, G. E. and Khan, J. Y. (2009). Adaptive token polling mac protocol for wireless underwater networks. In *2009 4th International Symposium on Wireless Pervasive Computing*, pages 1–5.
- [Catipovic, 2002] Catipovic, J. A. (2002). Performance limitations in underwater acoustic telemetry. *IEEE Journal of Oceanic Engineering*, 15(3) :205–216.
- [Chalkiadakis and Boutilier, 2004] Chalkiadakis, G. and Boutilier, C. (2004). Bayesian reinforcement learning for coalition formation under uncertainty. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, volume 3, pages 1090–1097, New York, NY, USA. IEEE.
- [Chang et al., 2016] Chang, M.-K., Chien, F.-T., Chen, T.-C., and Li, K.-L. (2016). Stochastic game-theoretical power allocation in d2d communications. In *2016 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, pages 1–4.
- [Chatain, 2016] Chatain, O. (2016). Cooperative and non-cooperative game theory. In *The Palgrave Encyclopedia of Strategic Management*, pages 1–3. Springer.
- [Chatterjee and Das, 2015] Chatterjee, S. and Das, S. (2015). Ant colony optimization based enhanced dynamic source routing algorithm for mobile ad-hoc network. *Information sciences*, 295 :67–90.
- [Chauvaud et al., 2018] Chauvaud, S., Chauvaud, L., and Jolivet, A. (2018). *Impacts des sons anthropiques sur la faune marine*. Éditions Quae.
- [Che et al., 2010] Che, X., Wells, I., Dickers, G., Kear, P., and Gong, X. (2010). Re-evaluation of rf electromagnetic communication in underwater sensor networks. *IEEE Communications Magazine*, 48(12) :143–151.
- [Chee et al., 2006] Chee, T. K., Lim, C.-C., and Choi, J. (2006). A cooperative game theoretic framework for resource allocation in ofdma systems. In *2006 10th IEEE Singapore International Conference on Communication Systems*, pages 1–5. IEEE.
- [Chithra et al., 2015] Chithra, K., Sireesha, N., Thangavel, C., Gowthaman, V., Narayanan, S. S., Sudhakar, T., and Atmanand, M. (2015). Underwater communication implementation with ofdm. *Indian Journal of Geo-marine sciences*, 44(2) :259–266.
- [Coates and Coates, 1990] Coates, R. F. and Coates, R. (1990). *Underwater acoustic systems*. Springer.
- [Cournot, 1863] Cournot, A. A. (1863). *Principes de la théorie des richesses*. Hachette.
- [Cover, 1999] Cover, T. M. (1999). *Elements of information theory*. John Wiley & Sons.
- [Cui, 2020] Cui, M. (2020). Introduction to the k-means clustering algorithm based on the elbow method. In *Geoscience and Remote Sensing*, volume 3, pages 9–16. Publisher information if available.

- [Cui et al., 2023] Cui, X., Yan, P., Li, J., Li, S., and Liu, J. (2023). Deep reinforcement learning-based adaptive modulation for ofdm underwater acoustic communication system. *EURASIP Journal on Advances in Signal Processing*, 2023(1) :1.
- [Dai et al., 2023] Dai, X., Shi, C., Wang, Z., and Zhou, J. (2023). Coalition game-theoretic power allocation strategy for multi-target detection in distributed radar networks. *Remote Sensing*, 15(15) :3804.
- [Fang et al., 2021] Fang, L., Li, H., and Wang, Z. (2021). Centralized resource allocation based on the bargaining approach. *Asia-Pacific Journal of Operational Research*, 38(02) :2050049.
- [Fudenberg and Tirole, 1991] Fudenberg, D. and Tirole, J. (1991). *Game theory*. MIT press.
- [Gao et al., 2022] Gao, C., Hu, W., and Chen, K. (2022). Research on multi-auvs data acquisition system of underwater acoustic communication network. *Sensors*, 22(14) :5090.
- [Gharsalli et al., 2023a] Gharsalli, K., Bouvet, P.-J., Najeh, S., Besbes, H., Le Pors, T., and Gazzah, H. (2023a). Non-cooperative resource allocation game-theoretic approach in underwater acoustics. In *2023 International Wireless Communications and Mobile Computing (IWCMC)*, pages 632–637. IEEE.
- [Gharsalli et al., 2023b] Gharsalli, K., Bouvet, P.-J., Najeh, S., Besbes, H., Pors, T. L., and Gazzah, H. (2023b). Non-cooperative resource allocation game-theoretic approach in underwater acoustics. In *2023 International Wireless Communications and Mobile Computing (IWCMC)*, pages 632–637.
- [Gharsalli et al., 2024a] Gharsalli, K., Le Pors, T., Najeh, S., Bouvet, P.-J., and Besbes, H. (2024a). Nash bargaining game for cooperative resource allocation in underwater acoustic communication. In *2024 International Wireless Communications and Mobile Computing (IWCMC)*, pages 1016–1021. IEEE.
- [Gharsalli et al., 2025a] Gharsalli, K., Najeh, S., Le Pors, T., Najjar, L., Bouvet, P.-J., and Besbes, H. (2025a). Coalition game-based reinforcement learning in underwater acoustic networks. *IEEE Access*, 13 :124602–124621.
- [Gharsalli et al., 2024b] Gharsalli, K., Pors, T. L., Najeh, S., Bouvet, P.-J., and Besbes, H. (2024b). Nash bargaining game for cooperative resource allocation in underwater acoustic communication. In *2024 International Wireless Communications and Mobile Computing (IWCMC)*, pages 1016–1021.
- [Gharsalli et al., 2025b] Gharsalli, K., Pors, T. L., Najeh, S., Najjar, L., Bouvet, P.-J., and Besbes, H. (2025b). Optimizing underwater acoustic communication with game theory and machine learning. In *OCEANS 2025 Brest*, pages 1–7.
- [Gibbons, 1997] Gibbons, R. (1997). An introduction to applicable game theory. *Journal of Economic Perspectives*, 11(1) :127–149.
- [Grindheim, 2018] Grindheim, J. V. (2018). Underwater acoustics - fundamentals, applications and limitations within underwater communication and positioning. NMBU, Ås, Norway. Conference presentation.
- [Guan et al., 2015] Guan, X., Li, A., Cai, Z., and Ohtsuki, T. (2015). Coalition graph game for robust routing in cooperative cognitive radio networks. *Mobile Networks and Applications*, 20 :147–156.

- [Han et al., 2018] Han, S., Li, X., Yan, L., Xu, J., Liu, Z., and Guan, X. (2018). Joint resource allocation in underwater acoustic communication networks : A game-based hierarchical adversarial multiplayer multiarmed bandit algorithm. *Information Sciences*, 454 :382–400.
- [He et al., 2019] He, C., Chen, Q., Pan, C., Li, X., and Zheng, F.-C. (2019). Resource allocation schemes based on coalition games for vehicular communications. *IEEE Communications Letters*, 23(12) :2340–2343.
- [Heidemann et al., 2012] Heidemann, J., Stojanovic, M., and Zorzi, M. (2012). Underwater sensor networks : applications, advances and challenges. *Philosophical Transactions of the Royal Society A : Mathematical, Physical and Engineering Sciences*, 370(1958) :158–175.
- [Hu et al., 2020] Hu, Y., Chen, M., Yang, Z., Chen, M., and Jia, G. (2020). Optimization of resource allocation in multi-cell ofdm systems : A distributed reinforcement learning approach. In *2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications*, pages 1–6. IEEE.
- [Huang et al., 2021] Huang, R., Si, J., Shi, J., and Li, Z. (2021). Deep-reinforcement-learning-based resource allocation in ultra-dense network. In *Proceedings of the 2021 13th International Conference on Wireless Communications and Signal Processing (WCSP)*, pages 1–5, New York, NY, USA. IEEE.
- [Hwang et al., 2006] Hwang, C.-S., Seong, K., and Cioffi, J. M. (2006). Opportunistic p-persistent csma in wireless networks. In *2006 IEEE International Conference on Communications*, volume 1, pages 183–188. IEEE.
- [Jain et al., 1984] Jain, R. K., Chiu, D.-M. W., and Hawe, W. R. (1984). A quantitative measure of fairness and discrimination. Technical Report DEC-TR-301, Eastern Research Laboratory, Digital Equipment Corporation, Hudson, MA. 21(1), 2022–2023.
- [Ji et al., 2025] Ji, M., Xu, G., and Wang, L. (2025). Cora : Coalitional rational advantage decomposition for multi-agent policy gradients. *arXiv preprint arXiv :2506.04265*.
- [Jornet et al., 2008] Jornet, J. M., Stojanovic, M., and Zorzi, M. (2008). Focused beam routing protocol for underwater acoustic networks. In *Proceedings of the 3rd International Workshop on Underwater Networks*, pages 75–82.
- [Ju et al., 2024] Ju, S., Zhang, Z., and Song, W. (2024). Dynamic resource allocation strategy based on k-means. In *2024 IEEE 2nd International Conference on Image Processing and Computer Applications (ICIPCA)*, pages 262–265. IEEE.
- [Kang, 2021] Kang, H. (2021). Performance analysis of ds-cdma system for underwater acoustic communications. *International Journal of Engineering Applied Sciences and Technology*.
- [Kaveripakam and Chinthaginjala, 2023] Kaveripakam, S. and Chinthaginjala, R. (2023). Clustering-based dragonfly optimization algorithm for underwater wireless sensor networks. *Alexandria Engineering Journal*, 81 :580–598.
- [Kazemi and Soleimani, 2025] Kazemi, U. and Soleimani, S. (2025). A new approach data processing : Density-based spatial clustering of applications with noise (dbscan) clustering using game-theory. *Soft Computing*, pages 1–16.
- [Khan et al., 2014] Khan, M. W., Zhou, Y., and Xu, G. (2014). Modeling of acoustic propagation channel in underwater wireless sensor networks. In *The 2014 2nd International Conference on Systems and Informatics (ICSAI 2014)*, pages 586–590. IEEE.

- [Khan et al., 2016] Khan, U. U., Dilshad, N., Rehmani, M. H., and Umer, T. (2016). Fairness in cognitive radio networks : Models, measurement methods, applications, and future research directions. *Journal of Network and Computer Applications*, 73 :12–26.
- [Khan et al., 2010] Khan, Z., Glisic, S., DaSilva, L. A., and Lehtomäki, J. (2010). Modeling the dynamics of coalition formation games for cooperative spectrum sharing in an interference channel. *IEEE Transactions on Computational Intelligence and AI in Games*, 3(1) :17–30.
- [Kim, 2018] Kim, S. (2018). A better-performing q-learning game-theoretic distributed routing for underwater wireless sensor networks. *International Journal of Distributed Sensor Networks*, 14(1) :1550147718754728.
- [Knudsen et al., 1948] Knudsen, V. O., Alford, R., and Emling, J. (1948). Underwater ambient noise. *J. mar. Res.*, 7(3) :410–429.
- [Kranich et al., 2005] Kranich, L., Perea, A., and Peters, H. (2005). Core concepts for dynamic tu games. *International Game Theory Review*, 7(01) :43–61.
- [Krishnaswamy and Manvi, 2019] Krishnaswamy, V. and Manvi, S. S. (2019). Fuzzy and pso based clustering scheme in underwater acoustic sensor networks using energy and distance parameters. *Wireless Personal Communications*, 108(3) :1529–1546.
- [Lam, 2003] Lam, S. (2003). Delay analysis of a time division multiple access (tdma) channel. *IEEE Transactions on Communications*, 25(12) :1489–1494.
- [Lee and Kim, 2023] Lee, I. and Kim, D. K. (2023). Marl-based resource allocation for heterogeneous traffic in v2x communications. In *2023 28th Asia Pacific Conference on Communications (APCC)*, pages 61–67. IEEE.
- [Liu and Ma, 2025] Liu, Q. and Ma, Y. (2025). Communication resource allocation method in vehicular networks based on federated multi-agent deep reinforcement learning. *Scientific Reports*, 15(1) :30866.
- [Luo and Wang, 2025] Luo, X. and Wang, H. (2025). Underwater acoustic cooperative communication network : A stackelberg game-based power control method. *Sensors & Materials*, 37.
- [Lurton, 1998] Lurton, X. (1998). *Acoustique sous-marine : présentation et applications*. Editions Quae.
- [Lurton, 2002] Lurton, X. (2002). *An introduction to underwater acoustics : principles and applications*. Springer Science & Business Media.
- [Ma et al., 2013] Ma, B., Cheung, M. H., and Wong, V. W. S. (2013). Interference management for multimedia femtocell networks with coalition formation game. In *2013 IEEE International Conference on Communications (ICC)*, pages 6112–6117, Piscataway, NJ, USA. IEEE.
- [Madani et al., 2022] Madani, K., Farhidi, F., and Gholizadeh, S. (2022). Bargaining power in cooperative resource allocations games. *Algorithms*, 15(12) :445.
- [Malhotra et al., 2025] Malhotra, S., Yashu, F., Saqib, M., Mehta, D., Jangid, J., and Dixit, S. (2025). Deep reinforcement learning for dynamic resource allocation in wireless networks. *arXiv preprint arXiv :2502.01129*.
- [Martin et al., 2018] Martin, J., Wang, J., and Englot, B. (2018). Sparse gaussian process temporal difference learning for marine robot navigation. In *Conference on Robot Learning*, pages 179–189. PMLR.

- [Medwin, 1975] Medwin, H. (1975). Speed of sound in water : A simple equation for realistic parameters.
- [Mendez and James, 2015] Mendez, P. A. and James, R. (2015). A comparative study of underwater wireless optical communication for three different communication links. *IOSR J Electron Commun Eng*, 10(3) :40–48.
- [Merchant et al., 2022] Merchant, N. D., Putland, R. L., André, M., Baudin, E., Felli, M., Slabbekoorn, H., and Dekeling, R. (2022). A decade of underwater noise research in support of the european marine strategy framework directive. *Ocean & Coastal Management*, 228 :106299.
- [Molins and Stojanovic, 2006] Molins, M. and Stojanovic, M. (2006). Slotted fama : a mac protocol for underwater acoustic networks. In *OCEANS 2006-Asia Pacific*, pages 1–7. IEEE.
- [Moose, 1994] Moose, P. (1994). A technique for orthogonal frequency division multiplexing frequency offset correction. *IEEE Transactions on Communications*, 42(10) :2908–2914.
- [Morozs et al., 2017] Morozs, N., Mitchell, P., and Zakharov, Y. V. (2017). Tda-mac : Tdma without clock synchronization in underwater acoustic networks. *IEEE Access*, 6 :1091–1108.
- [Muhammed et al., 2018] Muhammed, D., Anisi, M. H., Zareei, M., Vargas-Rosales, C., and Khan, A. (2018). Game theory-based cooperation for underwater acoustic sensor networks : Taxonomy, review, research challenges and directions. *Sensors*, 18(2) :425.
- [Myung et al., 2007] Myung, H. G., Lim, J., and Goodman, D. J. (2007). Single carrier fdma for uplink wireless transmission. *IEEE vehicular technology magazine*, 1(3) :30–38.
- [Nash et al., 1950] Nash, J. F. et al. (1950). The bargaining problem. *Econometrica*, 18(2) :155–162.
- [Nash Jr, 1950] Nash Jr, J. F. (1950). Equilibrium points in n-person games. *Proceedings of the national academy of sciences*, 36(1) :48–49.
- [Ng et al., 2008] Ng, H.-H., Soh, W.-S., and Motani, M. (2008). Maca-u : A media access protocol for underwater acoustic networks. In *IEEE GLOBECOM 2008-2008 IEEE Global Telecommunications Conference*, pages 1–5. IEEE.
- [Osborne and Rubinstein, 1994] Osborne, M. J. and Rubinstein, A. (1994). *A course in game theory*. MIT press.
- [Pan et al., 2012] Pan, C., Jia, L., Cai, R., and Ding, Y. (2012). Modeling and simulation of channel for underwater communication network. *Int. J. Innov. Comput. Inf. Control*, 8 :2149–2156.
- [Pan et al., 2024] Pan, G., Li, L., Fan, R., and Su, Y. (2024). A joint helper selection and power allocation scheme for secure underwater acoustic transmission. *Computer Networks*, 254 :110779.
- [Peters, 2009] Peters, H. (2009). Ntu-games. In *Optimization and Operations Research*, volume 3, pages 181–216. EOLSS Publishers. Encyclopedia of Life Support Systems (EOLSS).
- [Petroccia and Alves, 2024] Petroccia, R. and Alves, J. (2024). The janus underwater communications standard : From promulgation to present. In *OCEANS 2024 - Singapore*, pages 1–9.

- [Pompili and Akyildiz, 2009] Pompili, D. and Akyildiz, I. F. (2009). Overview of networking protocols for underwater wireless communications. *IEEE Communications magazine*, 47(1) :97–102.
- [Potter et al., 2014] Potter, J., Alves, J., Green, D., Zappa, G., Nissen, I., and McCoy, K. (2014). The janus underwater communications standard. In *2014 underwater communications and networking (UComms)*, pages 1–4. IEEE.
- [Pottier, 2018] Pottier, A. (2018). *Méthodes décentralisées d’allocation des ressources dans le canal d’interférence acoustique sous-marin*. PhD thesis, Ecole nationale supérieure Mines-Télécom Atlantique.
- [Pottier et al., 2016] Pottier, A., Socheleau, F.-X., and Laot, C. (2016). Distributed power allocation strategy in shallow water acoustic interference channels. In *2016 IEEE 17th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 1–5. IEEE.
- [Pottier et al., 2017] Pottier, A., Socheleau, F.-X., and Laot, C. (2017). Robust noncooperative spectrum sharing game in underwater acoustic interference channels. *IEEE Journal of Oceanic Engineering*, 42(4) :1019–1034.
- [Qarabaqi and Stojanovic, 2013] Qarabaqi, P. and Stojanovic, M. (2013). Statistical characterization and computationally efficient modeling of a class of underwater acoustic communication channels. *IEEE Journal of Oceanic Engineering*, 38(4) :701–717.
- [Rahman et al., 2025] Rahman, W. U., Gang, Q., Zhou, F., Tahir, M., Ali, W., Adil, M., Zong Xin, S., and Khattak, M. I. (2025). Energy-efficient mac protocol for underwater sensor networks using csma/ca, tdma, and actor-critic reinforcement learning (ac-rl) fusion. In *Acoustics*, volume 7, page 39. MDPI.
- [Rathi and Gupta, 2020] Rathi, R. and Gupta, N. (2020). Device to device communication using stackelberg game theory approach. In *2020 Research, Innovation, Knowledge Management and Technology Application for Business Sustainability (INBUSH)*, pages 100–103.
- [Ribeiro Junior et al., 2022] Ribeiro Junior, A., Rocha Filho, G., Guidoni, D., Grande, R., Sampaio, S., and Meneguette, R. (2022). A shapley value-based strategy for resource allocation in vehicular clouds. In *Proceedings of the 2022 Conference on Resource Allocation*.
- [Rosenberger et al., 2021] Rosenberger, J., Urlaub, M., and Schramm, D. (2021). Multi-agent reinforcement learning for intelligent resource allocation in iiot networks. In *2021 IEEE Global Conference on Artificial Intelligence and Internet of Things (GCAIoT)*, pages 118–119. IEEE.
- [Saad et al., 2008] Saad, W., Han, Z., Debbah, M., and Hjorungnes, A. (2008). A distributed merge and split algorithm for fair cooperation in wireless networks. In *ICC Workshops - 2008 IEEE International Conference on Communications Workshops*, pages 311–315. IEEE.
- [Saad et al., 2009] Saad, W., Han, Z., Debbah, M., Hjorungnes, A., and Başar, T. (2009). Coalitional game theory for communication networks. *IEEE Signal Processing Magazine*, 26(5) :77–97.
- [Schirripa Spagnolo et al., 2020] Schirripa Spagnolo, G., Cozzella, L., and Leccese, F. (2020). Underwater optical wireless communications : Overview. *Sensors*, 20(8) :2261.

- [Scutari et al., 2008] Scutari, G., Palomar, D. P., and Barbarossa, S. (2008). Optimal linear precoding strategies for wideband non-cooperative systems based on game theory—part ii : Algorithms. *IEEE Transactions on Signal Processing*, 56(3) :1250–1267.
- [Shannon, 1949] Shannon, C. (1949). Communication in the presence of noise. *Proceedings of the IRE*, 37(1) :10–21.
- [Sharma and Singh, 2020] Sharma, S. and Singh, B. (2020). Coalition game-based strategy for resource allocation and transmit power control in d2d communication. *National Academy Science Letters*, 43 :317–319.
- [Sharma and Teneketzis, 2011] Sharma, S. and Teneketzis, D. (2011). A game-theoretic approach to decentralized optimal power allocation for cellular networks. *Telecommunication Systems*, 47 :65–80.
- [Shherbakov et al., 2020] Shherbakov, A. V., Petruhin, G. D., Miroshnikova, N. E., and Titovets, P. (2020). Estimation of underwater optical communication link operating distance. *T-Comm*, 14(3) :54–60.
- [Shi et al., 2015a] Shi, Y., Zhu, G., Lin, S., Ding, J., and Xu, S. (2015a). Coalitional game theory for cooperative interference management in femtocell networks. *Mathematical Problems in Engineering*, 2015 :767649.
- [Shi et al., 2015b] Shi, Y., Zhu, G., Lin, S., and Xu, S. (2015b). Rssi-based dynamic coalition formation for cooperative interference management in femtocell networks. In *Proceedings of the 2015 International Wireless Communications and Mobile Computing Conference (IWCMC)*, pages 1400–1405, Dubrovnik, Croatia. IEEE.
- [Shoham and Leyton-Brown, 2008] Shoham, Y. and Leyton-Brown, K. (2008). *Multiagent systems : Algorithmic, game-theoretic, and logical foundations*. Cambridge University Press.
- [Soni et al., 2021] Soni, V., Sharma, A., and Singh, V. (2021). A critical review on nature inspired optimization algorithms. In *IOP Conference Series : Materials Science and Engineering*, volume 1099, page 012055. IOP Publishing.
- [Staddon, 2020] Staddon, J. (2020). The dynamics of behavior : Review of sutton and barto : Reinforcement learning : An introduction . *Journal of the Experimental Analysis of Behavior*, 113(2).
- [Stojanovic and Beaujean, 2016] Stojanovic, M. and Beaujean, P.-P. J. (2016). *Acoustic Communication*, pages 359–386. Springer International Publishing.
- [Stojanovic and Preisig, 2009] Stojanovic, M. and Preisig, J. (2009). Underwater acoustic communication channels : Propagation models and statistical characterization. *IEEE communications magazine*, 47(1) :84–89.
- [Sun et al., 2019] Sun, Y., Wang, F., and Liu, Z. (2019). Coalition formation game for resource allocation in d2d uplink underlaying cellular networks. *IEEE Communications Letters*, 23(5) :888–891.
- [Tang et al., 2021] Tang, C., Wang, T., Pan, C., Song, J., Zhang, C., and Yang, F. (2021). Spectrum allocation of multi-priority operators based on repeated game for future mobile communication. In *2021 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, pages 1–5.
- [Taywade, 2021] Taywade, K. (2021). Multi-agent reinforcement learning for decentralized coalition formation games. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 15738–15739. AAAI Press.

- [Tu et al., 2011] Tu, K., Duman, T. M., Stojanovic, M., and Proakis, J. G. (2011). Odfma for underwater acoustic communications. In *2011 49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 633–638.
- [van Walree, 2011] van Walree, P. (2011). Channel sounding for acoustic communications : techniques and shallow-water examples. *Forsvarets Forskningsinstitut, FFI-rapport*, 7 :2011.
- [Vegas Segura, 2020] Vegas Segura, A. B. (2020). The bargaining problem : Nash and other solutions.
- [Vengerov, 2007] Vengerov, D. (2007). A reinforcement learning approach to dynamic resource allocation. *Engineering Applications of Artificial Intelligence*, 20(3) :383–390.
- [Viterbi, 1995] Viterbi, A. J. (1995). *CDMA : principles of spread spectrum communication*. Addison Wesley Longman Publishing Co., Inc.
- [Von Neumann and Morgenstern, 1944] Von Neumann, J. and Morgenstern, O. (1944). Theory of games and economic behavior : 60th anniversary commemorative edition. In *Theory of games and economic behavior*. Princeton university press.
- [Wang et al., 2019] Wang, H., Li, Y., and Qian, J. (2019). Self-adaptive resource allocation in underwater acoustic interference channel : A reinforcement learning approach. *IEEE Internet of Things Journal*, 7(4) :2816–2827.
- [Wang et al., 2020] Wang, J., Zhang, Y., Kim, T.-K., and Gu, Y. (2020). Shapley q-value : A local reward approach to solve global reward games. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pages 7285–7292.
- [Winter, 2002] Winter, E. (2002). The shapley value. In *Handbook of Game Theory with Economic Applications*, volume 3, pages 2025–2054. Elsevier.
- [Wu et al., 2015] Wu, D., Cai, Y., Hu, R. Q., and Qian, Y. (2015). Dynamic distributed resource sharing for mobile d2d communications. *IEEE Transactions on Wireless Communications*, 14(10) :5417–5429.
- [Wynn et al., 2014] Wynn, R. B., Huvenne, V. A., Le Bas, T. P., Murton, B. J., Connelly, D. P., Bett, B. J., Ruhl, H. A., Morris, K. J., Peakall, J., Parsons, D. R., et al. (2014). Autonomous underwater vehicles (auvs) : Their past, present and future contributions to the advancement of marine geoscience. *Marine geology*, 352 :451–468.
- [Yang, 2011] Yang, T. (2011). Characteristics of underwater acoustic communication channels in shallow water. In *OCEANS 2011 IEEE - Spain*, pages 1–8.
- [Ye et al., 2024] Ye, F., Xu, H., and Gao, J. (2024). Relay selection in underwater acoustic sensor networks for qos-based cooperative communication using game theory. *Computer Communications*, 219 :104–115.
- [Zhai et al., 2012] Zhai, L., Ji, H., Li, X., and Tang, Y. (2012). Coalition graph game for joint relay selection and resource allocation in cooperative cognitive radio networks. In *2012 IEEE Global Communications Conference (GLOBECOM)*, pages 1818–1823. IEEE.
- [Zhang et al., 2015] Zhang, H., Jiang, C., Beaulieu, N. C., Chu, X., Wang, X., and Quek, T. Q. S. (2015). Resource allocation for cognitive small cell networks : A cooperative bargaining game theoretic approach. *IEEE Transactions on Wireless Communications*, 14(6) :3481–3493.
- [Zhang et al., 2013] Zhang, X., Han, X., Yin, J., and Sheng, X.-l. (2013). Study on doppler effects estimate in underwater acoustic communication. In *Proceedings of Meetings on Acoustics*, volume 19. AIP Publishing.

- [Zhou et al., 2019] Zhou, J., Jiang, H., Wu, P., and Chen, Q. (2019). Study of propagation channel characteristics for underwater acoustic communication environments. *IEEE Access*, 7 :79438–79445.
- [Zhou and Wang, 2014] Zhou, S. and Wang, Z. (2014). *OFDM for Underwater Acoustic Communications*. John Wiley & Sons, Hoboken, NJ, USA.
- [Štrumbelj and Kononenko, 2014] Štrumbelj, E. and Kononenko, I. (2014). Explaining prediction models and individual predictions with feature contributions. *Knowledge and Information Systems*, 41 :647–665.

Titre : Stratégies d'allocation de ressources basées sur la théorie des jeux et l'apprentissage par renforcement pour les réseaux acoustiques sous-marins

Mots clés : communications acoustiques sous-marines, allocation de ressources, théorie des jeux non coopératif, théorie des jeux coopératifs, apprentissage par renforcement, Q-learning

Résumé : Le milieu sous-marin constitue un environnement acoustique partagé entre une grande diversité d'acteurs, naturels (mammifères marins, faune benthique) et artificiels (modems de communication, sonars, capteurs, systèmes de télémétrie), contrairement aux radiocommunications terrestres, où le spectre est strictement encadré par des régulations internationales, l'environnement acoustique sous-marins demeure non coordonné, non normalisé, et soumis à une pression croissante du fait des activités anthropiques (océanographie, défense, prospection, exploitation des ressources marines, etc.). Dans ce contexte, la cohabitation non réglementée de nombreuses sources concurrentes peut conduire à un masquage mutuel des signaux, compromettant sévèrement la fiabilité des communications. Cette situation met en lumière une problématique majeure : le partage efficace et équitable des ressources acoustiques limitées (temps, fréquence, puissance) entre les différents utilisateurs du canal. En effet, sans mécanismes adaptés pour gérer l'accès multiple au canal, les interférences et la concurrence entre les utilisateurs peuvent rapidement dégrader les performances globales du système.

Cette thèse vise à résoudre le problème du partage des ressources dans les environnements acoustiques sous-marins en proposant des stratégies d'accès multiple à la fois robustes, adaptatives et décentralisées. Les solutions envisagées offrent la capacité aux systèmes de communication d'adapter automatiquement et intelligemment leurs paramètres de transmission selon l'état du canal et l'activité des utilisateurs, tout en s'affranchissant de toute coordination centralisée. Pour atteindre cet objectif, la thèse adopte une approche interdisciplinaire s'appuyant sur la théorie des jeux coopératifs et l'apprentissage par renforcement. Plutôt que de considérer les utilisateurs comme des compétiteurs opportunistes, elle explore des modèles de coopération entre agents acoustiques, où chaque nœud vise à maximiser la performance globale du réseau et à réduire les interférences grâce à des décisions locales mais coordonnées. Ce paradigme, inspiré des réseaux cognitifs terrestres, s'avère particulièrement adapté au contexte sous-marin, caractérisé par un environnement dynamique, bruité et à faible bande passante, où la gestion centralisée demeure difficile à mettre en œuvre.

Title : Resource allocation strategies based on game theory and reinforcement learning for underwater acoustic networks

Keywords : Underwater acoustic communications, resource allocation, non-cooperative game theory, cooperative game theory, reinforcement learning, Q-Learning

Abstract : The underwater environment is an acoustic medium shared by a wide range of actors, both natural (e.g., marine mammals, benthic fauna) and artificial (e.g., communication modems, sonars, sensors, telemetry systems), unlike terrestrial radio communications, where spectrum usage is strictly regulated by international standards, the underwater acoustic environment remains largely uncoordinated and unstandardized. It is also facing increasing pressure due to the expansion of human activities such as oceanographic research, military operations, and resource exploitation. In this context, the unregulated coexistence of numerous competing sources can result in mutual signal masking, severely degrading the reliability of underwater communications. This situation highlights a major challenge: the efficient and fair sharing of limited acoustic resources (time, frequency, power) among multiple users of the channel so without appropriate mechanisms to handle multiple access, the interference and the competition between users can degrade the overall performance of the system.

This thesis aims to address the resource allocation problem in underwater acoustic environments through the development of robust, adaptive, and decentralized multiple-access strategies. The proposed approaches allow communication systems to intelligently and autonomously optimize their transmission parameters according to channel dynamics and network activity, eliminating the need for centralized control. To achieve this goal, the thesis adopts an interdisciplinary approach based on cooperative game theory and reinforcement learning. Rather than considering users as opportunistic competitors, she explores cooperation models between acoustic agents, where each node aims to maximize the overall performance of the network and reduce interference through local but coordinated decisions. This paradigm, inspired by terrestrial cognitive networks, is particularly adapted to the underwater context, characterized by a dynamic, noisy and low-bandwidth environment, where centralized management remains difficult to implement.